International Journal of Research and Development in Applied Science and Engineering (IJRDASE)
ISSN: 2454-6844

An Improved Method for Foreground, Background and Object Identification

Utkarsh Shukla
Department of CSE
SRRT Mahila Polytechnic, Kanpur
utkarshshukla2013@gmail.com

***Abstract*--The concept of background difference is similar to the frame difference. 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. For producing Initial Object Mask, both of the BDM and FDM are used as input in to object detection. Object recognition can be achieved by building a representation of the scene called the background model and then finding deviations from the model for each incoming frame. Any significant change in an image region from the background model signifies a moving object.**

Keywords: Background Difference, Background Registration, Frame Differencing, Object Detection

1. Introduction:

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.

With the rising availability of video sensors and video processing hardware of high performance facilitates with great possibilities for handling numerous video understanding issues.

The need to evolve robust real-time video understanding processes is on peak which should be able in processing a lot of data feasible. We consider the motion detection problem in our thesis. An input video stream is assumed from any digital video camera or may be from a web cam. This video is converted into frames and then object detection takes place. This work, 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.

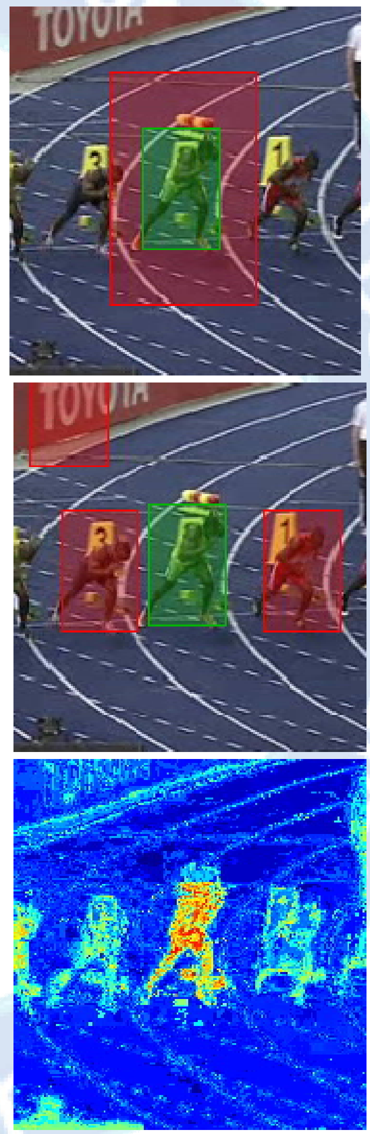


Fig. 1. Object Tracking (100 meter racing)

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.

2. Related Work:

Haritaoglu, *et al.* [1], model the background by representing each pixel with its most extreme value of intensity, least intensity value and difference of intensity values between back to back pixels. The disadvantages of a model like this are its susceptibility to the changes made in illumination.

An eigenspace model has been proposed by Oliver, et al. [2] for segmentation of moving object. In this technique, dimensionality of the space built from sample images is decreased by utilizing Principal Component Analysis (PCA).

It was claimed by them that, after the PCA application, the diminished space will represent just the static scene parts, yielding moving objects, in the case an image is expected on this space. In spite of having few achievements by the method in some applications, it cannot demonstrate dynamic scenes fully. Thus, it is not extremely suitable particularly for outdoor observation tasks.

Along with the mentioned statistical method Wren, et al. [3] proposed another one, which models all points in a scene utilizing a Gaussian distribution with an estimated mean intensity value. The disadvantage of the model is that it can just deal with unimodal distributions. Somewhat later on, in a convectional approach, a mixture of Gaussians is likewise proposed, rather than a single Gaussian [4].

Sample background images are used by Elgammal, *et al.* [5] for the estimation of the probability of the observation of values of pixel intensity in a nonparametric way with no assumption about the type of the background probability distribution. Truly, this theoretically method that is well established results numerous exact results under challenging conditions of outdoor.

The ultimate objective of a completely automated object tracking system is probably event recognition. Despite the fact that it is very critical and valuable to recognize an activity, it is difficult to characterize the motion type that is interesting and significant inside the sports context. Thus, there are numerous studies addressing diverse events types. Polana and Nelson [11] figure the optical flow fields between successive frames and sum up the vector magnitudes in regions of object to gain high dimensional component vectors that are utilized for recognition. Exercises are grouped by using the closest neighbour algorithm. To discover simple movement characteristics again attempted and in [13] proposes a "star" skeletonization strategy. The items are recognized by using background subtraction and then their boundaries are removed and a skeleton is created. The authors claim that skeletonization gives vital motion signs like posture of body and cyclic motion of skeleton segments, which thusly are used in finding human activities like walking or running [13].

Rather than making analysis of simplistic object motions, patterns of activity in time might also be observed. A state-based learning architecture was proposed in [2] with coupled hidden Markov models (CHMM), to model behaviours of object and communications between them. Object motion was represented by Johnson, *et al.* [14] using flow vectors, which include spatial location and instantaneous object velocity. Then, the trajectories are built as a grouping of flow vectors and a competitive learning network is adapted to model the probability density functions of flow vector sequences. In the similar way, [15] produce probabilistic models to describe the normal motion in the scene. The flow vectors are further quantized to get a prototype representation and trajectories are converted into prototype vector sequences. Thereafter, these sequences are evaluated using the probabilistic trajectory models.

A codebook of prototype representations was produced in [16] from input representations (*x, y, vx, vy,* size of object, binary mask) using on-line Vector Quantization (VQ). At that point, a co-occurrence matrix is characterized over the prototypes in the codebook and a hierarchical classifier is produced by making use of co-occurrence data. Lee, *et al.* [17] likewise work with prototype vectors and its objective is of the classification of both local and global trajectory points. Support Vector Machines are used by them for the detection of local point abnormality while the classification of global trajectories (sequences of vectors) is done by using HMMs. As a last step, a rule-based system consolidates local and global information to make the decision on the abnormality of the motion pattern [17].

3. Methodology:

The factor on which performance of an automated visual object tracking system extensively relies is on its capacity to detect moving objects in the observed environment.

A resulting action like tracking, identifying persons or analyzing the motion, needs an exact extraction of the foreground objects, making detection of a moving object in is an important system part.

The issue of detecting changes in a scene can be explained as: Same scene images are obtained in time by a static camera and the objective is to recognize changes between progressive frames. Pixels that have a critical difference in comparison to the prior ones are labelled as foreground pixels, though other pixels are marked as background, which results in a change mask. The pixels set in this change mask results the segmentation of the moving objects.

Keeping in mind the end goal to decide on whether a few regions in a frame are foreground or not, there should be a model for the background intensities. This model should also have the capacity to catch and store important background information.

Any change brought on by a new object, should be distinguished by this model, while non-stationary background regions, like leafs and branches of a tree or a flag waving in the wind, should be recognized as a background part. Many different methods are tested in this thesis to decide on their performance for such a problem of detection.

A. Frame Difference:

This is the simplest and easiest method for moving object detection. The model for the background is simply equal to the previous frame.

International Journal of Research and Development in Applied Science and Engineering (IJRDASE)
ISSN: 2454-6844

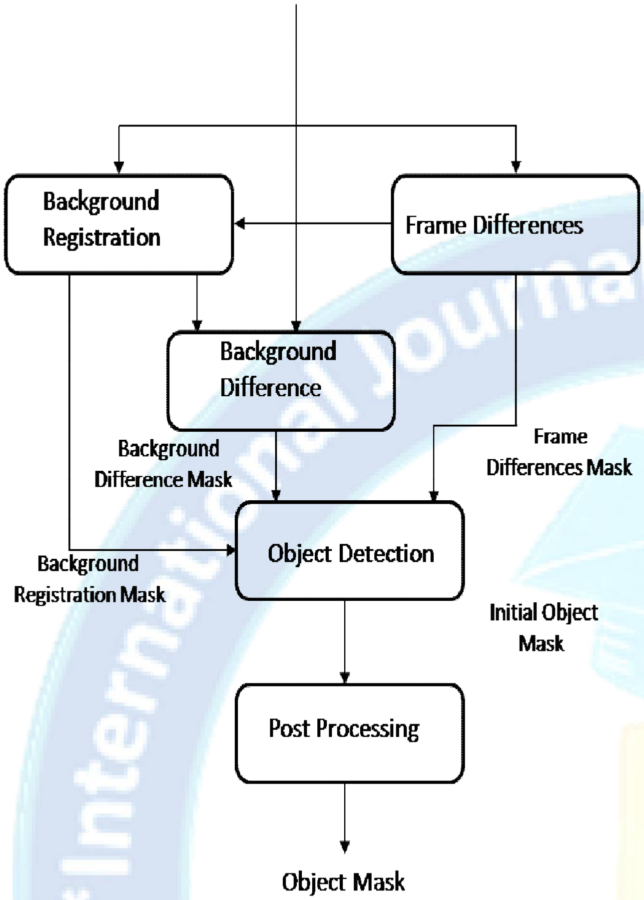


Fig. 2. Block diagram of the baseline mode

$$m(x,y,t)=\begin{cases}0 \text{ if } \left|I(x,y,t)-I(x,y,t-1)\right|<th\\ 1 \text{ if } \left|I(x,y,t)-I(x,y,t-1)\right|>th\end{cases} \quad (1)$$

In the above mentioned formula, $I(x,y,t)$ is the intensity at pixel location (x,y) at time t^{th} is the threshold value and $m(x,y,t)$ is the change mask acquired after thresholding. Rather than using the previous frame, a single frame, which exclude any moving objects, can also be utilized as a reference.

In spite of the fact that this method is very quick and has adjustment capacity to the changes in the scene, it has a generally low performance in dynamic scene conditions and its outcomes are exceptionally delicate to the threshold value, *th*.

B. 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.

C. Background Difference:

The concept of background difference is similar to the frame difference. 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.

D. Object Detection:

For producing Initial Object Mask, both of the BDM and FDM are used as input in to object detection.

E. Background Subtraction:

Object recognition can be achieved by building a representation of the scene called the background model and then finding deviations from the model for each incoming frame. Any significant change in an image region from the background model signifies a moving object.

4. Result and Discussion:

Let us suppose that we have sequence of images; $1 \leq t \leq T$, 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 previous 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 fig. 2 shows a block diagram of baseline mode. The baseline mode is divided in to five parts which are; Frame Difference, Background Registration, Background Difference, Object Detection, and Post processing.

Fig. 3. Pictorial representation of Lawn Tennis ground

International Journal of Research and Development in Applied Science and Engineering (IJRDASE)
ISSN: 2454-6844

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 Fig. 3.

Frame 1

Frame 2

Frame 3

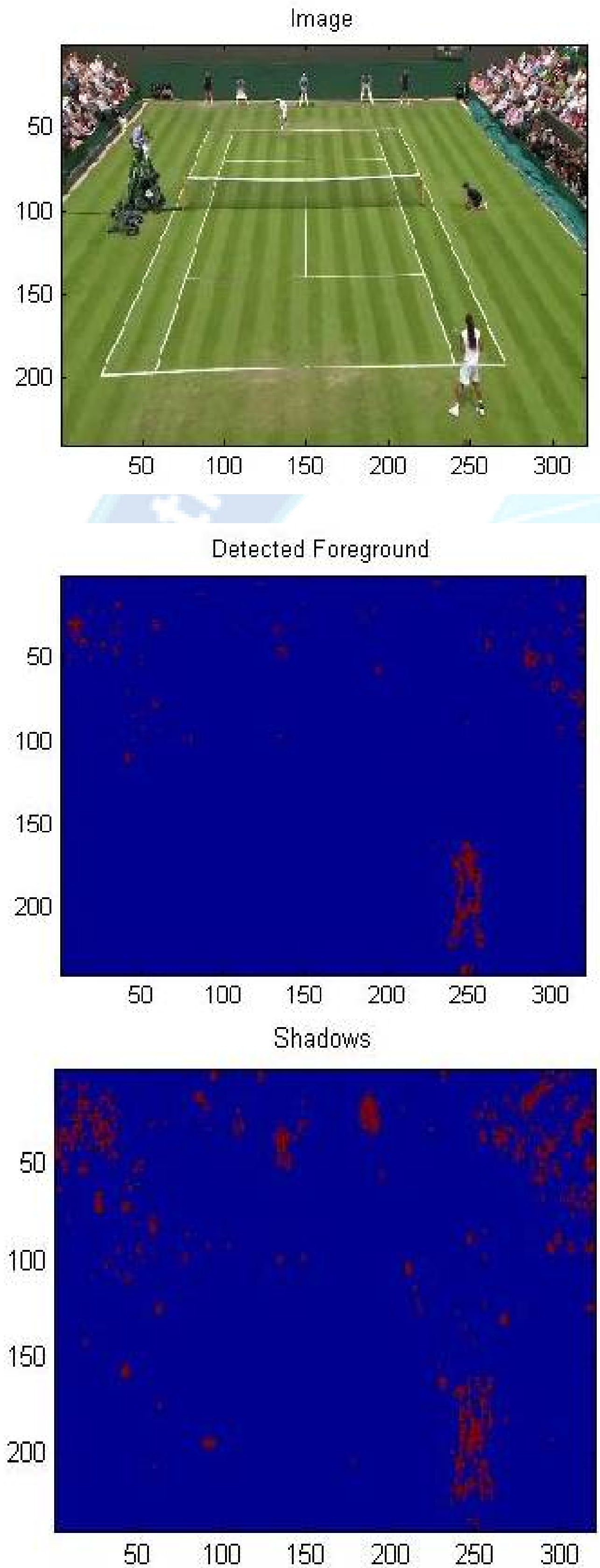
Frame 4

Frame 5

Frame 6

Fig. 3. Frame by Frame representation

International Journal of Research and Development in Applied Science and Engineering (IJRDASE)
ISSN: 2454-6844



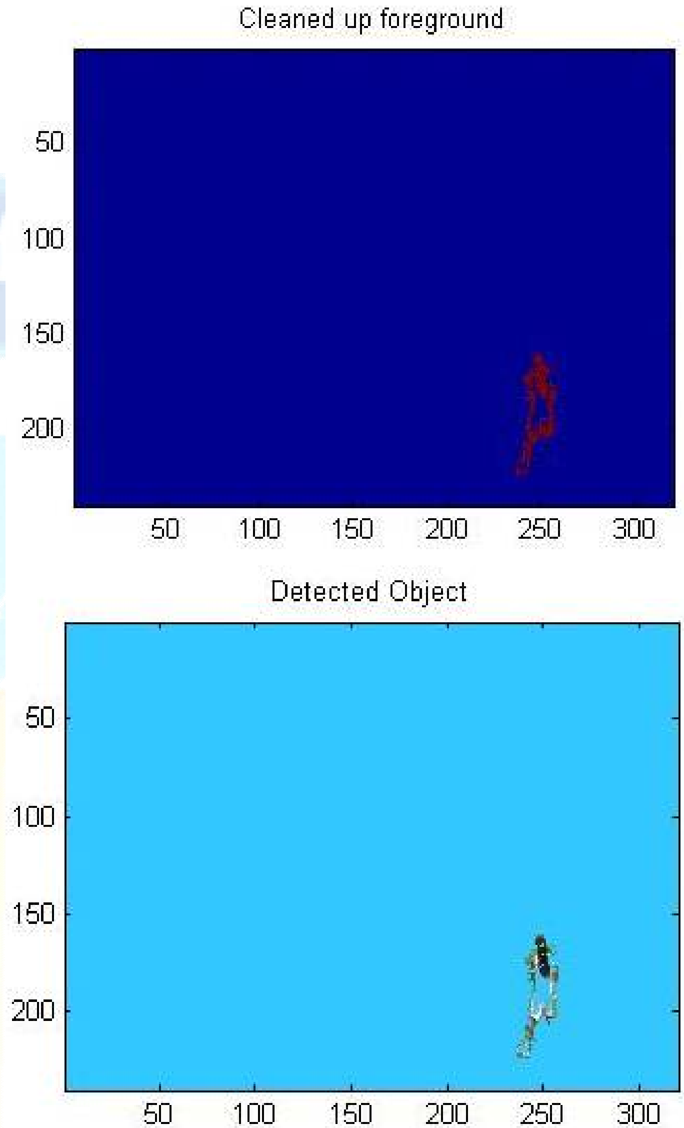


Fig. 4. Obtained pictures with number of layers as 5 and Euclidean distance 5

In fig. 4. 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.

This paper provides a methodology about how a mathematical can be used in players tracking in lawn tennis round. In the results it has been fund 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.

5. Conclusion:

This thesis 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 thesis discuses the baseline model for detecting foreground, shadow and object from sequence of frames. Simulation results are presented by considering a lawn tennis ground.

On the basis of the obtained results following conclusions can be made:

1. The considered model correctly detects object from a frame.
2. To detect object minimum number of required layers are 5.
3. Euclidean distances have good impact in shadow detection.
4. Results improved as the number of training images are increases.

The result obtained in the thesis are early results and set directions for the development of a system which can be used for lawn tennis coaching, player and ball tracking.

References:

[1] A. 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. In- tell., vol. 19, no. 7, pp. 780–785, Jul. 1997.

[3] C. Stauffer and W. Grimson, "Adaptive background mixture models for real- time tracking," in IEEE Comput. Soc. Conf. CVPR, 1999, pp.246–252.

[4] I. Haritaoglu, D. Harwood, and L. Davis, "W4: Real-time surveillance of people and their activities," IEEE Trans. Patt. Anal. Mach. Intell., vol. 22, no. 8, pp. 809–830, Aug. 2000.

[5] J. Jacques, C. Jung, and S. Musse, "Background subtraction and shadow detection in gray scale video sequences," in Eighteenth Brazilian Symp. Computer Graphics and Image Processing, Oct.2005, pp. 189–196.

[6] 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.

[7] O. Barnich and M. Van Droogenbroeck, "ViBe: A universal back- ground subtraction algorithm for video sequences," IEEE Trans. Image Process., vol. 20, no. 6, pp. 1709–1724, Jun. 2011.

[8] W. Kim and C. Kim, "Background subtraction for dynamic texture scenes using fuzzy color histograms," IEEE Signal Process. Lett., vol.19, no. 3, pp. 127–130,Mar. 2012.

[9]V. Reddy, C. Sanderson, and B. Lovell, "Improved foreground detection via block-based classifier cascade with probabilistic decision inte- gration," IEEE Trans. Circuits Syst. Video Technol., vol. 23, no. 1, pp.83–93, Jan. 2013.

[10] 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.

[11] J. Choi, Y. Yoo, and J. Choi, "Adaptive shadow estimator for removing shadow of moving object," Volume 3 Issue 11, November 2014.

[12] Z. Liu, K. Huang, and T. Tan, "Cast shadow removal in a hierarchical manner using MRF," IEEE Trans. Circuits Syst. Video Technol., vol. 22, no. 1, pp. 56–66, Jan. 2012.

[13] 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.

[14] Z. Zivkovic, "Improved adaptive Gausian mixture model for back- ground subtraction," in Proc. IEEE Int. Conf Pattern Recognition, Aug.2004, pp. 28–31.

[15] E. Salvador, A. Cavallaro, and T. Ebrahimi. Cast shadow segmentation using invariant color features. CVIU, 95(2):238–259, August 2004.

[16] O. Schreer, I. Feldmann, U. Goelz, and P. Kauff. Fast and robust shadow detection in videoconference applications. In Proc. IEEE VIPromCom, pages 371–375, 2002.

[17] A. Senior, A. Hampapur, Y.-L. Tian, L. Brown, S. Pankanti, and R. Bolle. Appearance models for occlusion handlinge. Image and Vision Computing, 24(11):1233–1243, 2006.