International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Selected Three Frame Difference Method for Moving Object Detection

Emrah Simsek*[1,2], Baris Ozyer[1]

Submitted: 04/02/2021 Accepted : 19/05/2021

Abstract: Three frame difference is one of the well-known methods that is used to perform moving object detection. According to the theory, the presence of a moving object is estimated by subtracting consecutive three image frames that provide moving object edges. However, these edges do not give complete information of the moving object which means that the method leads to loss of information. Some post-processing methods such as morphological operations, optical flow and combining these techniques are necessary to be apply for obtaining complete information of moving object. In this paper, we present a new approach called Selected Three Frame Difference (STFD) to detect moving object in video sequences without any post processing operations. We initially propose an algorithm that selects three images considering the local maximum value of frame differences. Instead of using consecutive three frames, these three selected image differences that include non-overlapping object frames are applied to the logical and operator. We mathematically prove that the entire moving object is always detected in the second selected image. We analyzed the proposed method on public benchmark dataset and the dataset collected from our laboratory. To validate the performance of our approach, we also compared with three frame difference method and background subtraction based traditional moving object detection methods on a few sample videos selected from different datasets.

Keywords: *Frame Difference, Three Frame Difference, Moving Object Detection*

This is an open access article under the CC BY-SA 4.0 license.
(https://creativecommons.org/licenses/by-sa/4.0/)

1. Introduction

Recently moving object detection in video sequences is required for making analysis more intelligent aspect in computer vision area such as surveillance in daily and wild life, retrieval and recognition [1-8]. The key issue of moving object in image frame can be defined as a region collected by a set of connected pixel that are segmented from the background. In general, the spatio-temporal changes between the consecutive images caused by moving object are considered to distinguish object in the scene from background [9-12]. However, there are so many difficulties obtaining robust model to detect object moving in the video sequences obtained from either indoor or outdoor scene. Lighting conditions, sudden fluctuations in direction of the object motion, complex background, shadow and non-rigid object deformation are some important problems that have been studied by researchers during the last decades in different aspect of view reported in [13].

Background subtraction, optical flow and frame difference are well-known algorithms for obtaining location of moving object in the video frames. Although mathematical approximation is very simple for all methods, in real applications, each method has advantages and disadvantages between each other. Background subtraction is basically a technique that detects object by subtracting background models which are estimated from consecutive image sequences by taking average of intensity [14-17]. This method performs well accuracy when the camera is positioned as static and only small movement in the background can be available. Optical flow requires significant displacement image points which can be defined as location of moving pixels over the image frames [18,19]. Optical flow is represented by two-dimensional vector which are velocities and direction of image points. Since it is sensitive to illumination changes and noises, additional processing is required to obtain distinguishing results that leads to complex computationally. The frame difference method, commonly used expression as temporal difference method, basically estimates presence of the moving object between two consecutive image [4,20]. Frame difference method [20,21], three frame difference method (TFD) [11, 22, 24] and four frame difference method [27] are such a different frame difference algorithm that have been proposed in the literature during last decades. The output of frame difference operations is the union of moving objects of two consecutive images [20]. The TFD methods are based on using logical “AND” (&) operator to the outputs of two frame difference obtained from the three consecutive images. The four frame difference method is based on the use of four consecutive images with the threshold parameter in order to avoid shadows in the TFD. Although effective results can be obtained by frame difference methods, there are some disadvantages such as leaving holes in foreground objects and detecting ghosts from background regions.

In classical approach of three frame difference method [11], moving objects' edges are approximately detected through consecutive of three images. However, it is not sufficient information to represent all boundary pixels of moving object. Edge combining algorithms [23], morphological analysis method [28], optical flow [19] are therefore used to improve the performance of three difference method for obtaining every pixel of moving object. Morphological analysis method such as erosion and expansion are used to clean noises and fill the edges. In addition, the tendency and trend directions are used to combine the detected edges. This

[1] *Computer Engineering Department, Ataturk University, Erzurum, Turkey ORCID ID : 0000-0003-0117-6983*
[2] *Digital Transformation Department, Erzurum Technical University, Erzurum, Turkey, ORCID ID : 0000-0002-1652-9553*
* *Corresponding Author Email: emrah.simsek@erzurum.edu.tr*

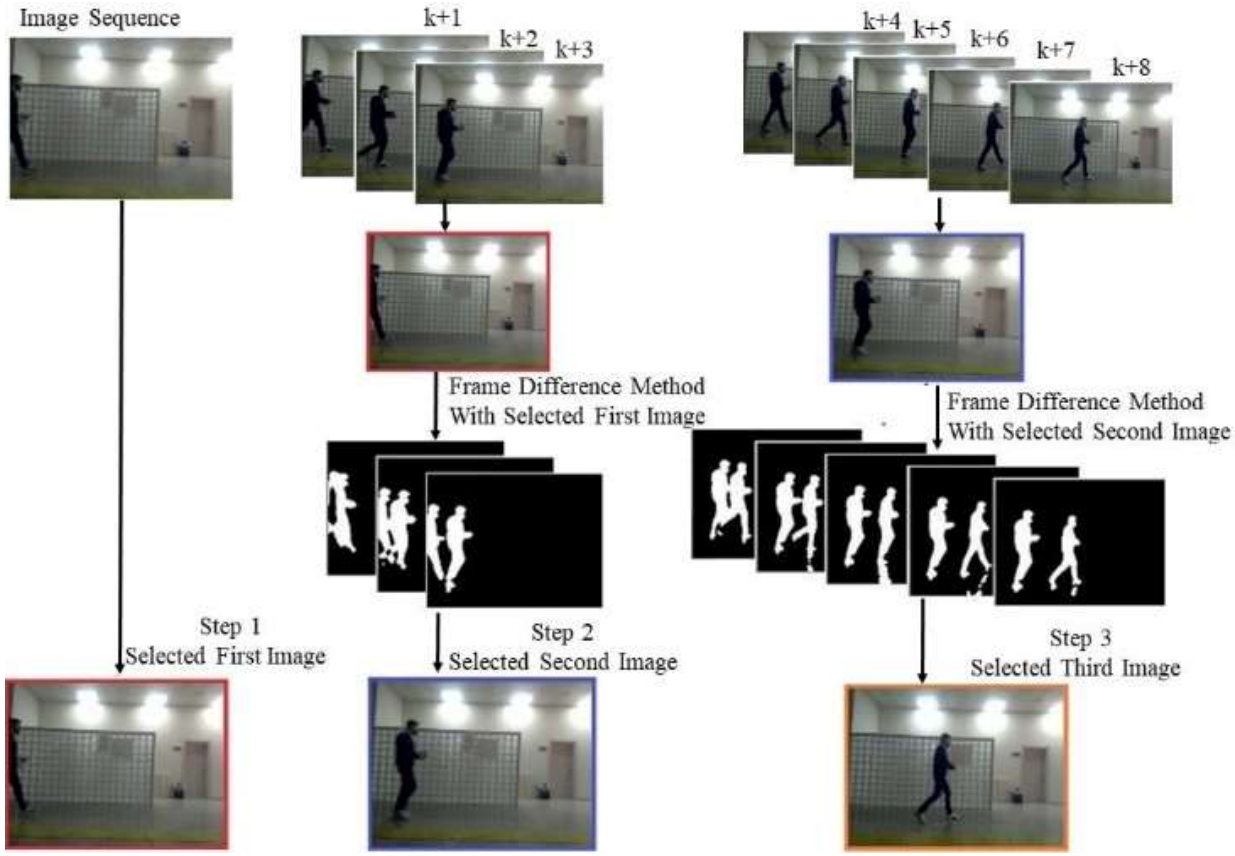


Fig. 1. The overview of the proposed image selection method.

method is inefficient where dynamic background is present [24,25]. Time-lapsed image selection method is applied to create the background model and identify the moving regions of the images [26]. Two-dimensional global thresholds have been applied on the interest areas to effectively clean the ambient noise. In a recent study [24], the background model is calculated by taking average intensity values of five consecutive images and then the outer line information of the moving object is obtained. However, it is not possible to detect more than one moving object. To overcome the limitations of conventional post processing algorithms and to obtain complete information of moving objects without applying any process after TFD method, we propose a new algorithm based on TFD method which is called selected three frame difference (STFD). The contribution of our proposed approach is provided by the selection of three images that is used for TFD methods. In our proposed method, images are selected by calculating local maximum differences between two images instead of selecting consecutive images. This contribution ensures to obtain the discrete and non-overlapping complete moving object regions of the output of the frame difference operation. We also mathematically prove that second selected image is always obtained after applying logical "AND" (&) operator to the frame difference methods.

The paper is organized as follows: We describe the proposed STFD algorithms in more details and present and mathematically prove our assumption for selecting images in Section 2. We show the performance of proposed algorithm on benchmark dataset and our own created dataset in Section 3. We summarize our findings and give future direction in Section 4.

2. Proposed Method

2.1. Three Frame Difference

In classic frame difference approaches [11,19], the moving regions is detected through subtracting consecutive images. It is assumed that image sequences are captured by a static camera. Let us consider that $I_k(x,y)$ is the intensity of the kth frame. The difference image $I_d(x,y)$ used to detect moving region is calculated as follows

$$I_d(x,y) = |I_{k+1}(x,y) - I_k(x,y)| \quad (1)$$

where x and y are the coordinates. In order to remove noises, the median filter is applied to the difference image. Then, global threshold value T_d is used to convert gray scale images to binary images defined as follows

$$I_f(x,y) = \begin{cases} 0: I_d(x,y) \leq T_d \\ 1: \quad otherwise \end{cases} \quad (2)$$

where I_f (x, y) is the binary image at the coordinate x and y. TFD method is originally developed by using frame difference and then applied logical operator from consecutive images. In this method, two frame difference outputs are obtained from three consecutive images. The moving regions are then detected by applying logical "AND" (&) operator on these output frames calculated as follow

$$I_b = I_{f1} \,\&\, I_{f2} \quad (3)$$

where I_b is the three frame difference binary output of images, I_{f1} is difference between k_{th} and $(k+1)_{th}$ images and I_{f2} is difference between $(k+1)_{th}$ and $(k+2)_{th}$ images. With the expressions of the TFD method, edges of the moving object in the $(k+1)_{th}$ image are obtained by applying logical "AND" (&) operator to the two frame difference output I_b [19,, 24].

2.2. Proposed Image Selection Method

Fig. 1 shows overall structure of our proposed image selection method. Our main contribution is that instead of using consecutive

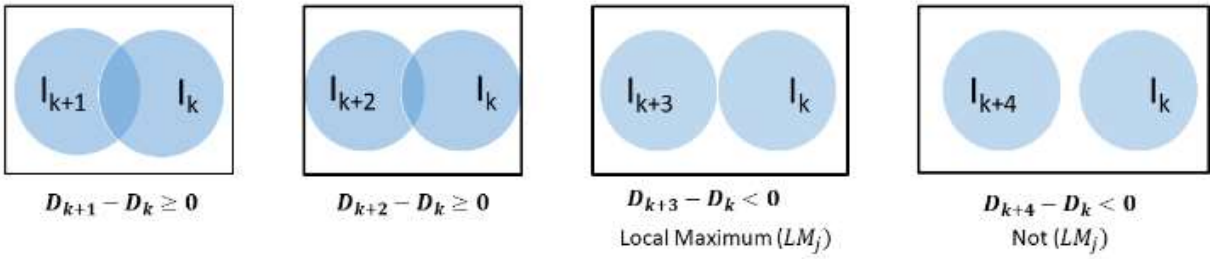


Fig. 2. Local maximum (LM) calculation among the consecutive images.

images, the selected images that contain the non-overlapping objects' regions will be used in the three-frame difference method. The details of proposed methods are described as follows. We have image sequences denoted as x= x_1, ..., x_n where n is the number of image frames.

Step 1: k_{th} is the first image in the consecutive images denoted as Id1 when moving object is detected in image sequences.

Step 2: Initially, the intensity difference between $(k+i)_{th}$ frame and selected first image frame k_{th} are calculated respectively. As the object moves within the scene, total amount of intensity changes D increases described in Algorithm 1. This is due to the fact that overlapped region of moving objects decreases. Fig. 2 shows the conceptional diagram of intersection region belonging to moving object according to intensity changes and differences. In order to select second image, we propose an algorithm described in Algorithm 2 that determines maximum total amount of intensity changes denoted as LM_j where j is the selected second frame. The second image denoted as I_{d2} is then selected corresponding to j_{th} frame. The differences between I_{d1} and I_{d2} images represented as I_{f1} is finally calculated.

Step-3: The third image I_{d3} is selected by computing next local maximum LM_{j+m} where m is the selected third image corresponding to selected j_{th} second image. In this case, the total amount of intensity changes is calculated according to the selected second image I_{d2}. The differences between I_{d3} and I_{d2} images represented as I_{f2} is finally calculated.

Result: Total amount of intensity changes *D* between $(k+i)_{th}$ frame and selected image *kth* frame initialization;
for *i = 1 to n-1* **do**
  Calculate frame difference I_{di} where $I_{di}=(I_k, I_{k+i})$;
  Compute median filter I_{mi} of I_{di};
  Obtain binary image I_{fi} of I_{mi};
  Calculate total amount of intensity changes D_i of I_{fi} where $D_i = \sum^{n} \sum^{m} I_{fi}(x,y)$;
end

Algorithm 1: The total amount of Intensity changes calculation *D*

Result: Local Maximum (LM) corresponding to previous selected frame;
for *j=1 to n-1* **do**
  If D_{j+1}-D_j >= *0* **then**
    Continue;
  else
    $LM_j = I_{fj}$;
  end
end

Algorithm 2: Searching negative intensity changes initialization to determine LM_j

Assumption:
The complete moving object region is detected in the second selected frame by taking the logical "AND" operator between non-overlapping moving objects in the frame of O_{f1} and O_{f2}.

Proof:
Let us define as O_{d1}, O_{d2} and O_{d3} are the non-overlapping moving objects of the first I_{d1}, second I_{d2} and third I_{d3} selected images. The intersection regions between for each moving objects of the selected images are calculated as empty set given as below.

$$O_{d1} \cap O_{d2} = \varphi, O_{d2} \cap O_{d3} = \varphi, O_{d1} \cap O_{d3} = \varphi \quad (4)$$

The moving object regions of the frame difference output I_{f1} and If2 is calculated as

$$O_{f1} = O_{d1} \cup O_{d2}, O_{f2} = O_{d2} \cup O_{d3} \quad (5)$$

Logical and operator is expressed as intersection of these regions defined as follows

$$O_{f1} \cap O_{f2} = (O_{d1} \cup O_{d2}) \cap (O_{d2} \cup O_{d3}) = (O_{d1} \cap O_{d2}) \cup (O_{d2} \cap O_{d3}) \cup (O_{d1} \cap O_{d2}) \cup (O_{d2} \cap O_{d3}) \quad (6)$$

This mathematical expression is simplified as

$$O_{f1} \cap O_{f2} = \varphi \cup \varphi \cup O_{d2} \cup \varphi \quad (7)$$

where we verify that the moving object region of the second image O_{d2} is obtained.

Result: Small Object Elimination
for *Every Output Binary Image* **do**
  Components=Find Connected _Components(Binary Image);
  for *Every Components* **do**
    Size=Calculate Size of Component ();
    if Size *is Bigger Than Image Size/35* **then**
      Bounding Box Coordinates=Find Coordinates (Component);
    end
  end
end

Algorithm 3: Connected component analysis is used to eliminate small objects in the scene

Connected component analysis is applied to determine bounding boxes of segmented moving detected objects. The aim of the analysis is to find the coordinates of the separated region in the binary image. The size of connected component is defined by bound box pixels that belongs to moving object or others objects. Algorithm 3 shows the details of connected component analysis approach for detecting moving objects that eliminates small size objects and reduce the detection time.

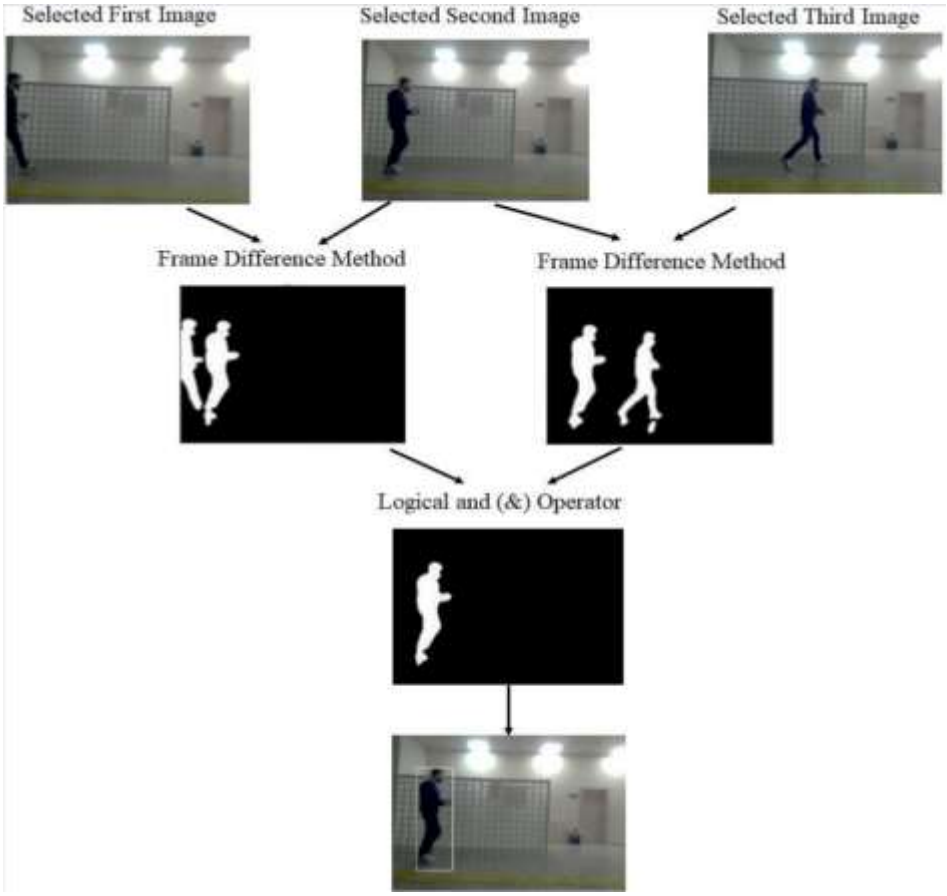


Fig. 3. The output images after applying logical **"AND" (**&) operator to frame difference output frames obtained from our proposed STFD method.

3. Experimental Results

We firstly tested our proposed image selection algorithm on three standard benchmark video datasets: CAVIAR[1] dataset, Fast Moving Object dataset (FMO)[2] and LASIESTA[3] dataset. We then analyzed the proposed STFD methods on real indoor movies dataset obtained from camera traps in our laboratory which is called CTRIN[4] Dataset. Table 1 shows detailed information of some videos of the datasets. In order to verify the computational simplicity and efficiency of our proposed method, we compare with existing difference based methods for all datasets. Our experiments are carried out with ANACONDA distribution and Python 3.7 version executed on an Intel(R) Xeon(R) X3440 cpu@2.53 GHz processor with 8GB RAM.

As seen overall structure of proposed method results in Fig. 1, the first image kth denoted as I_{d1} is directly selected when object motion is detected in the image frames. The second image I_{d2} and the third image I_{d3} is selected by calculating the local maximum L_{M1} and L_{M2} values in the consecutive images described in Algorithm 1 and Algorithm 2.

As seen in Fig. 1, the fourth image frame represented as $(k+3)_{th}$ in the consecutive images is the local maximum values. It is observed that the changes of difference are zero or negative in the next frame shown. The third image I_{d3} is selected by calculating the local maximum LM_2 according to the second selected frame I_{d2}. As seen in Fig. 1, the ninth image frame represented as $(k+8)_{th}$ in the consecutive images is the second local maximum value. The discrete and non-overlapping moving object regions of the output of the frame difference operation is obtained by using proposed image selection method. As seen in the Fig. 1, although the moving object regions is seen non overlapped in the frame difference outputs, $(k+5)_{th}$ or $(k+6)_{th}$ is not selected as the third image. This due to the fact that the image contains shadow of the moving object. At the end of the algorithm, the three image frame I_{d1}, I_{d2} and I_{d3} are selected in order to use in TFD method instead of using consecutive images. As shown in Fig. 3, moving objects of second selected image are detected by taking logical "AND" (&) operator between the selected output frame difference described in Equation (7). As seen from the results of output TFD, it is clearly observed that the complete information the moving object is obtained. In the experiments, we require three parameters for suppression of noise and small objects; global thresholding, the median filter box size and a threshold parameter for the disregarding small objects. Global threshold parameter is set to 17, the image size is set to 1/50 and the median filter box size is set to 1/35. It is observed that the median filter box size yielded successful results in many data sets in the value of 1/50 ratio. The accuracy of the proposed method was calculated by using recall, precision and F-score value represented in Equation (8), (9) and (10). True positives were determined according to the 0.5 and above value of the iou (intersection over union) of detected bounding boxes and the groundtruth of the dataset.

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative} \tag{8}$$

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive} \tag{9}$$

$$F - Score = 2\ x\ \frac{Precision\ x\ Recall}{Precision + Recall} \tag{10}$$

Table 1. Detailed Information of Benchmark Datasets 1,2 and own CTRIN dataset.

Video Name	*Num of Img*	*Num of Obj*	*Size*
Caviar Walk1	609	1375	384x288
Caviar Walk2	1559	2860	384x288
Caviar Walk3	1040	1039	384x288
FMO frisbee	100	18	1080x1920
FMO blue	53	22	720x1280
Lasiesta OCL02	425	419	352x288
CTRIN Video1	189	174	1200x1600
CTRIN Video10	175	141	1200x1600
CTRIN Video27	285	110	1200x1600

Table 2 represents precision, recall and F-score values of proposed the method applied on the benchmark CAVIAR, FMO and CTRIN datasets for testing the accuracy. As shown in the table, low precision value is obtained in the Caviar Walk2 video frames. The reason is that slowly moving object causes loss of information considering the different places that the moving target appeared in. On the other hand, due to containing comparatively fast moving objects in the image frames of FMO dataset, the accuracy increases. For our own CTRIN video dataset, it is clearly indicated that the precision and recall values are significantly high. One possible reason for the relatively good performance on these values might be elimination of shadow effect in the indoor environment. However, for all datasets, in the case of multiple moving objects available in the image frames, we observed that the accuracy of the proposed STFD method decreases. Table 3 shows the confusion matrix results of the proposed STFD method applied on all datasets.

Table 2. Accuracy of Benchmark Datasets[1,2] and CTRIN dataset

Video Name	Precision	Recall	F-Score
Caviar Walk1	0.98	0.97	0.97
Caviar Walk2	0.72	0.93	0.81
Caviar Walk3	0.98	0.83	0.90
FMO frisbee	1.00	1.00	1.00
FMO blue	0.95	0.84	0.89
CTRIN Video1	0.99	0.98	0.98
CTRIN Video10	1.00	0.96	0.98
CTRIN Video27	1.00	0.96	0.98

Table 3. Confusion Matrix of Benchmark Datasets[1,2] and CTRIN dataset

Video Name	TP	FP	FN
Caviar Walk1	1330	38	45
Caviar Walk2	2653	1016	207
Caviar Walk3	865	18	174
FMO frisbee	18	0	0
FMO blue	21	4	1
CTRIN Video1	171	1	3
CTRIN Video10	136	0	5
CTRIN Video27	110	12	0

Fig. 4 shows as a sample of output frames for the proposed STFD and TFD methods on all dataset. The goal of this comparison is to emphasize the effectiveness of our method on different video frames containing the various type of actions and environments. It can be observed that the proposed STFD method can detect moving objects with higher accuracy than three frame difference TFD method, providing only edges of the presence of moving objects

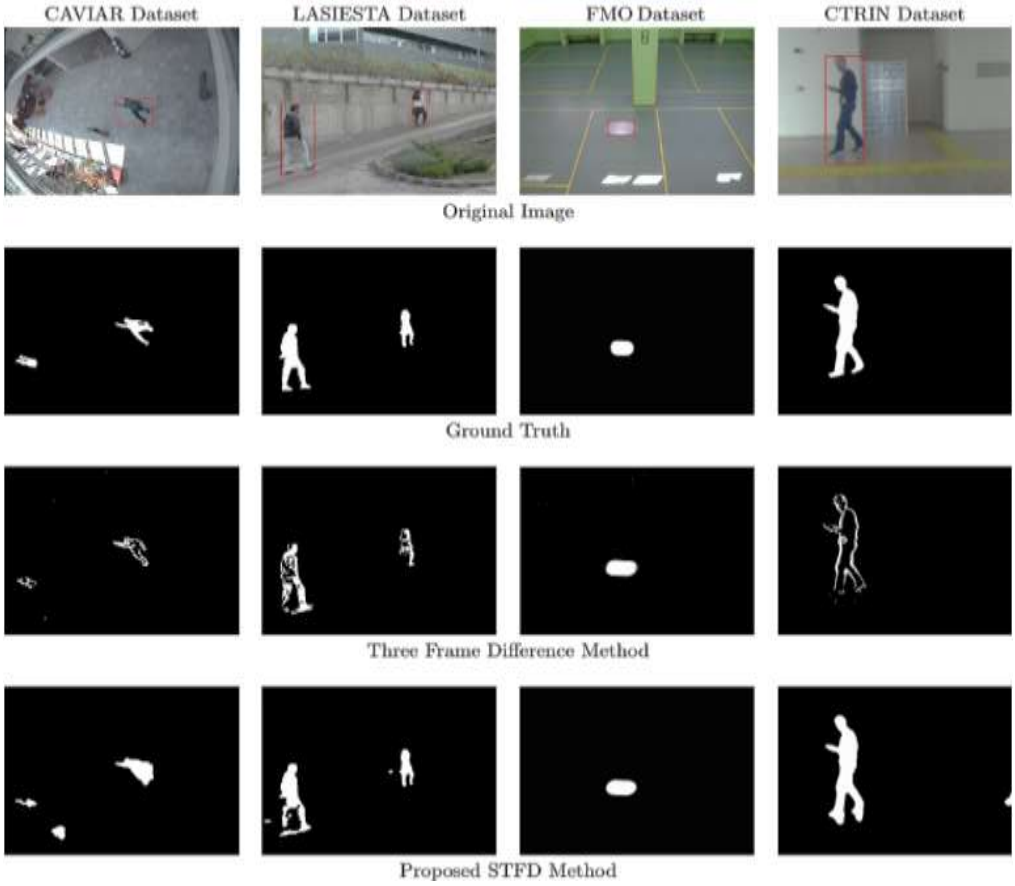


Fig. 4. Sample frames of the comparison between proposed TFD method and Three Frame Difference on benchmark datasets (CAVIAR, LASIESTA, FMO) and recently created CTRIN dataset in our laboratory.

[1] https://groups.inf.ed.ac.uk/vision/CAVIAR/CAVIARDATA1/
[2] http://cmp.felk.cvut.cz/fmo/
[3] https://www.gti.ssr.upm.es/data/lasiesta_database.html
[4] http://cogvi.atauni.edu.tr/ResearchLab/PageDetail/Our-CTRIN-Dataset-87

even the videos under taken from indoor or outdoor environments. We also compared proposed STFD method with TFD and background subtraction based traditional moving object detection methods; (GMG[29], MOG[30], MOG2[31], GRA[32], CNT[33], GSOC[34], KNN[35], LSBP[36]) in terms of F-score result on the LASIESTA3 dataset. Since the ground-truth of the LASIESTA3 dataset contains pixel-based segmentation results for each background and foreground class in all frames, the accuracy of the proposed approach can be analysed and compared more efficiently with other methods. Therefore, F-score values are calculated with each pixel classes evaluation in ground-truth instead of bounding box and iou based method. The detailed results are shown in Table 4. As seen from the table, the accuracy of the STFD is substantially higher than the TFD and background subtraction based traditional moving object detection methods.

In addition, we compared time complexity performance evaluation of the proposed STFD method with TFD method and background subtraction based traditional object detection methods on CTRIN dataset as shown in Table 5. As seen in table proposed method has the best processing time performance among the all approaches.

Table 4. F-Score and Time Complexity Comparison of Proposed Method, TFD and Background Subtraction Based Traditional Methods of Benchmark LASIESTA[3] dataset

	MOG	**GMG**	**MOG2**	**GRA**	**CNT**	**GSOC**	**KNN**	**LSBP**	**TFD**	**STFD**
I_BS_01	0,60	0,58	0,60	0,38	0,46	0,40	0,58	0,36	0,41	**0,72**
I_BS_02	0,74	0,66	0,81	0,42	0,49	0,65	**0,77**	0,43	0,54	0,75
I_CA_01	0,70	0,54	0,75	0,44	0,65	0,78	0,76	**0,79**	0,28	**0,79**
I_CA_02	0,55	0,63	0,56	**0,84**	0,56	0,75	0,60	0,67	0,06	0,64
I_IL_01	0,79	0,77	**0,81**	0,18	0,48	0,50	0,59	0,52	0,71	0,78
I_IL_02	**0,77**	0,62	0,76	0,39	0,51	0,53	0,66	0,41	0,44	0,73
I_MB_01	0,63	0,45	0,64	0,87	0,66	0,77	0,68	**0,92**	0,24	0,71
I_MB_02	0,73	0,64	0,79	0,78	0,73	0,77	0,77	0,81	0,43	**0,89**
I_MC_01	0,16	**0,67**	0,18	0,06	0,08	0,09	0,11	0,08	0,08	0,50
I_MC_02	0,21	**0,80**	0,24	0,08	0,09	0,11	0,15	0,08	0,33	0,46
I_OC_01	0,91	0,78	0,92	0,87	0,87	0,82	**0,93**	0,84	0,53	0,85
I_OC_02	0,90	0,77	0,89	0,42	0,89	0,87	**0,92**	0,94	0,51	0,80
I_SI_01	0,83	0,76	**0,90**	0,74	0,76	**0,90**	0,86	0,87	0,52	0,88
I_SI_02	0,77	0,59	0,82	0,82	0,78	0,80	**0,83**	0,71	0,37	0,77
I_SM_01	0,47	**0,61**	0,54	0,25	0,29	0,42	0,46	0,35	0,28	0,48
I_SM_02	0,42	**0,54**	0,51	0,27	0,29	0,43	0,45	0,36	0,28	0,49
O_CL_01	0,95	0,87	0,77	0,30	0,93	**0,96**	**0,96**	**0,96**	0,73	0,93
O_CL_02	0,77	0,76	0,75	**0,81**	0,69	**0,81**	0,76	0,75	0,44	0,76
O_MC_01	0,21	**0,64**	0,21	0,06	0,08	0,10	0,16	0,08	0,11	0,40
O_MC_02	0,39	**0,71**	0,43	0,09	0,15	0,11	0,30	0,10	0,18	0,54
O_RA_01	0,69	0,91	0,37	0,54	0,21	**0,95**	0,63	0,90	0,80	0,75
O_RA_02	0,93	**0,94**	0,61	0,45	0,52	0,90	0,85	0,92	0,75	0,88
O_SM_01	0,37	**0,44**	0,35	0,38	0,25	0,28	0,32	0,23	0,22	0,36
O_SM_02	0,36	**0,51**	0,34	0,37	0,24	0,24	0,29	0,19	0,31	0,35
O_SN_01	0,34	0,72	0,26	**0,96**	0,24	0,92	0,29	0,15	0,04	0,83
O_SN_02	0,34	0,51	0,31	**0,86**	0,18	0,84	0,32	0,04	0,02	0,78
O_SU_01	0,63	**0,83**	0,63	0,74	0,44	0,61	0,65	0,40	0,58	0,77
O_SU_02	0,79	0,74	0,79	0,68	0,75	**0,85**	0,80	0,84	0,44	0,81
AVERAGE	0,66	0,68	0,59	0,50	0,47	0,61	0,66	0,61	0,55	**0,73**
Time Complexity (Average Second Per Image)	0,0154	0,0188	0,0141	0,0127	0,0167	0,0245	0,0164	0,0351	0,0204	**0,0123**

Table 5. Time Complexity Performance Comparison on CTRIN dataset (Average Second Per Image)

	CTRIN	**Caviar**	**FMO (153**
CNT	0.0237	0.0278	0.0830
GMG	0,1752	0,0228	0,1909
GRA	0,0275	0,0118	0,2705
GSOC	0,2530	0,0309	0,4531
LSBP	0,4491	0,0153	0,1656
MOG	0,1120	0,0421	0,1127
MOG2	0,0436	0,0171	0,1295
KNN	0,0508	0,0144	0,0829
TFD	0,0432	0,0092	0,0641
STFD (Proposed)	**0,0158**	**0,0053**	**0,0509**

4. Conclusion

We proposed a new algorithm STFD to detect moving objects without using any post processing methods. The STFD method is basically improves three frame difference by selecting non-overlapping moving object in video frames. The proposed algorithm selects images calculating by the local maximum difference value between image sequences instead of using consecutive frames. We mathematically proved that the moving object is obtained the second selected image applying "AND" (&) operator of frame difference method. We tested our proposed STFD methods on benchmark datasets and the real data collected in our laboratory. We also compared the proposed method with TFD method and background subtraction based traditional moving object detection methods. We observed a remarkable result that the complete information of moving object is more accurately obtained compared with TFD method. We observed from the experimental results that regardless of indoor or outdoor environment, the accuracy of proposed algorithm is quite high where only one moving detected object is available. In the case of multiple detected moving object in the scene, the accuracy decreases and noisy results are obtained. Despite these negative result, the accuracy of STFD is still significantly higher than TFD method.

References

[1] T. B. Moeslund, "Introduction to Video and Image Processing: Building Real Systems and Applications" in Springer Science & Business Media, Heidelberg, NJ, USA: 2012.

[2] M. Ekinci, E. Gedikli, "Silhouette based human motion detection and analysis for real-time automated video surveillance" in Turkish

Journal of Electrical Engineering and Computer Sciences 2005; 13(2): 199-229.

[3] E. Simsek, B. Ozyer, L. Bayindir, G. T.Ozyer, "Human-animal recognition in camera trap images" in IEEE 26th Signal Processing and Communications Applications Conference (SIU); Cesme, Izmir, Turkey; 2018. pp. 1-4

[4] S. S. Sengar, S. Mukhopadhyay, "Moving object area detection using normalized self-adaptive optical flow" in Optik 2016; 127(16): 6258-6267.

[5] E. Simsek, B. Ozyer, G. T. Ozyer, "Foto-kapan Görüntülerinde Hareketli Nesne Tespiti ve Konumunun Belirlenmesi" in Erzincan University Journal of Science and Technology, 12(2), 902-919.

[6] V. Karavasilis, C. Nikou, A. Likas, "Visual tracking using spatially weighted likelihood of Gaussian mixtures" in Computer Vision and Image Understanding 2015; 140: 43-57.

[7] E. Simsek, B. Ozyer, G.T. Ozyer, "Object Detection on Camera-Trap Images with Local Features" in Gümüşhane University Journal of Science and Technology 2019; 9(4), 633-644.

[8] A.J. Lipton, H. Fujiyoshi, R. S. Patil, "Moving target classification and tracking from real-time video" in Proceedings Fourth IEEE Workshop on Applications of Computer Vision; Princeton, NJ, USA; 1998. pp.8-14.

[9] F. Moscheni, F. Dufaux, M. Kunt, "Object tracking based on temporal and spatial information" in IEEE International Conference on Acoustics, Speech, and Signal Processing Conference Proceedings; Atlanta, GA, USA; 1996. pp.1914-1917.

[10] Z. Wang, K. Liao, J. Xiong, Q. Zhang, "Moving object detection based on temporal information" in IEEE Signal Processing Letters 2011; 21(11): 1403-1407.

[11] Y. Zhang, X. Wang, BJPE.Qu, "Three-frame difference algorithm research based on mathematical morphology" in Procedia Engineering 2012; 29: 2705-2709.

[12] Z. Zhu, Y. Wang, "A hybrid algorithm for automatic segmentation of slowly moving objects" in AEU-International Journal of Electronics and Communications 2012; 66(3): 249-254.

[13] M. Yazdi, T. Bouwmans, "New trends on moving object detection in video images captured by a moving camera: A survey" in Computer Science Review 2018; 28: 157-177.

[14] M. Yokoyama, T. Poggio, "A contour-based moving object detection and tracking" in IEEE International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance; San Diego, California, USA; 2005. pp. 271-276.

[15] H. Wu, N. Liu, X. Luo, J. Su, L. Chen, "Real-time background subtraction-based video surveillance of people by integrating local texture patterns" in Signal, Image and Video Processing 2014; 8(4); 665-676.

[16] T. Bouwmans, "Traditional and recent approaches in background modeling for foreground detection: An overview" in Computer Science Review 2014; 11: 31-66.

[17] M. Piccardi, "Background subtraction techniques: a review" in IEEE International Conference on Systems, Man and Cybernetics; Hague, Netheirlands; 2004. pp. 3099-3104.

[18] S. S. Sengar, S. Mukhopadhyay, "Moving object detection based on frame difference and W4" in Signal, Image and Video Processing 2017; 11 (7): 1357-1364.

[19] X. Han, Y. Gao, Z. Lu, Z. Zhang, D. Niu, "Research on moving object detection algorithm based on improved three frame difference method and optical flow" in Fifth International Conference on Instrumentation and Measurement, Computer, Communication and Control (IMCCC); Qinhuangdao, China; 2015. pp. 580-584.

[20] N. Singla, "Motion detection based on frame difference method" in International Journal of Information and Computation Technology 2014; 4(15): 1559-1565.

[21] R. Jain, H. H. Nagel, "On the analysis of accumulative difference pictures from image sequences of real-world scenes" in IEEE Transactions on Pattern Analysis and Machine Intelligence 1979; (2): 206-214.

[22] J. Xiao, H. Cheng, H. Sawhney, F. Han, "Vehicle detection and tracking in wide field-of-view aerial video" in IEEE Computer Society Conference on Computer Vision and Pattern Recognition; San Francisco, CA, USA; 2010. pp. 679-684.

[23] L. Gang, N. Shangkun, Y. Yugan, W. Guanglei, Z. Siguo, "An improved moving objects detection algorithm" in International Conference on Wavelet Analysis and Pattern Recognition; Tianjin, China; 2013. pp. 96-102.

[24] H. Zhang, H. Zhang, "A moving target detection algorithm based on dynamic scenes" in IEEE 8th International Conference on Computer Science and Education; Colombo, Sri Lanka; 2013. pp. 995-998.

[25] J. S. Kulchandani, K. J. Dangarwala, "Moving object detection": Review of recent research trends. in International Conference on Pervasive Computing (ICPC). Pune, India; 2015. pp.1-5.

[26] W. Huang, Z. Liu, W. Pan, "The precise recognition of moving object in complex background" in IEEE Third International Conference on Natural Computation (ICNC 2007); Haikou, China; 2007. pp. 246-252 (2007).

[27] A. Abdelli, H. J. Choi, "A four-frames differencing technique for moving objects detection in wide area surveillance" in IEEE International Conference on Big Data and Smart Computing (BigComp),Jeju, South Korea; 2017. pp. 210-214.

[28] L. Shang, F. You, C. Han, X. Wang, S. Zhao, "Optimization of Three-Frame Difference Method and Improvement of Pedestrian Detection Code Book" in Journal of Physics: Conference Series 2019, August. (Vol. 1302, No. 2, p. 022014). IOP Publishing.

[29] Godbehere, A. B., Matsukawa, A., and Goldberg, K. (2012). Visual tracking of human visitors under variable-lighting conditions for a responsive audio art installation. In American Control Conference (ACC), 2012 (pp. 4305-4312). IEEE.

[30] KaewTraKulPong, P., and Bowden, R. (2002). An improved adaptive backgroundmixture model for real-time tracking with shadow detection. In Video-based surveillance systems (pp. 135-144). Springer US.

[31] Zivkovic, Z. (2004). Improved adaptive Gaussian mixture model for background subtraction. In Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on (Vol. 2, pp. 28-31). IEEE.

[32] Wren, C. R., Azarbayejani, A., Darrell, T., & Pentland, A. P. (1997). Pfinder: Real-time tracking of the human body. IEEE Transactions on pattern analysis and machine intelligence, 19(7), 780-785.

[33] Zeevi, S., Background Subtractor CNT: A Fast Background Subtraction Algorithm (2016), https://zenodo.org/badge/DOI/10.5281/zenodo.4267853.svg

[34] Qian, S., Cascitelli, A., (2017) Improve Background Subtraction with Aggregated Saliency. https://gist.github.com/Masterqsx/f63c2627873c2cdf7be9cb520c7060a7

[35] Zivkovic Z. F. van der Heijden. (2006) Efficient adaptive density estimation per image pixel for the task of background subtraction. Pattern recognition letters, 27(7):773–780.

[36] Guo, L., Xu, D., & Qiang, Z. (2016). Background subtraction using local SVD binary pattern. In Proceedings of the IEEE conference on computer vision and pattern recognition workshops (pp. 86-94).