# Human Detection Framework for Automated Surveillance Systems

**Redwan A.K. Noaman\*, Mohd Alauddin Mohd Ali\*, Nasharuddin Zainal\*, Faisal Saeed \*\***
\* Department of Electrical, Electronic, and Systems Engineering, Universiti Kebangsaan Malaysia, Malaysia
\*\* Department of Information Systems, Faculty of Computing, Universiti Teknologi Malaysia, Malaysia

## Article Info

## ABSTRACT

Vision-based systems for surveillance applications have been used widely and gained more research attention. Detecting people in an image stream is challenging because of their intra-class variability, the diversity of the backgrounds, and the conditions under which the images were acquired. Existing human detection solutions suffer in their effectiveness and efficiency. In particular, the accuracy of the existing detectors is characterized by their high false positive and negative. In addition, existing detectors are slow for online surveillance systems which lead to large delay that is not suitable for surveillance systems for real-time monitoring. In this paper, a holistic framework is proposed for enhancing the performance of human detection in surveillance system. In general, the framework includes the following stages: environment modeling, motion object detection, and human object recognition. In environment modeling, modal algorithm has been suggested for background initialization and extraction. Then for effectively classifying the motion object, edge detecting and B-spline algorithm have been used for shadow detection and removal. Then, enhanced Lucas–Kanade optical flow has been used to get the area of interest for object segmentation. Finally, to enhance the segmentation, some morphological processes were performed. In the motion object recognition stage, segmentation for each blob is performed and processed to the human detector which is a complete learning-based system for detecting and localizing objects/humans in images using mixtures of deformable part models (PFF detector). Results show enhancement in each phase of the proposed framework. These enhancements are shown in the overall performance of human detection in surveillance system.

*Corresponding Author:*

Redwan A.K. Noaman
Department of Electrical, Electronic, and Systems Engineering,
Universiti Kebangsaan Malaysia,
43600 Bandar Baru Bangi, Selangor, Malaysia.
Email: redwan_eng@yahoo.com

## 1. INTRODUCTION

Human detection is currently one of the active research topics in computer vision. Wide range of applications for tracking and behavior recognition were studied by many researchers. Automated system for estimating and tracking moving objects have received a lot of attention from industries and academia for its potential applications in the fields of surveillance and engineering such as video surveillance, content based image retrieval, and gait recognition [1-4]. Current active research in this area focuses on three main stages: low level (Detection), intermediate level (Tracking), and high level (Behavioral Analysis) [5]. This paper focuses on proposing a framework in order to enhance human detection stage. The efforts have been done in visual surveillance researches aimed to provide robust and efficient automated surveillance system. The

purpose of using the surveillance system is not only for monitoring subjected area, e.g. patients and children, but more importantly the analysis of the movement to understand suspicious behavior such as terrorist attacks, shop lifting, and robbery. Moreover, movement tracking is used to train machines to imitate different movement behaviors of biological objects such as flying and maneuvering. Unfortunately, human detection in surveillance systems suffer from two significant problems: effectiveness and efficiency. Effectiveness is defined as the ability of a system to classify the human objectsaccurately. The problem of effectiveness is due to many environmental noises such as characteristics of lights and shadow objects. The result is misclassification of the human in the scene and consequent high false alarmsand low accuracy. The criterion of efficiency is defined as the ability of the detectors to cope with the requirement of online surveillance systems such as detector speed with less computation overhead.

Human detection and tracking in surveillance applications require many disciplines (i.e. environmental modeling) in order to subtract the background, motion object detection, remove shadows and noises, and to do human classification. In this paper, a holistic framework is proposed to enhance human detection in surveillance systems. To detect moving objects, the first stage involves dividing the scene into two regions, foreground and background. The foreground contains only events of interest, and the background is relatively unchanging over time. Detection of moving objects from a video sequence is the first stage of the extraction of visual information. For video surveillance researches, background subtraction techniques are usually used for detecting motion in many real time tracking systems and video surveillance applications [1], [2], [6].

Many motion detection and tracking algorithms depend on the process of background subtraction, a technique which detects changes from a model of the background scene. Background subtraction takes advantage of both spatial and temporal cues to identify and track regions of interest (ROI) known as foreground objects like people running or walking, animals, cars and others. To achieve this separation of foreground and background, techniques such as motion detection or optical flow are used [7], [8]. Shadows affect the process of foreground object detection. They cause many problems in the object localization, segmentation, extraction of the objects and tracking [9]. Furthermore, shadows may cause merging of objects with each other; object shapes may be distorted, and as aresult, the background may be classified as foreground and missing objects. Initial background model can be obtained by using a short training sequence in which no foreground objects are present. However, in some monitoring areas, such as public area, crowded corridors, and traffic, it is difficult to control the area being monitored. In such cases, there may be a need to train the model using a sequence which contains foreground objects. An ideal background subtraction could produce good results while foreground regions are in motion during training sequence. There is also a need to maintain a background model to adapt to all possible changes in the monitoring area.
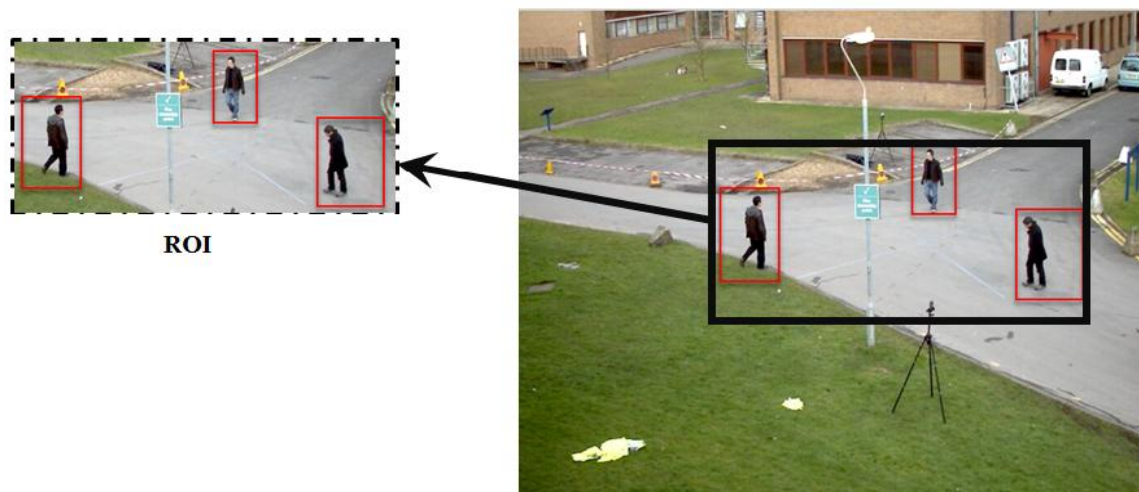


Figure 1. Region of Interest (ROI)

In general, human detection methods require large amounts of computational processing and complex models. In this framework, we proposed an enhanced Lucas–Kanade Optical Flow [10] in order to improve human detection results using a pre-processing stage before initiating a state-of-the-art detector. The

Lucas-Kanade method is widely used as a differential method for optical flow estimation for computer vision. However, the drawbacks of this method include errors regarding the boundaries of moving object where many unwanted vectors appear due to lighting and camera noise [11]. Researches in this area are interested in low level (detection), intermediate level (tracking) and high level (behavioural analysis) such in [12]. Instead of analyzing the whole frame, the enhanced Lucas–Kanade technique was used to obtain a Region of Interest (ROI) as shown in Figure 1. ROI is used as the input for the human detector. This enhancement reduced the processing time and increased the accuracy of the detector.

## 2.    RESEARCHMETHOD

The proposed framework can be described in three main phases as shown in Figure 2. First phase is background model which aims to extract the foreground part of the picture that contains the moving objects. The second phase is to effectively extract the moving objects by designing shadow removal model. Shadow removal model is responsible for optimizing the extraction process of the moving objects so that human objects can be effectively recognized. The third phase is to efficiently recognize the human objects among the extracted objects. For this purpose, an enhanced version of Lucas–Kanade optical flow technique is proposed to effectively and efficiently detect human objectsinthe video stream.
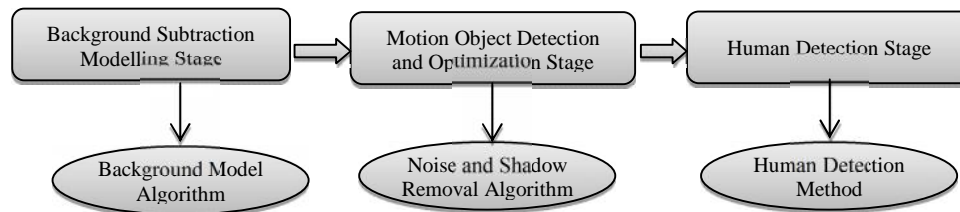


Figure 2. Enhanced Human Detection Framework

### 2.1. Background Subtraction Algorithm

Background subtraction is a method that takes advantage of both spatial and temporal cues to identify and track regions of interest called foreground object like people running or walking, animals, cars, etc. By comparing incoming image frames to a reference image, regions of the image which have changed are efficiently located. If these regions of interest can be detected precisely and effectively, then subsequent image processing stages will be presented with a much limited processing area within an image. This reduction will lead to better efficiency, accuracy and computational cost for the complete vision system. Currently, various techniques employ background subtraction [1, 2, 5, 13-15]. Most of these techniques use a background reference image to perform background subtraction. This reference image is obtained after the background is modeled mathematically. In the final step, the current image is subtracted from the reference image to produce a mask that highlights all foreground objects. The process of image acquisition, background modeling and finally subtracting the current image from the background reference image is implemented in what is known as the background subtraction algorithm. Initial background model is often obtained by using a short training sequence in which no foreground objects were presented. However, in some monitoring areas, such as public area, crowded corridors, or traffic, it is difficult or impossible to control the area being monitored. In such cases there is a need to train the model using a sequence which contains foreground objects. An ideal background subtraction could produce good results while foreground regions are in motion during training sequence. There is also a need to make maintenance for background model to adapt all possible changes in the monitoring area.

In this paper, in order to perform background model, a modal algorithm is developed based on brightness or the intensity value of the image in the current frame. The value that appears most often indicates the need for more redundancy during training sequences to represent the background model for that pixel. Let F represent a function that returns vector containinga number of frequent brightness values of a pixel within a sequence of t image, $B^t(z)$ is the brightness value of pixels at location z within the image sequence t. The initial background model M(z) for a pixel at location z can be obtained from the equation (1).

$$M(z) = \text{Most frequent value } F\left(B^t(z)\right)] \rightarrow \qquad\qquad \text{Eq. (1)}$$

where $F(B^t(z))$ represents a vector containingfrequency of intensitythat occurred in z location on image sequence t.  Figure 3 shows pseudo code for  the proposed modal algorithm.

```
Input(current_frame), Output(Foreground_Image)
Intilization
Set sliding_window_size = 20
Define History_Map (sliding_window_size)
Begin
For each frame is received from video do
If current_frame_no < sliding − window − size then
Foreground_Image = current_frame − previous frame
History_Map = History_Map_Insert(current_frame)
Else
//Calculate the background image from equation (1)
Background_Image = Get Most frequent intensity value[History_Map]
Foreground_Image = Current_Frame − Background_Image
//Update the history map to maintain Background
        History_Map = History_Map_Remove(current_frame − sliding_window_sisize)
History_Map = History_Map_Insert(current_frame)
End if
End if
End for
```

Figure 3. Pseudo Code for Modal Algorithm

Figure 4 shows the history map in which the first cell, labeled D, indicates the number of different brightness values taken for the pixel at location z during the training sequence. The second two cells concern the pixel at location z which are labeled A1 and B1respectively. Cell A1 indicates brightness value, whereas cell B1 shows the number of redundancy of that brightness value during the training sequence. The same process will continue for every frame for the rest of the training sequence, where every two cells are used to store the same information for other brightness values taken for the pixel at location z and their redundancy numbers respectively unless the pixel has brightness value already taken. The size of the array list assigned to the pixel at location z is dynamic because it depends on the number of possible different brightness values that can be taken for that pixel for all frames during the training sequence. The same process is repeated for each pixel and stored in the history map.
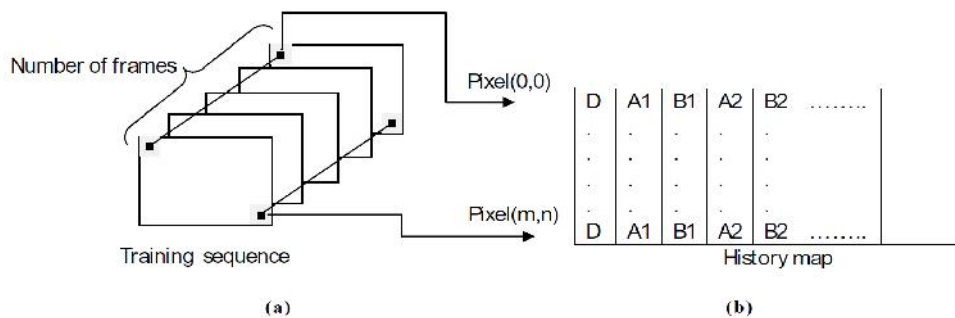


Figure 4. (a) Image Sequence (b) History map

Pixel intensity is the most commonly used feature in background modeling. The basic method is to monitor the change of the intensity of the pixel in sequential frames. However, noise, camera, jitter, automatic adjustment of the camera, illumination change, shadow, dynamic backgrounds and sleeping objects cause change in pixel intensity and leads to reduce the performance of the background subtraction models.

## 2.2. Noise and Shadow Removal Model

The next stage after getting the background reference image (*BR*) from previous stage is to deal with the shadow in the frame sequence. Shadows are cast on the background and change its color information which is why a Blue band from RGB color models has been chosen to be the main band. The RGB color model shows more significant shadow in the Blue band. First, the Blue band of the background reference

image $BR_{blue}$ is taken and moved to normalization process. Then, the Blue band of the current frame $CF_{blue}$ is taken and subtracted from $BR_{blue}$, the resultimage is $S_{blue}$. Then, a threshold applied for $S_{blue}$ to remove noise and the result image is $SB_{th}$.

Assume that the edges of the objects are not changed because objects often have strong edges than its shadow. Depends on that, a Prewitt edge filter is applied to get the edge information from $SB_{th}$. The Prewitt filter is chosen because it produces clear edges and it is more reliable to show the regions where the intensity changed.

For edge detection, the edges of the objects suffer from some disconnected points. B-spline algorithm is used in order to overcome this problem. A spline curve is a sequence of curve segments that are connected together to form a single continuous curve. For this reason, the B-spline technique has been used to reconstruct the edges for the objects which may disconnect and distorted through previous stages for shadow removal. B-spline algorithm goes through the whole image using a special kernel [15*30] and determines maximum number of edge points in each kernel, if the number of points is smaller than a certain threshold, means that there is a small connectivity, then B-spline works and connects the points in that kernel. If the number of points is greater than a certain threshold, means that there are more connectivity and common borders within that kernel, then assume the edges in that kernel is connected properly and no need to use B-spine there. The same steps are repeated until the special kernel moves to the end of the image.

### 2.3. Human Detection Method

Next stage is to propose a method to enhance the performance of the human detection mechanisms. The Lucas–Kanade optical flow technique [10] is enhanced and used to improve the human detector speed and accuracy. The Lucas–Kanade method is commonly used as a differential method for optical flow estimation in computer vision. It uses least squares criterion to solve basic optical flow equations for all the pixels in a neighbourhood. It is a purely local method assumes that the flow is constant in local neighbourhoods of the pixel under consideration. Its advantages are that it can make very fast calculations and accurate time derivatives. The disadvantage of the Lucas-Kanade method is the creation of errors regarding the boundaries of moving object.

As shown in Figure 5, every two frames were processed through the enhanced optical flow technique to get motion vectors (map of vectors). Then, Gaussian special filter was used to reduce unwanted motion vectors to avoid wrong segmentation. In order to use this new map to do segmentation and obtain the Region of Interest (RoI), the new map of motion vectors were converted to binary blobs. Furthermore, some morphological processes were performed to enhance segmentation stage by removing isolated noises. Finally, segmentation for each blob was performed and processed to human detector.
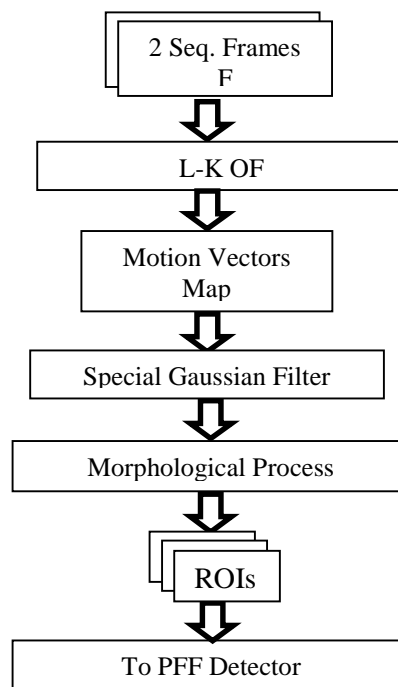
Figure 5. Steps for enhancing motion object detection performance

To achieve an efficient and effective human detection, PFF detector [12] is used to recognize human motion among all motion objects in the scene. PFF is described as a complete learning-based system for detecting and localizing objects and humans in images using combined multi-scale deformable part models. PFF has the ability to represent highly variable object classes and achieve state-of-the-art results in PASCAL object detection challenges. While deformable part models have become popular, their value had not been demonstrated using difficult benchmarks such as PASCAL with PETS datasets [19]. The PFF detector relies on new methods for discriminative training with partially labelled data. A margin-sensitive approach combined with data-mining hard negative examples and formalism is known aslatent SVM. A latent SVM is semi-convex and the training problem becomes convex once latent information is specified for the positive examples. This leads to an iterative training algorithm that alternates between fixing latent values for positive examples and optimizing the latent SVM objective function [12]. Different researches have been done toward pedestrian detection such as in [20-22].

## 3.    RESULTS AND DISCUSSION

The proposed environmentally model namely Modal algorithm produced encouraging results for a more robust, reliable and applicable approach for real-time background modeling for surveillance on both short and long time period through many video sequences which wereused to test it. A comparison between Modal, median and mean is shown in Figure 6 and Table 1, where one scene was chosen from CAVIAR datasets (a man walking in a hall stopped for a while and then walked back again) and the number of error pixels plotted. CAVIAR dataset contains a number of video clips which recorded the acting out of different scenarios of interest. These include people walking alone, meeting with others, window shopping, entering and exiting shops, fighting and passing out and last, but not least, leaving a package in a public place [16]. The plot in Figure 6 and Table 1 also shows that the Modal algorithm gives fewer error pixels compared to median and mean. The advantages of Modal algorithm over median and mean can be summarized as follows:

(I)    Modal algorithm is more flexible and applicable for real events because it depends on the highest redundancy of the brightness values while median filter requires every pixel to appear more than fifty percent during training sequence.

(II)   Modal algorithm overcomes the problem of blinding pixels which appear in mean algorithm.

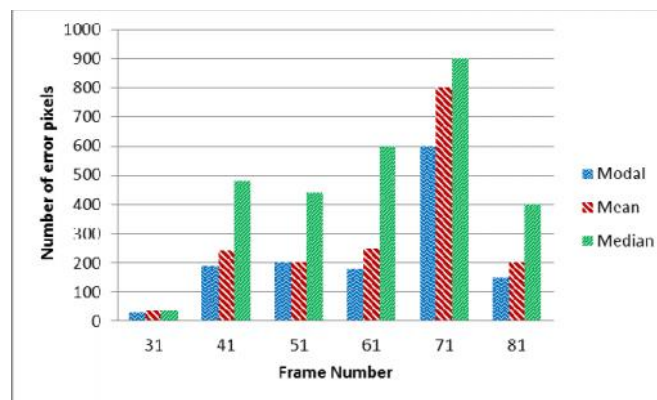(III) Modal algorithm gives fewer error pixels if the object stooped for a long time.



Figure 6. Comparison between Modal, median and mean

Table 1. Number of error pixels

| Threshold | Maximum number of error pixels |
|---|---|
| Modal | 892 |
| Median | 955 |
| Mean | 960 |

To evaluate the effectiveness of the proposed shadow removal method, the proposed method was compared with two methods, Horprasert's method [17] and Conaire's method [18]. Horprasert's method detected the shadow by proposing a computational color model that separates the brightness component from

the chromaticity component. Shadow is identified as pixels that have similar chromaticity but lower brightness than the corresponding pixels in the background image.

Conaire's method is based on the maximization of the agreement between two independent shadow detectors (Luminance & Saturation) without training data. However, appropriate thresholds were automatically chosen for shadow detection. As shown in Figure 7, the proposed method shows better performance and the advantages of the proposed method can be summarized:

(I)  Proposed method is more flexible and applicable for real events for surveillance
(II) Proposed method gives fewer error pixels
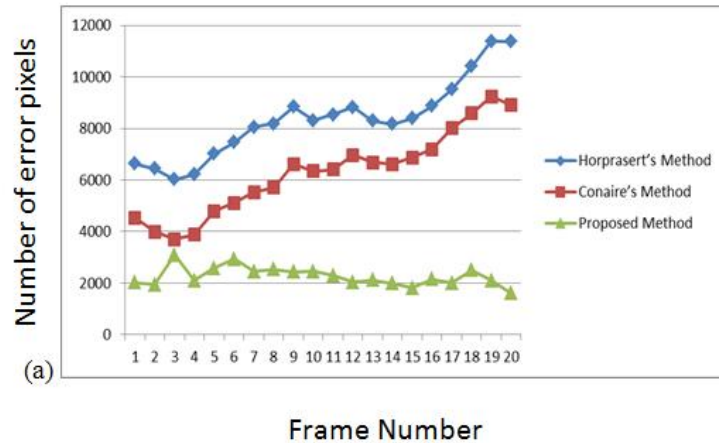(III) Proposed method requires less processing time because it works without the need for training stage.



Figure 7. Shadow removal performance comperison

In the final stage, the enhanced Lucas-Kanade optical flow technique was developed to improve human detection in terms of speed and accuracy. An experiment was performed using a benchmark dataset (PETS 2006) [19] that contained 744 frames. This database has been recognized as a standard in video surveillance environments that use public spaces such as shopping malls, train stations, airports, and outdoor parks. We combined object segmentation output with a human detector using an optical flow algorithm. The proposed technique used the optical flow to find the area of interest to complete object segmentation and used those results as an input for the human detector. This technique has been developed to be used in surveillance systems.

To evaluate the performance of the proposed algorithm, it was compared with the PFF standard algorithm following a Precision-Recall curve. A Precision-Recall curve is a trade-off between Precision and Recall. Precision is the number of true positives divided by the total number of elements labelled as belonging to the positive class (Equation 2). Recall is defined as the number of true positives divided by the total number of elements that actually belong to the positive class (Equation 3).

$$Precision = \frac{True\ Positive\ (TP)}{True\ Positive\ (TP) + False\ Positive\ (FP)} \rightarrow \qquad\qquad Eq.\ (2)$$

$$Recall = \frac{True\ Positive\ (TP)}{True\ Positive\ (TP) + False\ Negative\ (FN)} \rightarrow \qquad\qquad Eq.\ (3)$$
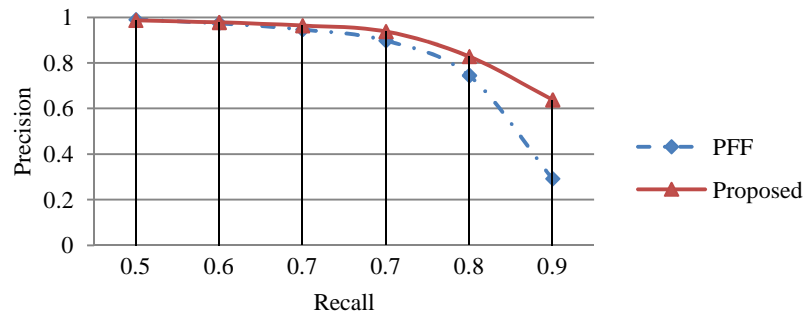
Figure 8. Human Detection performance comparison using Precision-Recall curve

An experimental test was performed for the proposed technique using PFF detector thresholds (0, -0.3, -0.5, -0.7, -1.0 and -1.5). As shown in Table 2, the number of False Positive (FP) decreased, which lead to higher precision and greater accuracy. The implemented method had a maximum true detection range once we increased the threshold. Additionally, the false positive for the proposed method was less than it was for the standard PFF detector. An average calculation was performed to determine the enhancement value for proposed method in terms of accuracy. Figure 8 shows performance comparison between the proposed detection method and the PFF detector in terms of precision and recall.As shown in Table 3, the results reveal that the proposed method is 37% faster than PFF standard.

Table 2. Comparison between the original PFF detector and the proposed method

| | nP | TP | FP | Precision | Recall |
|---|---|---|---|---|---|
| **Threshold = 0** | | | | | |
| PFF detector – Original - | 1165 | 708 | 7 | 0.99 | 0.61 |
| OF-PFF detector (The proposed Method) | 1165 | 634 | 8 | 0.987 | 0.54 |
| **Threshold = -0.3** | | | | | |
| PFF detector – Original - | 1165 | 793 | 21 | 0.974 | 0.68 |
| OF-PFF detector (The proposed Method) | 1165 | 733 | 16 | 0.978 | 0.63 |
| **Threshold = -0.5** | | | | | |
| PFF detector – Original - | 1165 | 851 | 48 | 0.947 | 0.73 |
| OF-PFF detector (The proposed Method) | 1165 | 786 | 29 | 0.964 | 0.68 |
| **Threshold = -0.7** | | | | | |
| PFF detector – Original - | 1165 | 900 | 102 | 0.898 | 0.77 |
| OF-PFF detector (The proposed Method) | 1165 | 856 | 57 | 0.938 | 0.74 |
| **Threshold = -1.0** | | | | | |
| PFF detector – Original - | 1165 | 1000 | 343 | 0.745 | 0.86 |
| OF-PFF detector (The proposed Method) | 1165 | 941 | 195 | 0.828 | 0.81 |
| **Threshold = -1.5** | | | | | |
| PFF detector – Original | 1165 | 1079 | 2616 | 0.292 | 0.93 |
| OF-PFF detector (The proposed Method) | 1165 | 1042 | 592 | 0.638 | 0.89 |

nP = Total number of objects in Ground truth
TP = True Positive
FP = False Positive

Table 3 Execution times for the original PFF detector and the proposed method

| Threshold | PFF (sec) | Proposed method(sec) |
|---|---|---|
| 0 | 2.337*1000 | 1.86*1000 |
| -0.3 | 2.942*1000 | 1.54*1000 |
| -0.5 | 3.203*1000 | 1.675*1000 |
| -0.7 | 2.272*1000 | 1.74*1000 |
| -1.0 | 2.296*1000 | 1.76*1000 |
| -1.5 | 2.366*1000 | 1.49*1000 |

## 4. CONCLUSION

In this paper, a framework for enhancing the performance of human detection in surveillance systems is proposed. In this framework, a number of models have been proposed in each stage of the detection process. Modal algorithm has been used for background initialization and extraction. Then, B-spline algorithm has been used for shadow detection and removal. Next, enhanced Lucas–Kanade optical flow has been used to get area of interest to do object segmentation. To enhance the segmentation, some morphological processes are performed. Finally, segmentation for each blob is performed and processed to the human detector which is a complete learning-based system for detecting and localizing objects/human in images using mixtures of deformable part models called as PFF detector. In future work, we will investigate the effectiveness and efficiency of tracking humans in video streams.

## REFERENCES

[1] Benenson, R., *et al.*, "Pedestrian detection at 100 frames per second. In Computer Vision and Pattern Recognition (CVPR)", *IEEE Conference on* 2012 (pp. 2903-2910).

[2] Doll´ar, P., *et al.*, "Pedestrian detection: a benchmark". *IEE Conf. on Computer Vision and Pattern Recognition*, Miami Beach, FL, USA

[3] Tafazzoli, F., and Safabakhsh, R., "Model-based human gait recognition using leg and arm movements", *Engineering Applications of Artifi- cial Intelligence* 23 (2010) 1237–1246.

[4] Hajer Fradi, *et al.*, "Spatio-temporal crowd density model in a human detection and tracking framework", *Signal Processing: Image Communication*, Volume 31, February 2015, Pages 100-111, ISSN 0923-5965

[5] Ger´onimo, D., *et al.*, "Survey of pedestrian detection for advanced driver assistance systems",*IEEE Trans. on Pattern Analysis and Machine Intelligence* 32(7), 1239–1258, 2010.

[6] Zhang, P., *et al.*,. "Moving people tracking with detection by latent semantic analysis for visual surveillance applications", *Multimedia Tools and Applications*, 68(3), 991-1021. 2010.

[7] Lihizelnikmanor, 'Motion Estimation', Cambridge University, year.2008.

[8] Huston SJ., *et al.*, "Visuomotor Transformation in the Fly Gaze Stabilization System", *Bielefeld University*, Germany. 2008

[9] Dalal, N., and Triggs, B., "Histograms of oriented gradients for human detection". *IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2005*. Vol. 1, pp. 886-893, 2005

[10] Lucas, B.D. and Kanade, T., "An iterative image registration technique with an application tostereo vision", *Proceedings of the 7th International Joint Conference on Artificial Intelligence*, IJCAI, 1981, Vancouver, BC, Canada, Vol. 2, pp.674–679. 1981

[11] Turaga, P., *et al.*, "Machine Recognition of Human Activities: A Survey", *IEEE Transactions on Circuits and Systems for Video Technology* , vol. 18, pp. 1473-1488, 2008.

[12] Felzenszwalb, P. *et al.*, "Object detection with discriminatively trained part-based models". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(9), 1627-1645. 2010.

[13] Horn, B.K., and Schunck, B.G., "Determining optical flow". *Technical Symposium East1981* (pp. 319-331). *International Society for Optics and Photonics*.1981.

[14] Huang Z, Xu Y, *et al.*, "Method for Opposing Flow Abnormal Event Detection Based on Optical Flow". *Video application and project*, 2011, 35(13):102-105.

[15] Moeslund, T.B., *et al.*, "A survey of advances in vision-based human motion capture and analysis". *Computer vision and image understanding*, 104(2), 90-126. 2006.

[16] CAVIAR dataset, http://homepages.inf.ed.ac.uk/rbf/CAVIARDATA1/, retrieved at 11/11/2015.

[17] T. Horprasert, *et al.*, "A robust background subtraction and shadow detection", *Proceedings of the ASIAN conf. on Computer Vision*, 2000.

[18] Conaire, C, *et al.*, "Detector adaptation by maximising agreement between independent data sources", *Proceedings of IEEE Conf. on Computer Vision and Pattern Recognition*, p.1-6. 2007.

[19] PETS (2006), Ninth IEEE International Workshop on Performance Evaluation of Tracking and Surveillance, http://www.cvg.reading.ac.uk/PETS2006/data.html, retrived at 12/11/2015.

[20] Bazmi, A., *Increasing the Accuracy of Detection and Recognition in Visual Surveillance.* International Journal of Electrical and Computer Engineering (IJECE), 2012. 2(3): p. 395-404.

[21] Feng-Mei, L. and T. Lin-Lin, *An Improved Moving Multi-Human Target Detection Algorithm.* TELKOMNIKA Indonesian Journal of Electrical Engineering, 2013. 11(12): p. 7166-7173.

[22] Ke-yang, C., M. Qi-rong, and Z. Yong-zhao, *Pedestrian Detection Based on Sparse and Low-Rank Matrix Decomposition.* TELKOMNIKA Indonesian Journal of Electrical Engineering, 2014. 12(2): p. 1544-1550.

## BIOGRAPHIES OF AUTHORS

Redwan A.K. Noaman received his BEng in Computerand Control Engineering from the University of Sana'a (Yemen), in 2003. In 2008, he joined as a researcher in the computer and network security group, UKM. He received his MEng in Computer and Communication Engineering from Universiti Kebangsaan Malaysia (UKM), Malaysia, in 2008. Currently, heis working toward his PhD in the Department of Electrical, Electronic, and Systems Engineering, UKM. His current research interests are mainly focusedon image processing and computer vision with emphasis on human detectionand tracking for surveillance.

Mohd Alauddin Mohd Ali is a Professor at the Department of Electrical Electronic and Systems Engineering, Faculty of Engineering and Built Environment, Universiti Kebangsaan Malaysia (UKM), Bangi, Malaysia. He was also the Head of the department and Deputy Dean of the faculty for the periods 1995–1997 and 1997–2002 respectively. He also served as the Director of the Institute of Space Science, UKM for the period 2009–2013. He received his BEng (Electrical), BSc (Mathematics) and MEngSc (Electrical) degrees from the University of Tasmania, Hobart, Australia in 1978, 1979 and 1984 respectively. His PhD degree was from the University of Nottingham, Nottingham, England in 1994. He worked in the area of microelectronics for his Masters degree and medical electronics for his PhD. His research interests include biomedical signal processing, instrumentation, integrated circuit design and testability, and ionospheric studies.

Nasharuddin Zainal received his BEng in Computer Science from Tokyo Institute of Technology (Japan), in 1998, MEng in Communication and Computer Engineering from Universiti Kebangsaan Malaysia (UKM), Malaysia, in 2003, and DrEng degree from Tokyo Institute of Technology (Japan) in 2010. He is currently a Lecturer at the Department of Electrical, Electronic and Systems Engineering, Universiti Kebangsaan Malaysia, Bangi, Malaysia. His main research interests include computer engineering, image and video processing and pattern recognition.

Faisal Saeed is a Senior Lecturer at the Department of Information Systems, Faculty of Computing, Universiti Teknologi Malaysia (UTM), Malaysia. He received his BSc in Computers (Information Technology) from Cairo University, Egypt, MSc in Information Technology Management and PhD in Computer Science from UTM, Malaysia. His research interests are machine learning, data mining, and information retrieval.