International Journal of Science and Research (IJSR)
ISSN: 2319-7064
ResearchGate Impact Factor (2018): 0.28 | SJIF (2019): 7.583

Moving Objects Detection and Tracking in Using Improved Odd frame Difference and Camshift Algorithm

Said I Mohamed[1], Li Xinghua[2]

[1, 2]State Key Laboratory of Precision Measuring Technology and Instruments, Tianjin University, Tianjin City, China

Abstract: ***At present, detection and tracking for moving objects from video sequences have been implemented in many fields due to its important role. Video surveillance, medical imaging, robot vision, traffic monitoring, biomedical image analysis, and others are among the fields on which object detection and tracking are applied. Different algorithms used, where background subtraction and frame differencing are most efficient for moving objects detections, and Camshift method for tracking. Traditional frame differencing tends to cause the noise in an image and false alarms, normally when the illumination changes. This paper presented the improved odd frame difference techniques which reduce the noise and false alarms in detects moving objects and tracking moving objects using Camshift techniques. The techniques use the odd frame difference, mean of the difference set thresholds and the global thresholds using the "bitwise OR" operator. The binary image presents the foreground if the consecutive sequence of binary intersects and the result is "HIGH, 1". The results obtained eradicate the effects caused by the traditional frame differencing, and it sounds the accuracy and efficiency of the techniques in detecting and tracking moving objects in the video sequences in the camera.***

Keywords: Object detection, tracking, odd frame difference, Camshift

1. Introduction

In recent years, the vast and rapid growth of technology yields the increase and broads of the research directions nowadays. The intelligent and automated surveillance system has become demanding, a great and very important area in the computer vision research field [1], [2]. Easily availability and cheap high-quality video cameras lead to an increment ofinterest in areas such as motion detection, object detection [3], face detection, recognition, image retrieval, industrialinspection[4], and tracking. Detection and tracking are important and have higher practical values for many applicationssuch as Medical therapy, security systems, and driver assistance systems.

Object detection is the first step towards the tracking process, the main function is to detect the location of an object ofinterest in a given video frame sequence. The object tracking is the process on which a moving object (or multiple objects) is locating over a sequence of videos. Tracking is usually performed on higher-level applications that require the locationand object in every frame. The main aim of the object tracking is to generate the trajectory of an object over time by locating it is a position in every frame of the video [5] and also to give the overview region in the image that is occupied by the object at every instant time despite they cause some problems in practical applications such as noise or disturbance [14].

There are myriad techniques used in object detection of interest from which a camera is fixed or dynamic, however, thedynamic camera is a difficulty to detect objects of interest[6]. Currently, different algorithms were proposed to ameliorate the efficiency of object detection and tracking such Background subtraction, Frame Differencing, Point Detector, andOptical flow as shown in **Fig. 1**. Background subtraction, also known as foreground detection is a popular widely usedtechnique in image processing and Computer vision where the image foreground is extracted for further processing. Normally, images of interest are Humans, Cars, text, etc. are obtained by subtracting the background pixels in the scene, and this approach is quite fast and accurate for the fixed background. Background subtraction is affected mostly by the dynamic (non-stationary) background, low resolution, and illumination changes [7].

The existing frame difference method is relatively simple and has some effects [13] such as noise in the image,objects similarity in the scene, interdependence of variable quantity, and false alarms, many algorithms were developed to overcome the effects. Uses of global thresholding lead the effect that tends to fail to scan all the moving objects in the frame, illumination changes are the great factor to that effect. Improved odd Frame Difference is the simplest technique that we can use to obtain the parts in the video frame which are moving, the absolute difference between successive frames gives a lot of information for further processing. The presented algorithm has better performance against those effects resulted from the traditional frame difference.

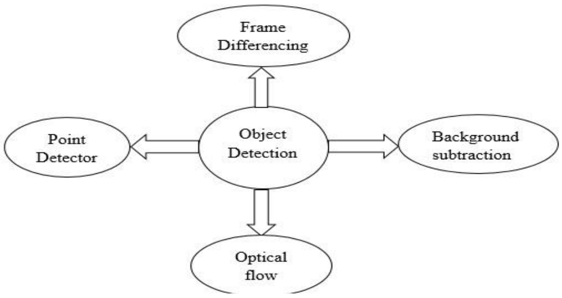

Figure 1: Object Detection Techniques.

Background subtraction and Frame Differencing are the methods which are most suitable for object detection technique.Every tracking mechanism required object detection in every frame or when the first objects appear in the video, PointTracking, Kernel Tracking, and Silhouette

Tracking, are the most popular tracking methods as shown in **Fig. 2.** Keeping track a moving is a hard and challenging issue as well, many parameters are responsible for making tracking movingobjects difficult [3] such as noise in the image, complex motion of objects, loss of information, real-time processing requirements, and occlusion.

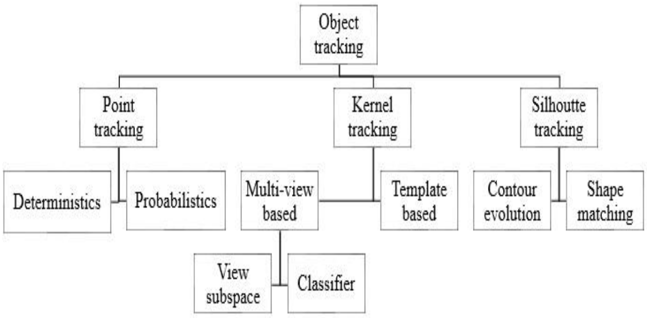

Figure 2: Object tracking techniques

2. Related Works

2.1. Object detection

Background subtraction and frame differencing are the high techniques for object detection due to its efficiency and accuracy in performance when applied for object detection. Tushar S. Waykole et al. [8] proposes the method for objectdetection and tracking which combines the advantages of Spatio-temporal differencing with basic background subtraction method, on which temporal features are used for the background updating, these lead the achievement of the background subtraction for the dynamically changing environment. Min Chen et al. [15] present the two prominent backgroundsubtraction method Gaussian Mixture (GMM) are applied to testify the approach. Given an input video sequence, frames were compressed at the ratio from 0% to 99%, GMM and ViBe methods are processed separately with every compression frames and the results are resized to its original. Zhanli Li et al.[2] presents the detection of moving objects using the three-frame difference method, takes any three consecutive frame images, and calculates the difference images of the two adjacent frames respectively. The threshold is been set to each of the differences to obtain the binary images andperform the logical operation in each corresponding pixel. Shreyamsh Kamate et al. [6] have suggested myriad algorithms such as background subtraction, Histogram of Oriented Gradients (HOG) descriptors, and adaptive background subtraction using Gaussian mixture models. Shuai Zhang et al. [9] have a focus on background modeling, adaptive Gaussian mixture modeling (AGMM), and MS-based segmentation. The MS uses the estimated initial location of the objects and improved, the holes will be inpainted and the object trajectory and boundary will be better aligned with trueobjects boundary, also separating the occluded objects in the mask by using depth information. MS perform color-based and different color reflect different segments and filter out the small regions in the original object mask obtained in AGMM. Suraj K Mankani et al. [10] focus on two methods, background subtraction and frame differencing. The implementations were on frame differencing to achieve the object detection, frames captured from the camera were pre-processed and the quality of the image is enhanced. The difference in the pixel values of the successive frame is calculated and compared with that of the threshold to assure whether the object is foreground or background. The binary object map is assigning '1' for the foreground pixel and '0' for background pixel.

2.2 Object tracking

Perhaps there are many object tracking algorithm implemented, Continuously Adaptive Mean Shift Tracking (Camshift)and Meanshift were much applied. Shreyamsh Kamate et al. [6] introduces the Camshift techniques based on mean-shift to perform the object tracking, the technique adjusts the window adaptively, moments can be used to find the objects newsize and update the scale during tracking, also makes the use of color histogram and texture of the objects as features for tracking. Hammad Naeem et al. [11] employs the tracking of the objects by using Camshift. An adaptive windows size is employed to deal with dynamically color distributions, each frame is applied with many iterations continuously and position and size are calculated for the next frame. Yan Zhang et al. [12] RGB color space is very sensitive, Camshift algorithm transforms the RGB color space into HSV color space and establishes the histogram and the feature match, setting the search window centroid as the zero moments.

3. Proposed Works

3.1 Object detection

In this paper, implementation the uses of background subtraction, $[BS(x,y)_n]$ andimproved frame difference using the consecutive odd frame, $[FD(x,y)_n]$.The widely used method for the object detection process on which the main objective is to obtain the foreground objects of interest. If there is no moving object represented then we treat the frame as a background $[F(x,y)_{n=0}]$ and it's subtracted from each frame in the video sequence. The method used when the objects are in motion, the mathematical principles of the methods are calculated as:-

$$BS(x,y)_n = F(x,y)_n - F(x,y)_n \quad (1)$$
$$FD(x,y)_n = F(x,y)_n - F(x,y)_{n-1} \quad (2)$$

$$where \quad n = 0,1,2,3, \dots \dots .. p$$

$F(x,y)_n$is a current frame, $F(x,y)_{n-1}$is a previous frame, and $FG(x,y)$is foreground frame

A pixel is considered to possess the motion if the difference is greater than the threshold and it is known as the foreground, see equation (3).

$$FG(x,y) = \begin{cases} 1, & FD(x,y) \geq T_h, at\ time\ t \\ 0, & otherwise \end{cases} \quad (3)$$

The proposed method is implemented as follows. **Fig 3.** Presents the flowchart diagram of the proposed odd frame difference algorithm.Assume the video sequence, $f_1, f_2, f_3, f_4, \dots \dots .. f_n$ for the implementation of the improved odd frames difference techniques, take the five consecutive frames and we find the difference between the sequences of frames as:

$$D_{k-2}, D_{k-1}(x,y) = abs(f_{k-2}(x,y) - f_{k-1}(x,y) \quad (4)$$

Paper ID: SR20713130204 DOI: 10.21275/SR20713130204

$$D_{k-1}, D_k(x,y) = abs(f_{k-1}(x,y) - f_k(x,y) \quad (5)$$
$$D_k, D_{k+1}(x,y) = abs(f_k(x,y) - f_{k+1}(x,y) \quad (6)$$
$$D_{k+1}, D_{k+2}(x,y) = abs(f_{k+1}(x,y) - f_{k+2}(x,y) \quad (7)$$

In this improving method, the foreground is presented when the binary image is greater than to the mean threshold, μ_{th} bitwise OR the global threshold,T_h see **table 1** below.

Table 1: Bitwise OR operations

T_h	μ_{th}	$T_h \mid \mu_{th}$
1	1	1
1	0	1
0	1	1
0	0	0

The mean threshold is obtained by finding the mean of the set threshold of each of the difference, for the five sequence video frame assume the difference threshold is as follows: $t_{(k-2,\ k-1)h}$,$t_{(k-1,\ k)h}$, $t_{(k,\ k+1)h}$, and $t_{(k+1,\ k+2)h}$.

The mean threshold, μ_{th} *is*:

$$\mu_{th} = \sum \frac{threshold\ of\ the\ difference}{total\ number} \quad (8)$$

$D_{k-2}, D_{k-1}(x,y), D_{k-1}, D_k(x,y), D_k, D_{k+1}(x,y), and\ D_{k+1}, D_{k+2}(x,y)$ Its binary images are $b_{k-2}, b_{k-1}(x,y), b_{k-1}, b_k(x,y), b_k, b_{k+1}(x,y), and\ b_{k+1}, b_{k+2}(x,y)\ respectively.$

$$b_{k-2}, b_{k-1}(x,y) = \begin{cases} 1, & D_{k-2}, D_{k-1}(x,y) > \mu_{th} | T_h \\ 0, & D_{k-2}, D_{k-1}(x,y) \leq 0 \end{cases} \quad (9)$$

$$b_{k-1}, b_k(x,y) = \begin{cases} 1, & D_{k-1}, D_k(x,y) > \mu_{th} | T_h \\ 0, & D_{k-1}, D_k(x,y) \leq 0 \end{cases} \quad (10)$$

$$b_k, b_{k+1}(x,y) = \begin{cases} 1, & D_k, D_{k+1}(x,y) > \mu_{th} | T_h \\ 0, & D_k, D_{k+1}(x,y) \leq 0 \end{cases} \quad (11)$$

A binary image contains only two gray levels, we take a logical intersection operation to the presented consecutive two sequence binary in each pixel and achieve the results after the odd frame difference i.e. in this we take five frames, "1" represents the foreground regions and "0" the background.

$$B(x,y) = \begin{cases} 1\ if\ intersection\ of\ binary = 1 \\ 0 \quad\quad otherwise \end{cases} \quad (13)$$

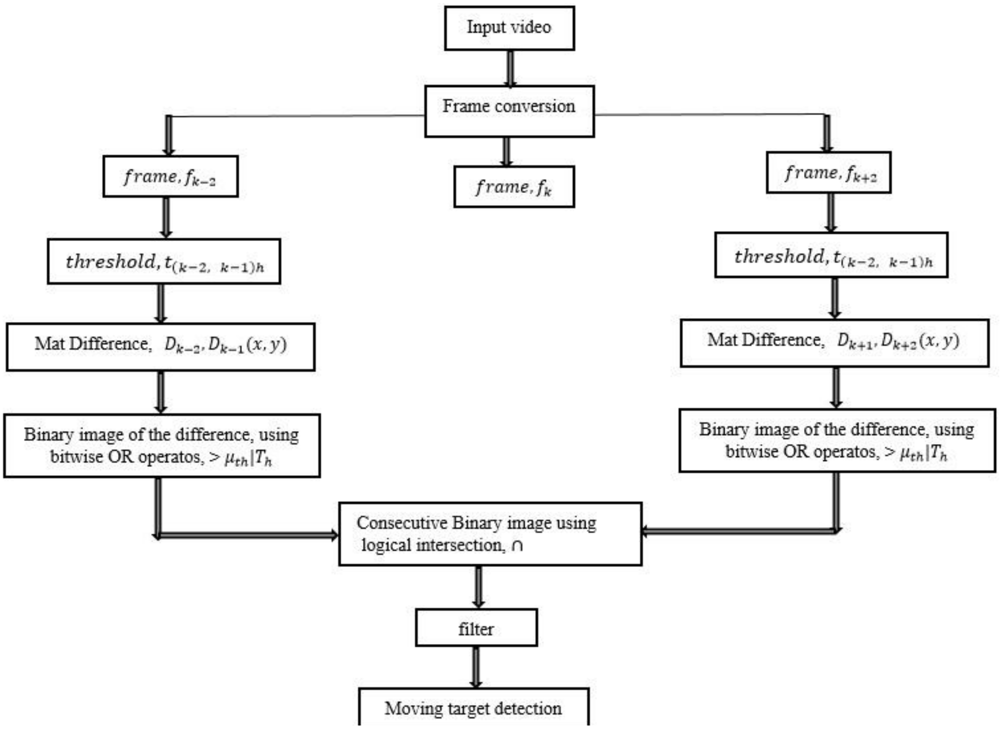

Figure 3: Flowchart of the proposed algorithm

The implementation steps of the improved odd frame difference method in simplified are as follows:

First, takes the odd consecutive frames, in this paper five consecutive frames in the video were used $f_{k-2}(x,y), f_{k-1}(x,y), f_k(x,y), f_{k+1}(x,y), and\ f_{k+2}(x,y).$

Convert the frames to the gray color space, and then calculate the difference images as:

$$D_{k-2}, D_{k-1}(x,y), D_{k-1}, D_k(x,y), D_k, D_{k+1}(x,y), and\ D_{k+1}, D_{k+2}(x,y)$$

For each two consecutive adjacent frame sequence and the result is binarized.

Set threshold for each difference image and the global threshold, finds the mean of the set threshold of the difference, the binary image is obtained by using bitwise OR operator of the mean and global threshold.

Takes a logical intersection operator to the binary and get the binary image,$B(x,y)$ "1" represents the foreground region and "0" represents the background regions.

The great advantage of this improved algorithm is the image scanned in a single and runned count the number of pixels found on each intensity, extract all thedynamic frames objects as foreground, eradicate all the noises in the image, and ignore the static one, as it assumes all static is the background.**Fig. 4, Fig. 5, and Fig.6**present some experimental results for the proposed techniques using different video sequences and different frame numbers.

Figure 4 (a):1000th frame (b) 1001th frame

Figure 4 (c):1002th frame (b) 1003th frame

Figure 4 (e):1004th frame

Figure 4 (f): Detected moving objects

Figure 5 (a): 100th frame **(b)** 101th frame

Figure 5 (c): 102th frame (d) 103th frame

Figure 5 (e): 104th frame

Figure 5 (f): Detected moving objects

Figure 6 (a): 200th frame (b) 201th frame

Figure 6 (c): 202th frame (d) 203th frame

Figure 6 (e): 204th frame

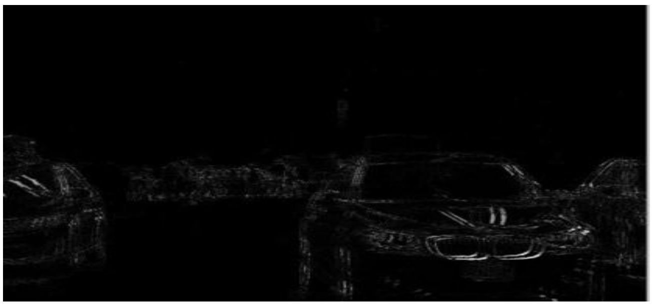

Figure 6 (f): Detected moving objects

Video sequences with different parameters were used as input for the proposed method, **table 2**. Shows the video parameters as follows.

Table 2: Video parameters

Video name	*Total frame*	*Five frames numbers used*	*Frame per second*	*Video length in sec*
videotest.mp4	4040	1000-1004	29.97	135
gari.mp4	332	100-104	25	13
Car5.mp4	682	200-204	25	27
Carme3.mp4	250	200-204	25	10
Car4.mp4	522	50-54	29.97	15

The normal background subtraction method used for the video sequence to present the foreground in the sequences, **Fig.7** display the results achieved

Figure 7 (a): Original Video frame

Figure 7 (b): Results of background subtraction.

The system diagram shown in **Fig. 8**, describes the various parts and their functionalities for the moving objects detection in very simplified.

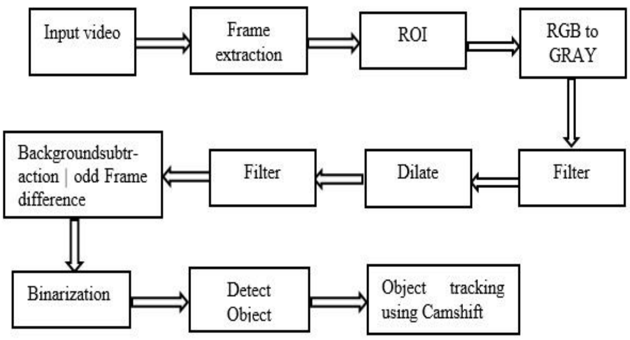

Figure 8: Block diagram of the proposed system.

Input video: The sequences of videos frame are used as input parameters.

Frame extraction: After inputted the video, the important parameter to start working with is the extraction of the frame. In this, we use the VideoCapture object in OpenCV library, let assume the frame intensity at time t and $t+1$ are: $I(x,y)_t$ and $I(x,y)_{t+1}$ the object seems to appear if the difference between the two frames is not equal to zero, otherwise no displacement of an object between frames.

ROI (Region of Interest): Are samples within a data set needed for the detection and tracking purposes, such as the size.

RGB to GRAY conversion: Images are made up of pixels and channels, RGB possesses three channels. We need to convert the color image into the gray value, we use the cvtColor functions in C++ to perform the conversion, using the OpenCV library.

Filter: We use a Gaussian blur filter to remove noise, smoothing the image, and feature extraction. The images should be convolved with a Gaussian kernel matrix to produce the smoothed image. Each channel in the original image is processed independently based on the Gaussian function.

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \qquad (14)$$

Dilate: It is a morphological operation used to eliminate small holes and fills the gaps in contours. The function dilates the source image using the specified structuring element that determines the shape of a pixel.

Background subtraction / odd frame difference: Background algorithm used for the detection of moving objects, we use this method because it is simple and widely used for generating a foreground mask. BackgroundSubtractor class is implemented to perform the BS algorithm in C++ using the OpenCV library. Odd Frame difference is also used to generate the foreground mask and store the result as a foreground mask.

Binarization: Image binarization converts an image of up to 256 gray levels to a black and white image, it used a pre-processor. It checks threshold values and classifies pixels with values above this threshold as white and others are black.

3.2. Object tracking

Moving object tracking calculates the object position of each frame, connects the same object in different frames to obtain the trajectory of the moving objects. There are different algorithms used to perform the object tracking mechanism. Meanshift method, Lucas-Kanade optical flow tracking [6], Block matching algorithm (BMA) [5], Bayesian Kalman filter with Gaussian mixture algorithm [9], etc. each method have its pros and cons and is selected based on the environment.

We make the use of the Camshift algorithm to track the only one moving object at a time, to minimize the influence of brightness change of the object RGB color space conversion to HSV color space must be performed. Camshift uses the HUE component to establish a histogram and feature match. The Camshift first gets the ROI region, calculates the color probability of the region centered at the Meanshift for new size and position of the search window, and finally iterate to find the centroid and store the zeroth moment and centroid location. If (x,y) is the pixel position in the search window,

$I(x, y)$ is the intensity value at the projection(x, y). The mass center calculation is:

Zero[th]moment M_{00}:

$$M_{00} = \sum_x \sum_y I(x, y) \qquad (15)$$

First moment for x and y as:

$$M_{10} = \sum_x \sum_y x\, I(x, y) \qquad (16)$$

$$M_{01} = \sum_x \sum_y y\, I(x, y) \qquad (17)$$

The mean search window location is:

$$(x_c, y_c) = \left[\frac{M_{10}}{M_{00}}, \frac{M_{01}}{M_{00}}\right] \qquad (18)$$

The Camshift algorithm is implemented in C++ using the OpenCV library and the single moving object and successfully tracked.

4. Experimental Verification

The new odd frame-difference algorithm using five consecutive adjacent framesis implemented and shows better performance in solving the existing problems. The experiment has been conducted in the hope of platform, Visual Studio 2019, C++ programming language, and OpenCV version 4.3.0. The different video parameters are used as input. Some experiments have been conducted to withstand the performance of the proposed method and the result obtained is presented in **Fig. 4,Fig. 5,** and **Fig. 6**above in the proposed works section. To verify the proposed method, consider **fig. 9** and **fig. 10** below, which presents the results of the existing frame difference and the proposed methods respectively.

(a) 1000th frame (b)1001th frame

(c) Moving objects

Figure 9: Existing frame difference

(a) 1000th frame (b) 1001th frame (c) 1002th frame

(d) 1003th frame (e) 1004th frame

(f) Moving objects

Figure 10: Proposed Frame difference method

The above figures present the moving detection results, it is seen that fig.10 possesses noise as well as the strong interdependence variable, i.e. it is strongly correlated. The fig. 11 shows the much improvements of the effects despite there is minor noise in the moving objects detection. The obtained results knocked that the proposed system has better performance in the detection part.

The Camshift method is used to track the moving objects, **Fig.11** shows the experimental results obtained, and it sounds that the method works well in real-time, speed, and it's robust.

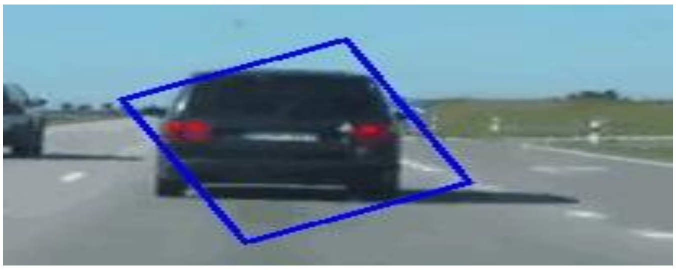

Figure 11 (a): tracked object with its block effect.

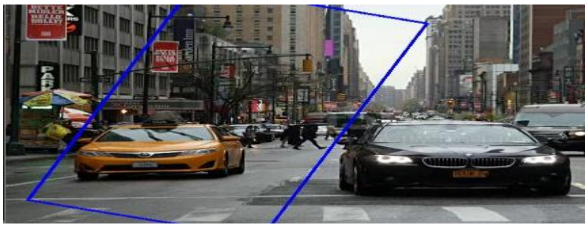

Figure 11 (b): tracked object with its block effect.

Figure 11 (c): Tracked object with its block effect.

5. Conclusion

This paper proposes the odd frame difference method using five consecutive frames to detect the moving objects and overcomes the weakness shown in the existing frame difference techniques such as noise, objects similarity in the scene,the interdependence of variable quantity, and false alarms. Also,background subtraction used as another method to detect moving objects,The Camshift method used to track the moving objects in the video sequences, the algorithm iteratively, and calculate the best matching window and spatial information. Experimental results show that the method this paper proposes is better performed and can accurately detect and extract all the objects (foreground) in the video sequences. It assists in the detection, hence based on the experimental results can conclude that the method works accurately.

References

[1] P. P. Gangal, V. R. Satpute, K. D. Kulat, and A. G. Keskar, "Object Detection and Tracking Using2D -DWT and Variance Method" Visvesvaraya National Institute of Technology, Nagpur-440001, India.
[2] Zhanli Li, Fang Yang, Hong-an Li, " Improved moving object detection and tracking method," Proc. SPIE 10011, First International Workshop on Pattern Recognition, pp. 1001106-2-1001106-3, (11 July 2016);.
[3] Apurva S. Samdurkar, Shailesh D. Kamble, Nileshsingh V. Thakur, Akshay S. Patharkar, "Overview of Object Detection and Tracking based on Block Matching Techniques" Proceedings of the Second International Conference on Research in Intelligent and Computing in Engineering pp. 313–315, Vol. 10 ISSN 2300-5963.
[4] B. N Krishna Sai, Sasikala T, "Object Detection and Count of Objects in Image using Tensor Flow Object Detection API" pp. 542-545,Second International Conference on Smart Systems and Inventive Technology (ICSSIT 2019) IEEE Xplore Part Number: CFP19P17-ART; ISBN:978-1-7281-2119-2.
[5] P. S. Khude, S. S. Pawar," Object Detection, Tracking and Counting using enhanced BMA on static background videos" 2013 IEEE International Conference on Computational Intelligence and Computing Research.
[6] Shreyamsh Kamate, Nuri Yilmazer, "Application of Object Detection and Tracking Techniques for Unmanned Aerial Vehicles" pp. 437-439, Conference Organized by Missouri University of Science and Technology 2015-San Jose, CA 2015.
[7] Shraddha Mane, Prof.SupriyaMangale," Moving object detection and tracking Using Convolutional Neural Networks" pp.1809-1810. Proceedings of the Second International Conference on Intelligent Computing and Control Systems (ICICCS 2018) IEEE Xplore Compliant Part Number: CFP18K74-ART; ISBN:978-1-5386-2842-3
[8] Tushar S. Waykole, Yogendra Kumar Jain, " Detecting and Tracking of Moving Objects from Video"pp.23-25.International Journal of Computer Applications (0975 – 8887) Volume 81 – No 18, November 2013
[9] Shuai Zhang, Chong Wang, Shing-Chow Chan, "New Object Detection, Tracking, and Recognition Approaches for Video Surveillance Over Camera Network" pp.2681-2687. IEEE SENSORS JOURNAL, VOL. 15, NO. 5, MAY 2015
[10] Dong -Hyung Kim, Youngmyung Lee, Ji-Yeong Lee, Gyung-Jin Park, Chang-Soo Han, Sunil K. Agrawal " Detection, Motion Planning and Control of Human Tracking Mobile Robots" pp.1965-1967. The 8th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI) Nov.23-26,2011 in Songdo ConventiA, Incheon, Korea.
[11] Hammad Naeem, Jawad Ahmad, and Muhammad Tayyab, "Real-Time Object Detection and Tracking" pp.149-153. HITEC University Taxila Cantt, Pakistan.
[12] Yan Zhang, Lijie Li, " A Pedestrian Tracking Algorithm Based on Camshift and Kalman Filtering" p.2256 International Journal of Science and Research (IJSR) ISSN (Online): 2319-706.
[13] Lei Shang, Fucheng You, Chao Han, Xuewei Wang, and Shuai Zhao, "Optimization of Three - Frame Difference Method and Improvement of Pedestrian Detection Code Book" pp. 1-7, Journal of Physics: Conference Series.
[14] K. Rasool Reddy, K. Hari Priya, N. Neelima," Object Detection and Tracking – A Survey"
[15] Min Chen, Andy Song, Shivanthan A. C. Yhanandan, and Jing Zhang,"Background subtraction using compressed Low-resolution Images"Journal of Latex Class files, Vol. 14, No. 8,