Volume 5, Issue 5
October 2016

ITEE Journal
Information Technology & Electrical Engineering

ISSN: - 2306-708X

Video Processing Based Decision Making and Learning in Lawn Tennis

[1] Utkarsh Shukla [2] Swati Srivastava

[1] Naraina Vidyapeeth Engineering Management Institute, Kanpur, INDIA

2 Naraina Vidyapeeth Engineering Management Institute, Kanpur, INDIA

E-mail: [1]utkarshshukla2013@gmail.com,

ABSTRACT

Object tracking, in the last couple of years, has evolved as an area of one of the most active research in computer vision, particularly because of the developing significance of visual surveillance for the purpose of security. Object tracking could be defined as a general framework that comprises various distinctive computer vision projects that aims to track, recognize, classify and characterize objects of interest from sequences of image, and on the following level to comprehend and explain the behaviour of these objects. The final objective in designing brilliant object tracking systems is to take place of the current passive surveillance and to uproot, or if nothing else, minimize the requirement for an individual observer to control and analyze the visual data. In the present day environment the use of internet has grown tremendously. The use of internet is everywhere, in entertainment, learning e-commerce etc. This paper, presents a lawn tennis system for foreground, background and object detection system. The system can be used for lawn tennis coaching, player and ball tracking. Thus the presented system can also be used for line calls and tracing of Aces. The presented system is also able to detect shadows which are very effective in TV referrals.

Keywords: *Foreground, Background, shadow and object detection*

1. INTRODUCTION

Video and Image processing is now become very important n sports. This is due to the fact that now a days in various sports videos are used to make critical decisions. In this work, it is shown how video and image processing can be used in Lawn tennis in decision making and providing training to the learner. The design of video based training and use of simulation cannot be easy. The one major issue is of selecting exact video clips; more generally selecting video clips by using suitable camera angle is very difficult. Some of the camera angels which are generally used in previous and present research are aerial or broadcast angle. The broadcast angle is more popularly used because it is easily available through television. The difficulty of the scenarios also needs to be considered when selecting clips. The inclusion of a subjective difficulty ratings should be included, or an item analysis post testing, to remove any clips that no participant scored on below chance. One of the prime issues in the designing part of decision-making programs on video-based is making the athlete engaged and feel like they are making decisions in a real match position. For detecting the moving objects in the scene like in video surveillance [2, 3], optical motion capture [4, 5-6] and multimedia [7, 8-10], the background modelling is used very frequently. The easiest way to model the background is to obtain a background image which doesn't contain any moving object.

In some cases the background does not exist and can be changed under crucial situations like illumination change, objects being introduced or removed from the scene.

Let us suppose that we have sequence of images; $1 \le t \le T$, Now the scene can be divided in to two parts one is foreground (k=1) and other is background (k=2). The terms foreground and background are used loosely; the foreground layer contains regions occluding the background. On the other hand in foreground layer multiple moving objects that do not occlude each other appears frequently. For stable situations the baseline mode is designed [6, 11]. This is still used in cameras where no light changing and no shadow are formed. Basically this is based on two techniques these are background registration technique and change detection technique. In this change detection algorithm, the change detection mask, here is not simply generated from frame difference between current and earlier frame, but also from the frame difference between the background frame and current frame here background frame is generated from background registration technique. Since the background used here is stationary so it is well behaved and reliable in comparison to previous frame. The figure 2 shows a block diagram of baseline mode. 5 parts in which the baseline mode is divided in are as follows; Frame Difference, Background Registration, Background Difference, Object Detection, and Post processing, concept explained in [6, 11].

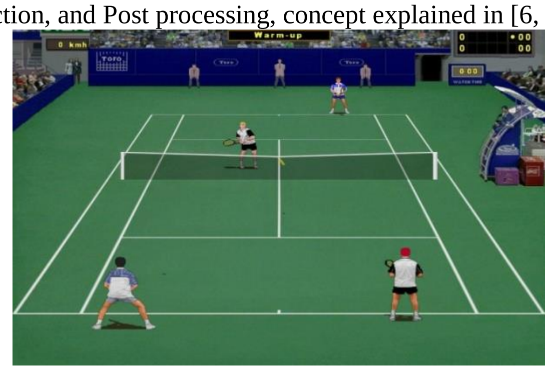


Fig. 1 Pictorial representation of Lawn Tennis groun

2. QUALITY OF SERVICE (QoS)

Frame Difference

Under this heading calculation shows the frame difference between the current frame and previous frame, which is to be stored in to the frame buffer. It can be presented as,

$$F_D(x,y,t) = \left|I(x,y,t) - I(x,y,t-1)\right| \quad (1)$$

$$F_{DM}(x,y,t) = \begin{cases} 1 & if \quad F_D \geq T_h \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

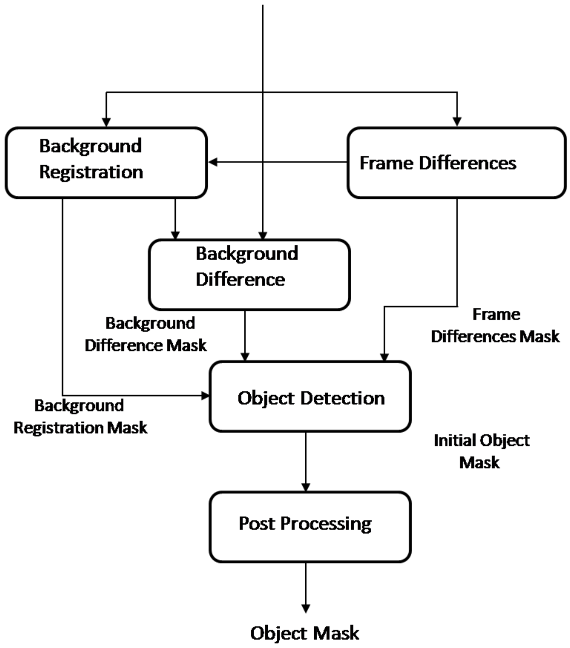


Figure.2. Block diagram of the baseline mode

In equations (1) and (2) *I* represent frame data, F_D shows frame difference and F_{DM} represents Frame *Difference Mask*. Here one thing is important that the pixels which belong to the F_{DM} are in the category of moving pixels. Here one thing is noticed that the parameters which are required in this case are set in advanced.

Background Registration

This is used to extract the background information from video sequences. The pixels which are not moving for a long time extent are considered to be as reliable background pixels. The procedures of background registration are discussed under the following equations.

$$S_I(x,y,t) = \begin{cases} S_I(x,y,t-1)+1 & \text{if} \quad F_{DM}=0 \\ 0 & \text{if} \quad F_{DM}=1 \end{cases} \quad (3)$$

If $S_I(x,y,t) = F_{th}(x,y,t-1)$

$$B_G(x,y,t) = S_I(x,y,t-1)+1 \quad (4)$$

If $S_I(x,y,t) > F_{th}(x,y,t-1)$

$$B_G(x,y,t-1)(1-\alpha) + I(x,y,t)\alpha \quad (5)$$

Else

$$B_G(x,y,t) = B_G(x,y,t-1) \quad (6)$$

$$B_I(x,y,t) = \begin{cases} 1 & \text{if} \quad S_I(x,y,t) = F_{th}(x,y,t\text{-}1) \\ B_I(x,y,t-1) & \text{else} \end{cases} \quad (7)$$

If $S_I(x,y,t) > F_{th}(x,y,t-1)$

$$F_{th}(x,y,t) = S_I(x,y,t-1) \quad (8)$$

Else

$$F_{th}(x,y,t) = F_{th}(x,y,t-1) \quad (9)$$

Here in the above equations S_I represents the stationary index, B_I is background indicator, B_G shows background information , F_{th} shows threshold and finally "α" represent parameter indicator. Initially the all values of S_I , B_I , and B_G are set to be "0", the value of F_{th} set to "5". In the case if the pixel is in the background region then the stationary index records the possibility. In the case if the value of S_I goes high then the corresponding possibility also goes high; otherwise on the other hand the possibility goes low. This shows that the chance of pixels belong to the background is high, In the case when a pixel "Keeps stationary" for many successive frames. When the possibility is high enough, the present pixel information of the position is registered into the background buffer B_G, which is the main concept as in [5]. For checking that the background information of present position is ready or not B_I is used. Here F_{th} used in the equation is different from the original one. The F_{th} is used as constant in the original algorithm.

But in the case if a foreground object stays stationary for a long time, that means the value of S_I is equal to the F_{th} , hence the background information for that point will be set to pixel information of foreground object. In this paper F_{th} is used as a dynamic variable. In the case if the value of S_I is larger than the value of F_{th} and , the value goes on expanding in respect to the time , then F_{th} will be updated according to the value of S_I. This concept is explained in [6]. The value of S_I used in the background pixel should be larger than the value which is used in the foreground pixel. This happens because of background is much stable in comparison to the foreground. By using this method, the pixel information of foreground is not registered in to B_G. The concept which is explained in [4] is quite different from the original one, too. For updating the background with respect to change in light we uses α.

Background Difference

The concept of background difference is similar to the frame difference. This concept is broadly discussed in [11]. But the difference between the background difference and the frame difference is that the, previous frame is substituted by background frame. After background difference, background difference mask is generated, which is one of the other change detection masks. The following equations show the operations of background difference.

$$B_D(x,y,t) = I(x,y,t) - B_G(x,y,t-1)$$

$$B_{DM}(x,y,t) = \begin{cases} 1 & if \quad \text{B}_D \geq T_h \\ 0 & \text{otherwise} \end{cases} \quad (9)$$

Volume 5, Issue 5
October 2016

ITEE Journal
Information Technology & Electrical Engineering

ISSN: - 2306-708X

Here B_D represents the background difference, B_G shows the background frame and B_{DM} represents Background Difference Mask.

Object Detection:

For producing Initial Object Mask, both of the BDM and FDM are used as input in to object detection. The whole process of object detection is represented by following equations.

$$I_{OM}(x,y,t)=\begin{cases}B_{DM}(x,y,t) & if \quad B_I(x,y,t)=1\\ F_{DM}(x,y,t) & \text{otherwise}\end{cases} \qquad (10)$$

Here in the equation I_{OM} represents whether pixels belongs to foreground or not. In the case if the value of I_{OM} is equal to 1 then the pixel at position (x, y) belongs to foreground position otherwise it belongs to background.

Background Subtraction

Object recognition can be acquired with the formation of a representation of the scene known as the background model and after this discovering deviation from the model for all reaching frames. With any noticeable variation in a region of image from the background model gives the signification of a object which is not stationary. For more processing the pixels forming the regions in which changes are going on are marked. In general for obtaining connected regions correspond to objects, a connected component algorithm is applied. The process discussed here is called the background subtraction process and is explained in [11]. At the moment we derive the background model, for every pixel (x, y) in the input frame, the likelihood of its colour coming from N(μ(x, y), σ(x, y)) is evaluated, and the pixels that undergoes deviation from the background model are labelled as the foreground pixels. The classification of pixel is made on the basis of the fact that whether the matched distribution provides the representation of the background process. Figure 3 illustrates the moving regions, which are discovered with the help of this approach, along with the background models.

In table I author discuss criteria for object detection. In this table for first two cases the background information is not yet available; hence for separating object from background, the frame difference information is used as criteria for this.

Table 1: Criterion of Object Detection

Situation	FDM	BDM	BI	IOM
Stationary	0	-	0	0
Moving	1	-	0	1
Background (BG)	0	0	1	0
Moving object	1	1	1	1
Still object	0	1	1	1
Uncovered BG	1	0	1	0

The criteria used in the decision table, for cases 3 to 6 are background difference this is why because the background information exists. The pixel is the part of moving object, in the case if the background difference and the frame difference are significant. On the other hand in some another cases if these differences (frame and background) are insignificant the pixel is not be the part of object mask. This is why for third and fourth cases in decision table, the result is exactly same as the result when we uses only, the frame difference for change detection. The frame difference based change detection cannot handle correctly this situation shows in the cases 5 and 6 of decision table, here one thing which is noticed is that; in this situation the background difference works properly. One of the issues that creates the confusion in the conventional change detector (only frame difference issued) is that the object may stop moving temporarily or move very slowly. In all cases, if we are focused only on the frame difference then the motion information disappeared. On the other hand in the case of background difference information, we noticed that the pixels which are used frequently; belongs to the object region and can be integrated in the object mask. In case 6, both regions (background and moving object) have considerable luminance change, in the case if only the frame difference is given then this is very difficult to uncovered background from object. In this algorithm, the open background region is handled accurately because we know that this region matches the background information even though frame difference suggests significant motion.

3. FRAMEWORK FOR THE SIMULATION

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

ITEE Journal
Information Technology & Electrical Engineering

ISSN: - 2306-708X

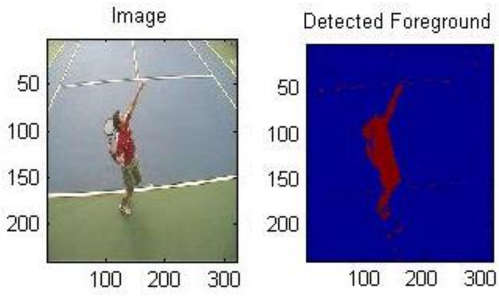
Frame 4

Frame 5

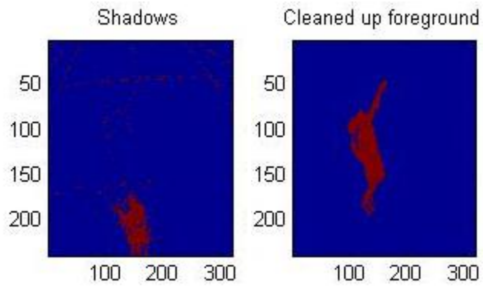
Frame 6

Fig. 3: Frame by Frame representation

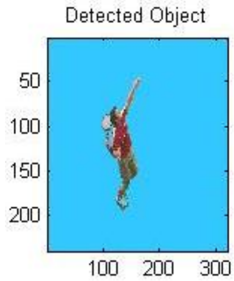
In the first experiment all six frame were used in the training and frame 6 was under investigation as marked as image is figure 4 (a). Detected foreground image is shown in figure 4(b).



(a) (b)



(c) (d)

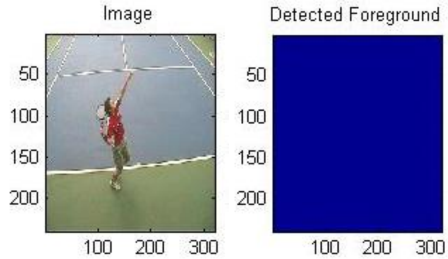


(e)

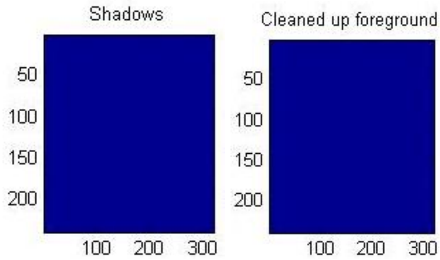
Fig 4 Foreground, shadow and object separation with 6th frame under investigation with all six frames in training

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).

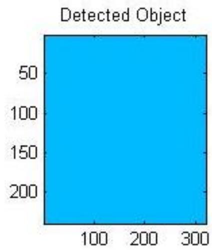
In the second experiment first four frames were used in the training and frame 6 was under investigation as marked as image is figure 4 (a). Detected foreground image is shown in figure 4(b).



(a) (b)



(c) (d)



(e)

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

ITEE Journal
Information Technology & Electrical Engineering

ISSN: - 2306-708X

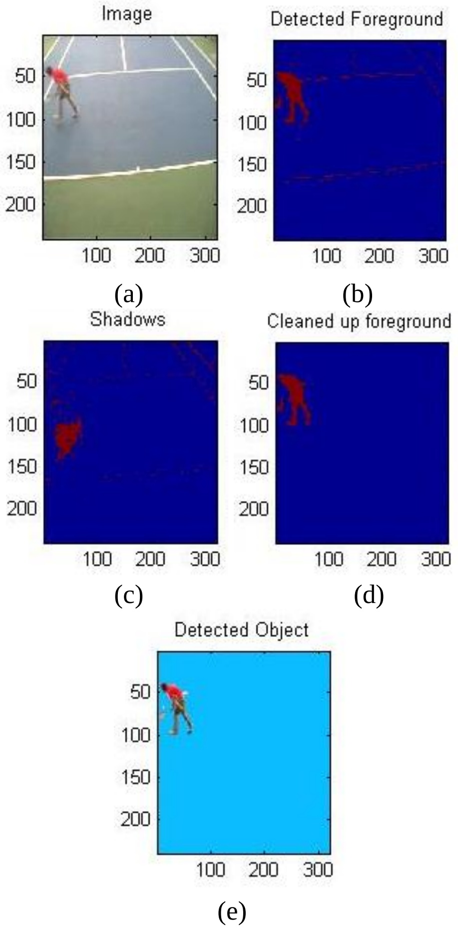


Fig 6 Foreground, shadow and object separation with 5th frame under investigation with first five frames in training

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).

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).

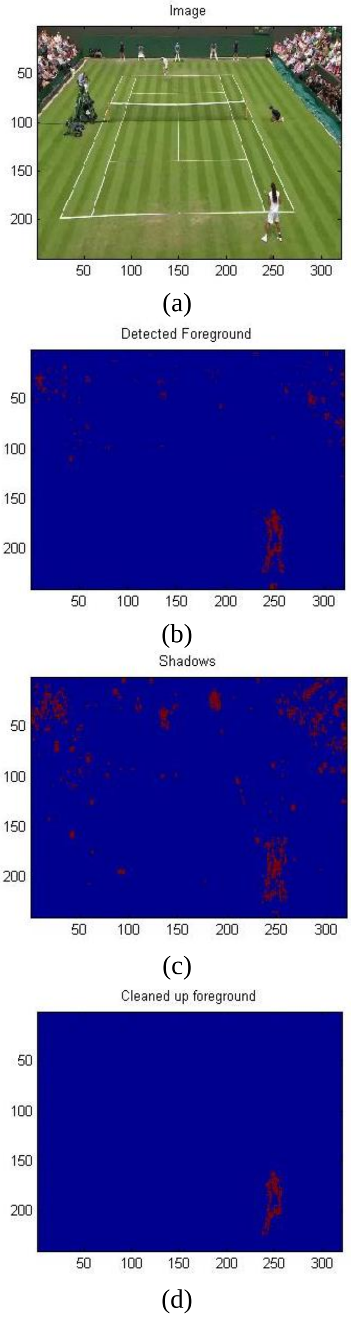
The shadow of the object is shown in figure 6(c) and image with clear foreground is shown in figure 6(d), however, the detected object is shown in figure 6(e).

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.

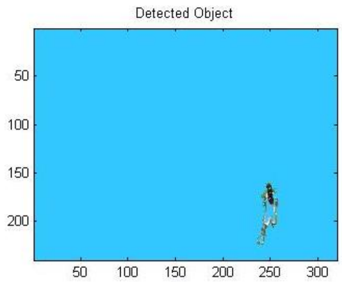
Fig.7 Snapshot of video clip

Fig.8 Snapshot of generated frames



ITEE Journal
Information Technology & Electrical Engineering

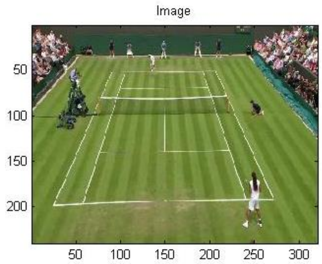
ISSN: - 2306-708X



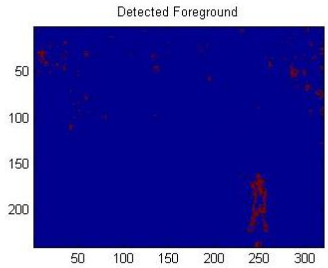
(e)

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

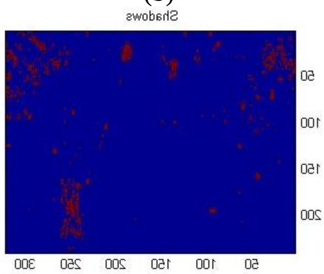
In figure 9 results are obtained while considering number of layers as 5 and Euclidean distance as 3. Figures show, input image, detected foreground, shadow, cleared foreground and detected object. Dustin Brown was correctly detected.



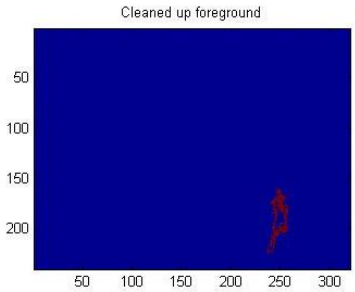
(a)



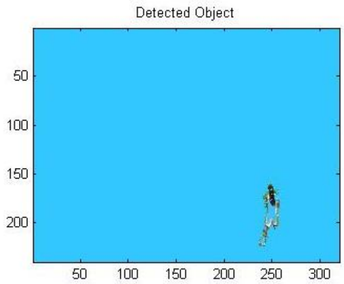
(b)



(c)



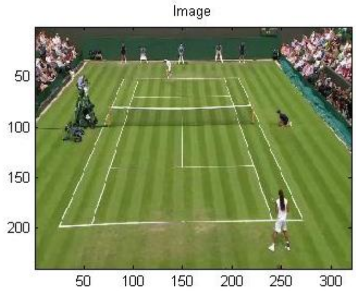
(d)



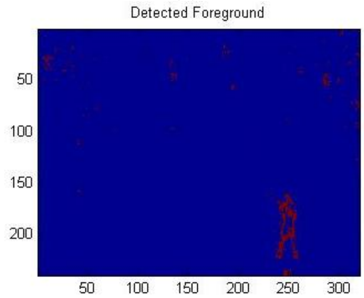
(e)

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

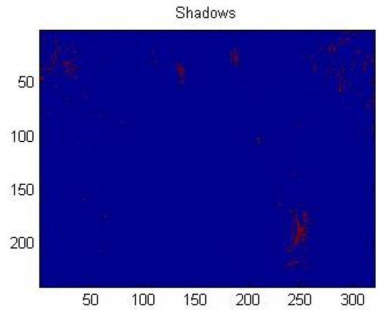
In figure 10 results are obtained while considering number of layers as 5 and Euclidean distance as 5. Figures show, input image, detected foreground, shadow, cleared foreground and detected object. Dustin Brown was correctly detected. Most of the images look similar in figure 9 and 10 except shadow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



(a)



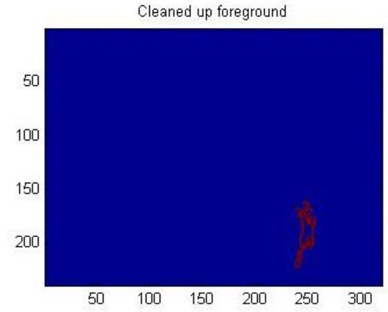
(b)



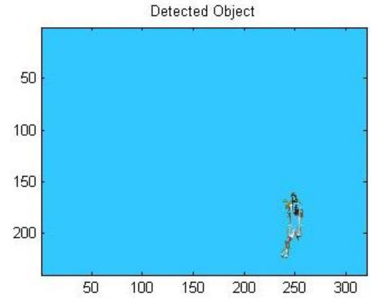
(c)

ITEE Journal
Information Technology & Electrical Engineering

ISSN: - 2306-708X



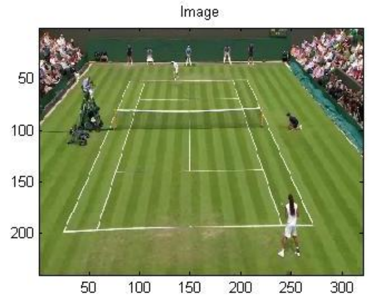
(d)



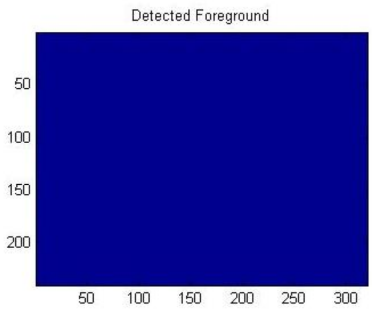
(e)

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

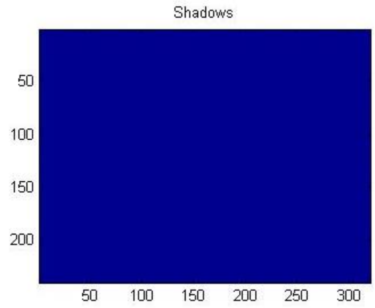
In figure 11 results are obtained while considering number of layers as 5 and Euclidean distance as 7. Figures show, input image, detected foreground, shadow, cleared foreground and detected object. Dustin Brown was correctly detected. Most of the images look similar in figure 5.8 and 5.9 except shadow image which differ in two cases. Thus it can be inferred that Euclidean distance affects shadow image.



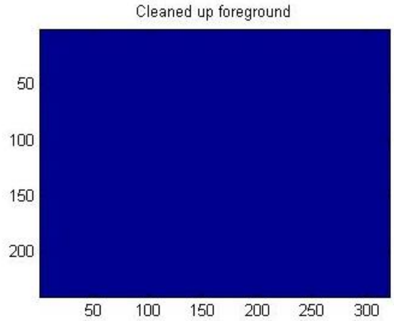
(a)



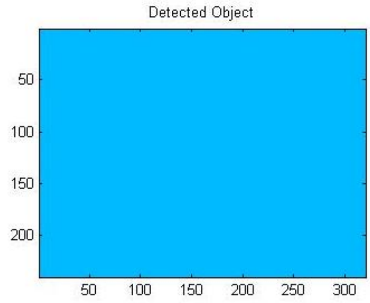
(b)



(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.

4. CONCLUSIONS

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. It has been discovered from the results that to detect object the minimum number of required layers are 5. In the object detection, Euclidean distance place an important role and shadow of an image is heavily depends on Euclidean distance.

REFRENCES

[1] A. Adams, J. Baek, and M. A. Davis. Fast high-dimensional filtering using the permutohedral lattice. Comput. Graph. Forum, 29(2):753–762, 2010.
[2] X. Bai, J. Wang, and G. Sapiro. Dynamic color flow: A motionadaptive color model for object segmentation in video. In ECCV, pages V:617–630, 2010.
[3] X. Bai, J. Wang, D. Simons, and G. Sapiro. Video snapcut: Robust video object cutout using localized classifiers. In SIGGRAPH, pages 70:1–70:11, 2009.
[4] Y. Boykov, O. Veksler, and R. Zabih. Fast approximate energy minimization via graph cuts. PAMI, 23(11):1222–1239, Nov 2001.
[5] A. Buades, B. Coll, and J. Morel. A non-local algorithm for image denoising. In CVPR, pages 2:60–65, 2005.
[6] D. J. Butler, J. Wulff, G. B. Stanley, and M. J. Black. A naturalistic open source movie for optical flow evaluation. In ECCV, IV, pages 611–625, 2012.
[7] A. Criminisi, G. Cross, A. Blake, and V. Kolmogorov. Bilayer segmentation of live video. In CVPR, pages 1:53–60, 2006.
[8] T. Darrell and A. Pentland. Cooperative robust estimation using layers of support. PAMI, 17(5):474–487, 1995.
[9] P. Kr¨ahenb¨uhl and V. Koltun. Efficient nonlocal regularization for optical flow. In ECCV, pages I:356–369, 2012.
[10] F. Kschischang, B. Frey, and H.-A. Loeliger. Factor graphs and the sum-product algorithm. IEEE Trans. IT, 47(2):498–519, Feb 2001.
[11] M. Kumar, P. Torr, and A. Zisserman. Learning layered motion segmentations of video. IJCV, 76(3):301–319, March 2008.