

# Review of Human Motion Detection based on Background Subtraction Techniques

Arwa Darwish Alzughabi  
-Taibah University

Department of Computer Science and Information, Community College  
Madinah, Kingdom of Saudi Arabia  
-University of Technology Sydney  
Faculty of Engineering and Information Technology

Hanadi Ahmed Hakami  
University of Technology Sydney

Faculty of Engineering and Information Technology

Zenon Chaczko

University of Technology Sydney  
Faculty of Engineering and Information Technology

## ABSTRACT

For the majority of computer vision applications, the ability to identify and detect objects in motion has become a crucial necessity. Background subtraction, also referred to as foreground detection is an innovation used with image processing and computer vision fields when trying to detect an object in motion within videos from static cameras. This is done by deducting the present image from the image in the background or background module. There has been comprehensive research done in this field as an effort to precisely obtain the region for the use of further processing (e.g. object recognition). This paper provides a review of the human motion detection methods focusing on background subtraction technique.

## Keywords

Motion detection methods, Background subtraction method, Moving object detection

## 1. INTRODUCTION

The general means by which objects in motion are detected within videos from fixed cameras is referred to as foreground detection or background subtraction. This is broadly used as a way to recognize the actions of humans, object tracking, monitoring traffic, computer vision application, and also human computer interaction. For this mechanism, objects in motion, also referred to as the foreground, are required to be disjoined from the static information. This static information is referred as the background, and the background subtraction approach is the process frequently used for this [2]. The image, which is located in the background, ought to be an illustration of the setting with immobile objects, and should be updated regularly in order to adjust to the changing geometry settings and luminance conditions. The notion of background subtraction has been developed to mean more than its actual meaning with challenging models. As a matter of fact, this process has greatly developed into a common and popular method for

detecting motion. This innovation utilizes the contrast between the background and current image to determine the region of motion. It also has the ability to supply data which includes information with regards to the object. The easiest method is by using the following equation for the application of differences between the background and current frame.

$$|I(x, y, t) - I(x, y, t - 1)| > Th \quad (1)$$

Where the previous frame, also known as the background frame, is  $I(x,y,t-1)$  and deducted from the present or current frame,  $I(x,y,t)$ . The threshold ( $Th$ ) is then applied to the complete difference in order to receive the foreground mask. The main parameter in the thresholding method is the preferred threshold value. This can be selected to be either automatic or manual. The justification of this method is the detection of the foreground or objects in motion from the compression between the background frame and input video frame. Nonetheless, any system which detects motion established on background deduction considers the following:

- Background deduction should separate the important objects upon initial appearance in a setting.
- A suitable pixel-level criterion needs to be established. A pixel that meets the satisfaction of this criterion is considered background and is disregarded.
- The model of the background should adhere to light within the setting and the progressive and immediate changes within the background.
- Background models need to consider the differences in spatial scales.
- The shadow regions that are protruded by the objects in the foreground are detected as objects in motion.
- Numerous objects shifting within a setting for short and long intervals.

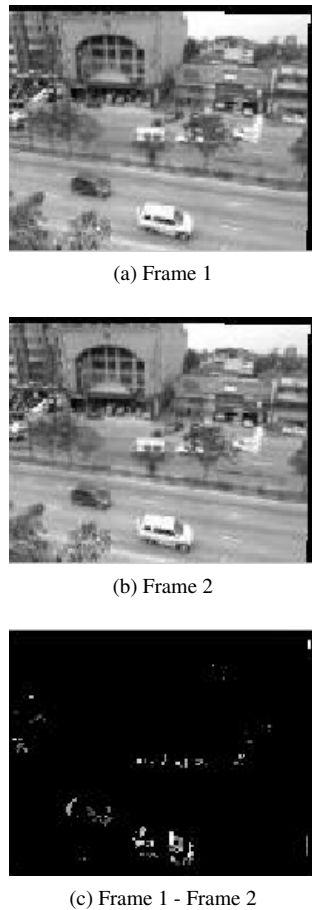


Fig. 1: Temporal Difference Method  
Source [10]

## 2. MOTION DETECTION METHODS

Motion detection methods are classified according to the method of finding a moving object [8]. Different motion detection methods are described as follows:

### 2.1 Temporal difference

Frame difference is the easiest model of background deduction. It attributes to taking the previous frame and then deducting the current frame. In the event that the pixel-value for a pixel given surpasses the threshold ( $T_h$ ) value, that specific pixel is treated as though it is a part of the foreground [3]. Temporal difference has the ability to adapt greatly notably for a widespread group of dynamic environments. Despite this, there is a common difficulty in deriving an accurate breakdown of an object in motion, and there is a fault within the obtained detected movement. Fig. 1 shows the result of using the temporal difference method.

### 2.2 Optical flow

The optical flow process is put in place in order to compute the field of the image optical flow and also to administer clustering processing in keeping with the distribution of the optical flow

nature of the image [4]. The optical flow process aids in gaining accuracy and

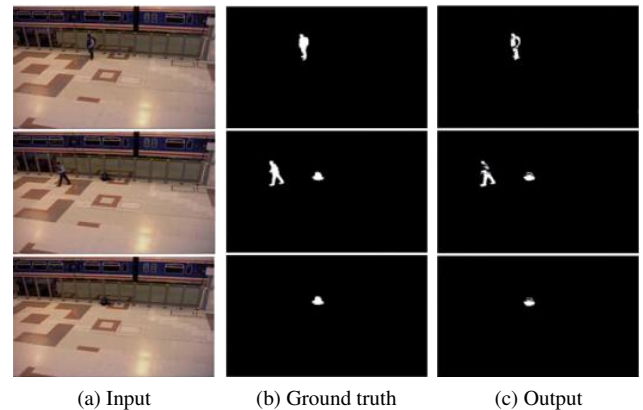


Fig. 2: Background subtraction method  
Source [11]

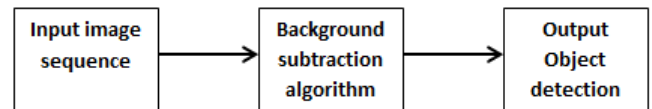


Fig. 3: system block Diagram

adjacent to the background. Despite this, owing to the process inadequate anti-noise performance and noise sensitivity, it is appropriate for studies with the demand for real-time detection.

### 2.3 Background subtraction

Background subtraction is a general technique for the separation of motion within fixed images. It detects regions in motion by deducting the needed current image pixel-by-pixel from an attributed background image. This is established by equating multiple images from the start. The main point of background subtraction is the detection of objects in motion. This is done by initially setting a background, then deducting the present or current frame which contains the objects in motion that are to be detected. This process is easy and it acts in accordance to a simple protocol. It accurately aids in obtaining features of target data, nevertheless, it had sensitivity towards small changes within the external environment. Hence, its only usage is within situations in which the background can be predicted or is known [4]. Fig. 2 shows the background subtraction method.

## 3. MOTION DETECTION SYSTEM DESIGN

The usage for background subtraction is detecting objects in motion that are within a frame sequence. In this detection, there are two images, both having the same sizes preferably, acquired from within the video. One of the images is initially set as the background image to which the object in motion is missing. The object is contained in the other image, which is referred as the

current image. Both images individually have two models; one is the background and the other the foreground model. Fig. 3 depicts a diagram of the system block. Within the background model, the object in motion is absent; on the other hand, the object in motion is present for the foreground model. The initial stage for detecting motion is the initialization of the image. This process initially sets the background image. For instance, within the video, the frame count with regards to the time. One of these frames is initially set as the background image by means of an assumption. Therefore, initially setting the background is an important pre-processing procedure for detecting motion. The median filter carries out the pre-processing which reduces image noise. This is carried out on each frame. When this process is finished, the background subtraction algorithm is given the frames. The image deducted is partitioned by the use of the threshold.

## 4. BACKGROUND SUBTRACTION METHOD

### 4.1 Background Image Initialization

There are numerous other means to gain the first background image. Some of these ways include the background being the first frame, the common brightness of the pixel of first few frames being the background or utilizing background image series lacking the faith of the substances in motion to guess the parameters of the background model. Of the various previously stated methods, the most common method is used for the initialization of a background image [3]. However, there are underlying issues that will be present; this can be eliminated by the median method. Below shows the equation:

$$\beta_{init}(x, y) = \text{median}fk(x, y)_{k = 1, 2, \dots, n} \quad (2)$$

beta init is depicted as the initial background while n is depicted as the sum of the selected frames.

### 4.2 Background Update

The background model has a great ability to adapt to the changes of time. There needs to be an update for the background in order to extract the object in motion. When detecting, the pixels are considered to be a part of the object in motion. They manage the original background grey values and are not updated. As it relates to the pixels that are considered the background, the following equation is used to update the background module:

$$\beta_{k+1}(x, y) = \beta\beta_k(x, y) + (1 - \beta)F_k(x, y) \quad (3)$$

Where beta belongs to (0, 1), a renewed parameter, and  $F_k(x, y)$  is referred to as the pixel grey value within the existing framework.  $\beta_k(x, y)$  is the background value of the existing frame, while  $\beta_{k+1}(x, y)$  is the background value of the other frame. Using the static camera, the background model has the ability to remain fixed for a long time period. This method, with great effect, can avert the unforeseen unpredicted occurrence of the background. Examples include the immediate form of an object within the background, which is not a part of the novel background. Nonetheless, by the pixel grey value update of the background has the possibility to be adapted effectively, through the light impact and variations of the external environment which includes the weather.

### 4.3 Moving Object Extraction

Background subtraction is a regular method which separates the important objects within a frame. The method includes deducting an image containing the object found in the earlier background

image and has no important foreground objects. The region found within the image plane where a noticeable contrast can be seen in these images shows the location of the pixel of the object in motion [6]. Expressed by clusters of pixel, the objects are disjoined from the background image with the use of the threshold method. When the background image  $\beta_k(x, y)$  is retrieved, it is deducted from the existing frame  $F_k(x, y)$ . In the event that the pixel distinction is large compared to the set threshold  $T_h$ , the pixels then appear in the objects in motion. If this does not occur, they appear as background pixels within the frame. After this process, objects in motion are able to be detected. The following shows the expression:

$$Dk(x, y) = 1 \text{ if } |Fk(x, y) - \beta_k(x, y)| > T, Dk(x, y) = 0 \text{ Otherwise} \quad (4)$$

Depicting the complementing human part deriving out of the region of motion is the main purpose of human body detection. The region of motion can be correlated to multiple objects in motion. These include birds, vehicles, swaying trees, and floating clouds [9]. This is further decided on whether or not the object in motion is a human being. The shape features of regions in motion may be used. The following shows the criteria for judging:

- The set threshold is smaller than the object area.
- The region of the objects aspect ratio needs to adhere to the set ratio.

Once these prerequisites are met, the confirmation is made of whether or not the object in motion is a mobile human body.

### 4.4 The Removal of Noise

Owing to the issues relating to the background subtraction method, the variation of the image obtained includes the noise amount in addition to the region of motion. These noises may or may not be contained in the illumination changes or environmental factors during transmitting the video from the camera for further processing. Hence, the removal of noise is needed. The noise may be filtered by adopting a median filter [5]. The region of motion includes flying birds, swaying tress, human beings, non-body parts, and also flying clouds. Further processing may be done using morphological methods [7]. Corrosion operation is done to filtrate the majority of non-body regions of motion and in addition, without injury, maintain the structure of human motion. After this, the inaccessible areas of the image and the intervention of small pieces are disregarded in an effort to gain a precise human motion region.

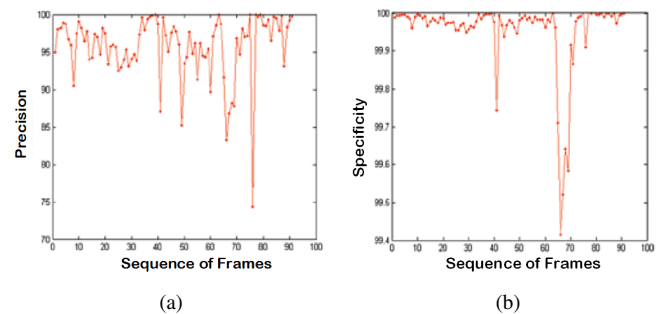


Fig. 4: Performance evaluation: (a) Precision (b) Specificity  
Source [11]

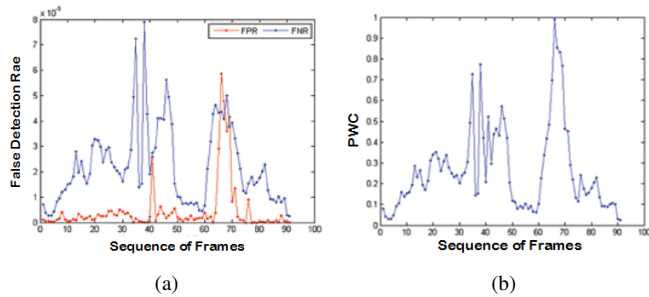


Fig. 5: Performance evaluation: (a) FPR and FNR (b) PWC  
Source [11]

#### 4.5 Understanding Behaviour

After the success of the detection of mobile humans from multiple frames within an image sequence, there is a natural issue to understand the behaviors of humans from image sequences. The understanding of behavior includes description and action identification. This can aid in the enhancement of numerous analysis systems of human motion. This process recognizes and analyses patterns within human motion and produces a top-standard description of interactions and actions. This is also the key area of any future study in human motion analysis.

#### 5. SURVEILLANCE APPLICATION: A CASE STUDY

The Performance Evaluation of Tracking and Surveillance (PETS2006) is in charge of providing the benchmark surveillance dataset that was used for examining the execution of the suggested method [1, 11]. Video footage was recorded from multiple cameras with a PAL standard resolution in all sequences (768x 576 pixels, 25 frames per second), compressed as JPEG image sequences. In this case, the analyzed video consists of a surveillance sequence located at a station depicting multiple moving objects (people) walking forth and back at the station and an abandoned bag, which can be considered, in this case, a security menace. In addition to PETS2006 dataset, a set of ground truth images is also included for quantitative evaluation, recognizing a set of 1,200 frames sequence for testing. After analyzing the exactitude of recognition of this method, results are presented in Fig. 4 and 5. Fig.4 depicts two plots. In the first place, since it is known that for the majority of frames more than 90 percent precision is obtained, the figure represents the results based on precision measure for a set of 900 frames, taking the samples 10 frames apart. Secondly, the figure also represents the results obtained for the same set of frames. In here, all the frames amount to a total of 99 percent scores of specificity. Fig. 5 shows for a set of 900 frames (with samples taken 10 times apart) a false detection rate contrasted against the number of frames. In here, the FPR (false positive rate) is represented by a red line, while the FNR (false negative rate) is represented by a blue line. In the same Fig. 5, a plot of the percentage of wrong classification (PWC) is shown with the same set of frames, concluding with a desired result that the values of PWC are fairly small. Summarizing the results shown in both Fig.4 and Fig.5 illustrate how the proposed method is competent and efficient regarding the accuracy of detection of moving objects.

#### 6. ANALYSIS OF BACKGROUND SUBTRACTION PERFORMANCE

Programmed/automatic visual recognition of items is an essential undertaking for a huge scope of home, business, and modern applications. Video cameras are among the most generally utilized sensors for reconnaissance to savvy spaces for feature conferencing. Moving target identification intends to recognize moving items from the background picture to the persistent feature video. Following moving target items intends to discover different areas of the moving question in the feature. There is a necessity to create calculations for undertaking, for example, moving items identification. Presently utilized strategies as a part of in-motion item location are basically the frame subtraction technique, the foundation subtraction system, and the optical stream strategy. Frame subtraction technique is the contrast between two continuous frames to focus the vicinity of in-motion items. Its estimation is straightforward and simple to create. For a mixture of element situations, it has solid flexibility; however it is generally hard to acquire a complete diagram of in-motion items. Thus, the discovery of moving items is not precise. Optical flow systems ascertain the picture optical stream field and do much preparing, as indicated by the optical stream appropriation components  $f$  picture. This strategy facilitates the complete development data and identifies the in-motion items from the background more superior, because of a huge amount of computation, affectability to commotion, and poor hostile to calmer execution, making this system not appropriate for constant requesting events. Background subtraction is conventional when it comes to identify moving objects in videos. This method is accomplished by identifying the difference among the current frame as well as a background model, or background image, along with an understandable and uncomplicated algorithm. As a consequence, it will also produce accurate information about the specific object when the case of that specific background is known beforehand. This procedure is adequate and efficient to upgrade the development of identifying the moving object. We have mentioned an exact as well as a real-time background subtraction approach in this article, which adequately excludes the collision of changes of light. This simple and intelligible algorithm is capable of identifying the moving object greater as well as it has a wide relevance. This procedure is often used in video surveillance applications.

#### 7. CONCLUSION

This review paper detailed the simple approach for background subtraction methods. During the process, morphological operations and median filter can be used to completely get rid of noise. It has been explained how this method works by block processing, meaning that a fair amount of the present frame is disposed from further analysis in the subsequent background subtraction step. This is done to save both the amount of processed data and processing time, due to the fact that pixel statistics are computed exclusively for those regions of the frames that qualify during block processing. Threshold method has been used to identify moving objects, background initialization, and updating the background in real-time. However, what need to be the main focus of current research are robust methodologies which can be used to get rid of environmental noise. This process is easy and it acts in accordance to a simple protocol. It accurately aids in obtaining features of target data. Nevertheless, it had sensitivity towards small changes

within the external environment. Hence, its only usage is within a situation in which the background can be predicted or is known.

## **8. REFERENCES**

- [1] Pets 2006:benchmark dataset: gth iee international workshop on performance evaluation of tracking and surveillance; <http://www.evg.rdg.ac.uk/pets2006/data.html>.
- [2] T. Bouwmans. Traditional approaches in background modeling for static cameras. *journal*, 2015.
- [3] M Kalpana Chowdary, S Suresh Babu, and Haidar Khan. Fpga implementation of moving object detection in frames by using background subtraction algorithm. pages 1032–1036, 2013.
- [4] Koji Kinoshita, Masaya Enokidani, Masanori Izumida, and Kenji Murakami. Tracking of a moving object using one-dimensional optical flow with a rotating observer. pages 1–6, 2006.
- [5] BSM Madhavi and MV Ganeswara Rao. A fast and reliable motion human detection and tracking based on background subtraction. *IOSR Journal of Electronics and Communication Engineering*, 1(1):29–35, 2012.
- [6] M. Madhusudhan. Human motion detection using background subtraction algorithm. *journal*, 4(2):991–996, 2013.
- [7] PD Mahamuni, RP Patil, and HS Thakar. Moving object detection using background subtraction algorithm using simulink.
- [8] Mr Mahesh C Pawaskar, Mr NS Narkhede, and Mr Saurabh S Athalye. Detection of moving object based on background subtraction. 2014.
- [9] Rupali S Rakibe and Bharati D Patil. Background subtraction algorithm based human motion detection. *International Journal of scientific and research publications*, 3(5), 2013.
- [10] Aresh T. Saharkhiz. Low complexity background subtraction using frame difference method. 2010.
- [11] Soharab Hossain Shaikh, Khalid Saeed, and Nabendu Chaki. Moving object detection: A new approach. pages 25–48, 2014.