Precise Foreground Detection Algorithm using Motion Estimation, Minima and Maxima inside the Foreground Object

Muhammad Nawaz, Prof. John Cosmas, Pavlos Lazaridis, Zaharias D. Zaharis, Hamdullah Mohib, Yue Zhang

Abstract— **In this paper the precise foreground mask is obtained in a complex environment by applying simple and effective methods on a video sequence consisting of multi-colour and multiple foreground object environment. To detect moving objects we use a simple algorithm based on block based motion estimation, which requires less computational time. To obtain a full and improved mask of the moving object, we use an opening-and-closing-by-reconstruction mechanism to identify the minima and maxima inside the foreground object by applying a set of morphological operations. This further enhances the outlines of foreground objects at various stages of image processing. Therefore, the algorithm does not require the knowledge of the background image. That is why it can be used in real world video sequences to detect the foreground in cases where we do not have a background model in advance. The comparative performance results are not only confined to a few conventional performance measures such as precision, recall and area under the curve, and they finally demonstrate the effectiveness of the proposed algorithm.**

Index Terms— Block-based motion estimation, Foreground detection, Morphological Operations, Opening-and-closing-by-reconstruction

I. INTRODUCTION

THE foreground is the more visible and prominent part of the scene in a picture or video. In contrast to the background, the foreground can be defined as that part of the scene in an image, which consists of bits closer to the viewer, while the background refers to the bits at the back or further away from the viewer. The foreground may refer to an image object relatively closer to the camera. This idea is fully explained in this paper under the opening-and-closing-by-reconstruction topic. Foreground detection in a video is the identification of the Region of Interest (ROI), or the identification of the moving objects (foreground) and the static parts (background). Due to its motion, a human is considered as a foreground by the surveillance systems. Therefore, the challenge in detecting a foreground is to fully cover the shape of the moving object in various motion styles, e.g., walking, sitting or jumping.

The foreground detection is the prerequisite step for many video analysis systems such as intelligent video surveillance or vehicular traffic analysis, human detection and tracking, or gesture recognition in human-machine interface and video compression. So far, different algorithms have been proposed but none of them can be considered as a comprehensive solution for different situations and application scenarios. Furthermore, the level of complexity in the foreground detection may depend on the level of complexity of the videos under observation.

The following situations in the analysis of the videos should be considered in order to construct an efficient algorithm for foreground detection. Since, the videos can be of different nature due to the application scenarios involved, the designed algorithm should be efficient enough to capture the details accordingly:

1. In talk shows the background is usually static most of the time, while the foreground consists of moving objects.
2. There can be situations where the background and foreground are both moving at the same time. For example, in the mobile video sequence, it is possible that the objects in the background are moving in the same direction as the ones in the foreground. Similarly, the camera movement might be involved during the capture of the video sequence.
3. There might be a situation where the ROI in the video under analysis is static while the camera, which is capturing the video, is moving. This can happen in the situations of aerial surveys.
4. Foreground detection in rush hours or traffic jams is another situation where the video is analysed for foreground detection, keeping in view the vehicles' as well as the pedestrian's relative movement.
5. Light variation should be considered by the foreground detection algorithm. This is because, the video quality changes with changes in the light intensity. For example, in cases of cloudy weather the video frames might need a

M. Nawaz. is final year PhD student at Brunel University, West London (phone: +44(0) 7411862891; e-mail: m.nawaz@imsciences.edu.pk).

J. Cosmas. John Cosmas, Professor of Multimedia Systems, Associate Editor IEEE Transactions on Broadcasting, Director Wireless Networks and Communications Centre, Director MSc in Advanced Multimedia Design and 3D Technologies, School of Engineering & Design, Brunel University, Uxbridge, Middlesex. UB8 3PH United Kingdom (e-mail: john.cosmas@brunel.ac.uk).

different approach to extract the desired foreground compared to frames in cases of sunny weather. A similar situation can arise when a video is captured in cases where the objects are moving from a dark or semi-dark environment to high intensity light locations.

Colour and texture changes are other important features, which should be taken into account. Since, there are multiple objects in the same frame that might be of the same colour or texture, these objects need to be individually identified by the foreground detection algorithm.

The outline of this paper is as follows. An overview of the foreground detection algorithms is described in Section II. Experimental work is discussed in Section III. Section IV presents the various performance measure results and discussion to analyze the performance of sate of the art algorithms with proposed one solution in terms of accuracy. Finally, the conclusion is presented in Section VI.

II. AN OVERVIEW OF FOREGROUND DETECTION ALGORITHMS

The foreground detection algorithms can be classified into two major categories [47], i.e., the derivative and background subtraction algorithms, as shown in Fig. 1.

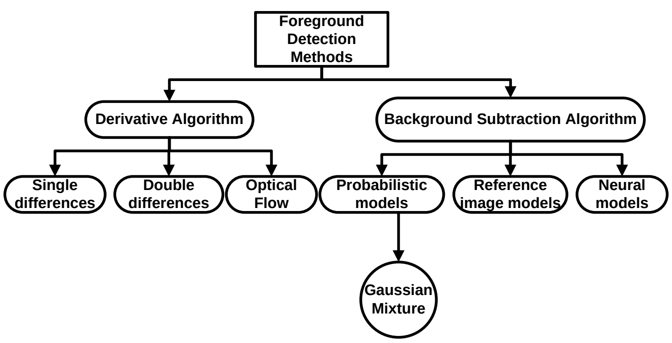


Fig 1.1 Foreground detection algorithms classification

The *derivative algorithms* work under the assumption that the foreground objects compared to the background change rapidly while the background remains static or changes very slowly. The drawback of these algorithms is that they cannot extract accurately the foreground, when the background changes very fast, too.

The *background subtraction* or *frame differencing algorithms* provide one of the most convenient ways for foreground detection [14, 15, 47] due to their simple implementation and processing. In these algorithms, the frame under analysis is compared to (subtracted from) foreground-free frames, as given below in the Equation (1.1), :

$$Foreground = R_i - R_j \qquad (1.1)$$

Where R_i is a foreground-free (reference) image and R_j is the image taken when the foreground object is present. Thus, a simple subtraction between the two images results in the foreground object. In these methods, each video frame is continuously compared to the reference image (background model). The situation where pixels of the current frame deviate significantly from respective pixels of the reference image points to a moving object in the current image. Furthermore, the subtraction algorithms are used to manipulate the obtained foreground pixels for object location and tracking. The main drawback of these algorithms is that in real situations it is not always possible to have foreground-free images. Also, the algorithms encounter problems in several cases of background variation, e.g., cases of camera motion, background that contains shadows, wavering of plant branches or illumination changes. To overcome these problems, the Gaussian function can be applied for optimum results. The Gaussian function describes the distribution of colour in the stable background of an object. This process is performed on each pixel of the object of interest [3, 4]. To follow the changes in the background of the video, the Gaussian model parameters are recursively updated.

The derivative algorithms can be further classified into the following three sub-types:

1) *Single difference algorithms*

These algorithms compare the pixels between the current and previous frames of the video sequence in such a way that pixels, whose difference is significant (difference equal to 1 or 2) and based on pre-threshold, become the background [7-9].

2) *Double difference algorithms*

The double difference algorithms consider the variations in three or more adjacent frames of the video under analysis. One of the advantages of applying these algorithms is that they filter sudden changes occurring due to image noises [10, 11].

3) *Optical flow algorithms*

The optical flow algorithms are primarily based on motion vectors and use the spatiotemporal derivatives of pixel values or block matching techniques [12]. This procedure is capable of detecting the person in a changing background. Thus, the method has the capability of extracting the foreground from complex outdoor scenes that contain non-stationary vegetation [6]. Since this method does not use background subtraction, it produces good results in cases where the background image is not available [16, 17, 21-24]. Obviously, in these cases the traditional background methods fail. The Optical flow models are based on a two-frame differential method for motion estimation. This method estimates the motion between two frames, which are taken at a time interval *t*. Optical flow methods are very useful in pattern recognition, computer vision and other image processing applications.

All of the above methods consider only the changing parts of an image as the foreground. This is not always true and can further cause two types of problems respectively called *foreground aperture* and *false foreground detection*. The case of foreground aperture occurs when the foreground (moving part) is much bigger compared to the background and thus is assumed as the background of the video sequence. This situation might happen because the object in the frame is temporarily still or because it shares the same texture or colour, and thus the motion is only detected in the borders. The case of false foreground detection occurs when there are light variations in the background or very small waving of trees branches.

The background subtraction algorithms can be further classified into the following three sub-types:

1) Probabilistic models

In these models, the background of the video under observation is represented as a probability distribution. The probabilistic representation of the background can be applied by using either a parametric or a non-parametric approach. The parametric approach adopts the Gaussian distributions while the non-parametric uses Kernel Density Estimation (KDE). In these methods, the current frame is initially compared with the background and the probability of each pixel is computed according to the background probability distribution. Then, all those pixels whose probabilities are below a threshold are considered as the foreground.

2) Reference image models

These models consider the background in the form of a single or multiple frames. The comparison is performed between the current and the background reference frames by taking the colour space distance between any two corresponding pixels. The selection of the foreground is based on each pixel distance above a threshold.

3) Neural models

In these models, Artificial Neural Networks (ANNs) are used to identify the foreground in the video sequence. The ANNs are trained by using a set of random frames. After the training procedure, the ANNs can further classify the pixels into background and foreground.

Researchers have recommended background subtraction methods to solve several problems [16-26, 47]. Some popular methods use Gaussian average, temporal median filter, mixture of Gaussian (MoG), Kernel Density Estimation (KDE), Sequential Kernel Density Approximation (SKDA), co-occurrence of image variation and eigen-backgrounds. Comparisons in terms of performance based on essential parameters such as speed, memory and accuracy show that Gaussian average or a temporal median filter is the fastest method provides acceptable accuracy and requires less memory than other methods. The MoG and KDE methods exhibit good accuracy but KDE requires high memory usage. SKDA and KDE methods have almost the same accuracy, but SKDA requires less memory than KDE. The co-occurrence of image variation and eigen-background methods exhibit fair accuracy, and they require reasonable computational time and memory [13].

MoG is one of the most recent methods proposed for foreground detection. This method produces very promising results in outdoor scenes. MoG was initially introduced by Stauffer and Grimson [18] and its improved version was given by Kaewtrakulpong and Bowden [19]. In MoG, the colours of the background objects' pixels are represented by multiple Gaussian distributions. Many researchers have reported that more than two Gaussians can badly degrade the foreground object extraction [4, 5]. The main disadvantages of MoG are the computational complexity of the method and the fact that the variables require careful setting. Thus, the method requires more time in processing. Also, MoG is very sensitive to sudden changes in global illumination and thus produces inaccurate results. Consequently, when the scene is still for a long time, a rapid change in global illumination may turn the whole frame into foreground [20]. Criminisi's algorithm [48] uses depth information to separate the foreground from background object inside an image. Now foreground detection is getting popularity in 3D video area [48, 49] and is a challenging task for the researchers.

Constraints

Most of the foreground extraction approaches are based on the assumptions that stationary objects are included in the background and that their colour and intensity may change gradually with the passage of time. A simple way to overcome this issue and make the colour of the background pixels smooth is the application of an Infinite Impulse Response (IIR) or Kalman filter [1, 2].

III. Approach to Experimental Work

The objective in any foreground detection algorithm is to find areas of the video sequence where motion exists. The next task is to identify sufficiently the mask of the moving object. This second goal is more challenging than the first one.

The motion of objects is estimated by a block-matching algorithm, which finds matching blocks in a video sequence. Such algorithms are the cross search, full search algorithm, spiral search, exhaustive search, three step search, new three step search, simple and effective search, four step search, two dimensional logarithmic search, binary search, orthogonal search, hierarchical search, and diamond search [27, 36]. Each algorithm has its own merits and demerits but their performance is measured by their accuracy and computation time.

In the present study, we adopt the Adaptive Rood Pattern Search (ARPS), which is based on the fact that motion in a frame is generally coherent, i.e., if the macro-blocks around a given macro-block move to a certain direction, this macro-block is highly probable to have a similar motion vector (MV). In the ARPS, each macro-block benefits from the MV of its adjacent left one to guess its own MV.

The ARPS estimates the four endpoints of the four-armed rood pattern of its diamond (Small Diamond Search Pattern (SDSP) or Large Diamond Search Pattern (LDSP)) along with the predicted point (from the neighboring block) of the motion vector (MV) to measure the current block motion tendency. At primary step a minimum SAD (sum of absolute difference) is found and it becomes the center for unit sized rood pattern. The four endpoints of the four-armed rood pattern (in both cases i.e., that of SDSP or LDSP) are then calculated and compared with the SAD to find a new minimum SAD. This is repeated in order to find the minimum SAD at the rood center.

Search pattern has a very important role in searching algorithms and it size has its own significance. A small search pattern is useful primarily for small motion detection and will result in false estimates while probing a large motion vector (MV). In such a case a large search pattern is suitable. Consequently, search pattern size and magnitude of motion vector should be adaptable to the various situations.

In the prediction of accurate MV of the current block the region of support (ROS) and the algorithm to predict the motion vector, are very important. The current block motion vector is predicted from the MVs of the ROS, i.e. the neighboring blocks. Normally, a macro block is expected to be at the same position in the current block as that of the reference block but keeping all the references in memory is increasing the run time. The other solution is to focus on few but the most important blocks around the current block, above, above right, above left and left blocks. The MV of these blocks are used as reference. Further details can be seen in [37].

The next step is the search pattern, where initially adaptive search is performed and then fixed pattern is chosen for local search. The four search points located at the four vertices as in the figure 2.1 depict the rood pattern symmetry.

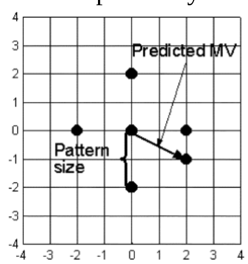


Figure 2.1 Adaptive rood pattern (ARP)

The size of the rood shape is referred to the distance between the center and any vertex point. It has been noticed that the MV distribution in horizontal and vertical directions are higher than that in other directions, [38]. The search can detect fast the motion in the horizontal or vertical directions as these are the most probable motions of cameras. Also, a MV is possible to be decomposed into its horizontal and vertical components. The rood shape can detect the main tendency of motion which is the purpose of the initial search. Summarizing, the adaptive pattern has a rood-shaped pattern (with four vertex points) and a search point which is specified by the predicted MV.

The initial adaptive rood search leads to the final step of local search, avoiding the extra intermediary searches. There are many searching algorithms that can be used e.g. SDSP in Diamond search DS, [39]. The advantage of these algorithms over DS is that if the predicted motion vector is at point (0, 0), it does not waste computational time in LDSP, and it rather directly starts using SDSP. Furthermore, if the predicted motion vector is far away from the center, then again ARPS is saving on computations by directly jumping to that vicinity and using SDSP, whereas DS is wasting time doing LDSP, [40].

In Equation (2.1) below, M_{part} represents the Motion part and S_{part} is the Static part of the foreground object, which is the eventual objective of any foreground detection algorithm.

$$Foreground = M_{part} + S_{part} : \;\; S_{part} \in M_{part} \qquad (2.1)$$

In the Figure (2.2), $\boldsymbol{u}$ is the universal set that contains all the elements being considered in a particular image.

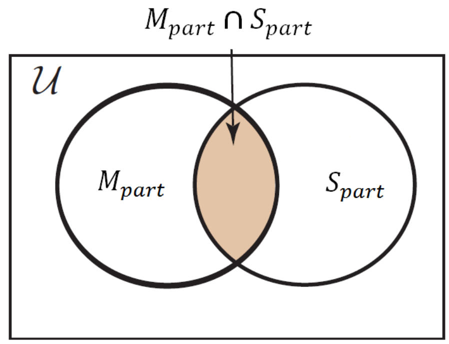


Fig2.2 Static and motion part of a frame

Foreground area up to Equation (2.2), can be easily detectible by our motion estimation technique, with few miscalculated or over calculated areas of motion that will be assumed to be noise. To a greater extent this noise can be reduced by using certain morphological operations.

$$M_{part} \cap S_{part} = \{x : x \in M_{part} \; and \; x \in S_{part}\} \qquad (2.2)$$

Equation (2.3) below is a challenging task, which ultimately covers the full mask of foreground object. For the solution of Equation (2.3), we determine the minima and maxima of the foreground object. The minima and maxima of the foreground object can be determined by morphological operation of opening-and-closing by reconstruction. By minima and maxima of the foreground object we mean the area inside the foreground where the values of intensities are low and high respectively. However this does not cover the background area.

$$Perfect_{foreground} = \{S_{part} \cup (M_{part} \cap S_{part})\} \qquad (2.3)$$

It is well understood that image segmentation in terms of foreground and background separation is among one of the interesting but demanding areas, from the implementation point of view, in the image processing field. However foreground detection is the prerequisite process for many image processing procedures. The present state of the art in foreground detection algorithms does not produce the same good quality results for different types of images [25, 28-32] due to the varying nature of images and end user requirements. For these reasons, the segmentation process is much more difficult when dealing with videos, having numerous frames, having a range of luminance, contrast, texture, color and a varying number of moving objects (ranging from low to high speed). The aforementioned difficulties appear also in the selection of video for the implementation and testing of the proposed foreground detection algorithms. In the existing research on foreground detection, researchers have selected simple videos with a limited number of foreground objects and movement with static background. Moreover, the number of frames selected is always very small, [33-35]. On the contrary, this research is conducted on multi-featured videos in order to test the performance of our algorithms for various types of videos.

Morphological operations (MOs) are used on binary images to remove noise or irrelevant detail. In general, dilation expands, while erosion shrinks the pixel areas with the defined radii or structuring element in the given image respectively. Mathematically, dilation of an image α by factor β is defined as in Equation (2.4).

$$\alpha \oplus \beta = \{x | (\beta')_x \cap \alpha \neq \emptyset\} \qquad (2.4)$$

Dilation has the effect of increasing the size of an object. Erosion of the image α by a factor β is defined mathematically as in Equation (2.5), where α is the image and β is the structuring element and $\boldsymbol{\alpha^c}$ is the complement of $\boldsymbol{\alpha}$.

$$\alpha \ominus \beta = \{x | (\beta)_x \cap \alpha^c \neq \emptyset\} \qquad (2.5)$$

The proposed work first computes motion estimation and then the minima and maxima of the foreground object in the video sequence are determined frame by frame. The motion estimation process is block based, whereas the second process is pixel based. The objective for both processes is to compensate for the missing areas of foreground object. Noise is removed from the original frames using MO, opening-and-closing-by-reconstruction. In order to obtain pixel based foreground, regional minima and maxima were used. For this purpose, MOs are applied to the segmented image for different intensity values, where the lowest and the highest intensities are used to determine the foreground maxima and minima respectively inside each frame. Both minima and maxima are added to obtain a sufficient mask of the object in Figure 2.2(b) and (c). The resultant binary mask is combined by an OR logical operator with block-based motion estimation mask to generate the final binary mask as shown in the Figure 2.2(d).

Figures 2.3 (a), (b), (c) and (d) demonstrate the block-matching estimation result with miss and over-calculated blocks. The foreground object minima, maxima masks are obtained after the opening-and-closing-by-reconstruction process to obtain the full mask of the foreground objects.

Fig. (2.3a) Fig. (2.3b) Fig.(2.3c) Fig.(2.3d)

Figure 2.3 Motion estimation, minima and maxima

Noise Removal

In Figure 2.4(a), the result of the block-based motion estimation on the video sequence obtains segmentation of the foreground objects from background with a large amount of noise. For this purpose, a couple of MOs such as clean, bridge, dilation and erosion are applied to remove the isolated pixels, bridge them if unconnected, expand, and shrink pixels respectively. As a result, a sufficiently noiseless block-based motion estimated foreground is obtained as shown in the Figure 2.4(b) as compared to Figure 2.3(a), but still with some missing areas of foreground objects.

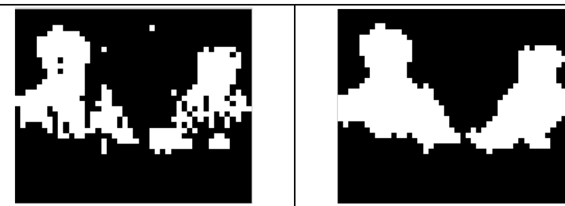


Fig. 2.4(a) Fig. 2.4(b)

Figure 2.4 Motion estimation and noise removal

Figure 2.5, depicts the overall layout of our algorithm, where both motion estimation and opening-and-closing-by-reconstruction operations are applied on the same frames simultaneously. A sufficient mask of the foreground is eventually obtained for the frames under observation Figure 2.3(d).

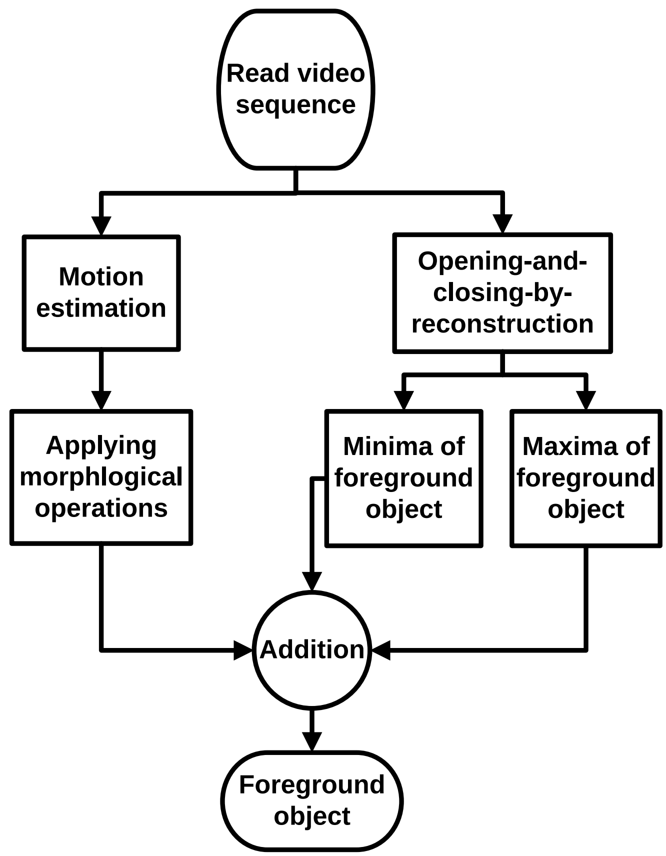


Fig. 2.5 Overall system diagram

Mathematically, MO opening is defined in Equation (2.6), where α is the image and β is the structuring element.

$$\alpha \circ \beta = (\alpha \ominus \beta) \oplus \beta \qquad (2.6)$$

Similarly MO closing is defined in Equation (2.7), where α is the image and β is the structuring element.

$$\alpha \bullet \beta = (\alpha \oplus \beta) \ominus \beta \qquad (2.7)$$

Algorithm to find maxima of the foreground object:

Step 1: Define structuring element (β),
Step 2: Apply MO opening on (α)
Step 3: Apply MO closing on the resultant of step 2
Step 4: MO Reconstruct results from step 2 and 3
Step 5: Apply closing operation on resultant of step 2
Step 6: Dilate reconstructed resultant from step 4
Step 7: Reconstruct complemented results from step 4 and 6
Step 8: Complement resultant of step 7
Step 9: Apply regional maxima operation on step 8

Algorithm to find minima of the foreground object:
Figure (2.6) demonstrates the step by step algorithm for computing the minima of foreground object.

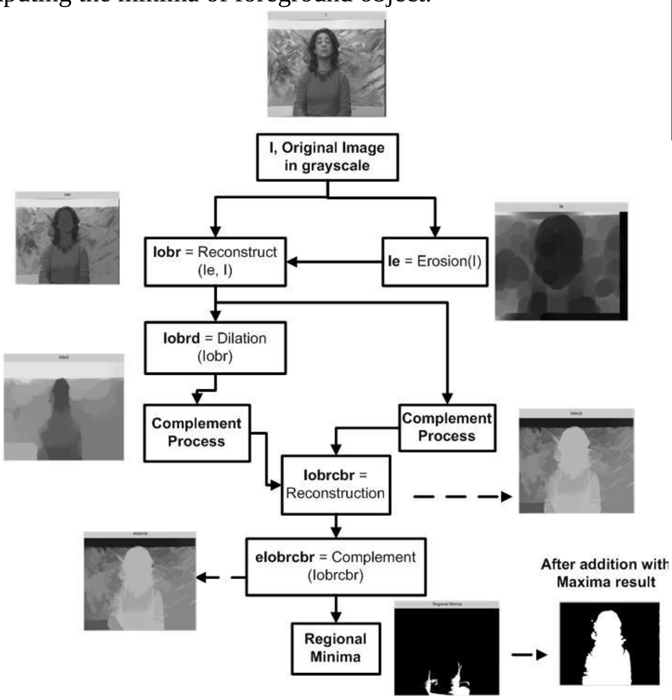


Fig. 2.6, Minima and maxima

Figure (2.7) are the original frames of video sequence, Figure (2.8) shows the ground truth for respective frames and Figure (2.9) to Figure (2.12) demonstrate respective frames foreground detection results by various state of the art algorithms [29-32]. Results of the proposed algorithm for foreground detection are given in the Figure (2.13).

Fig 2.7 Original video sequence

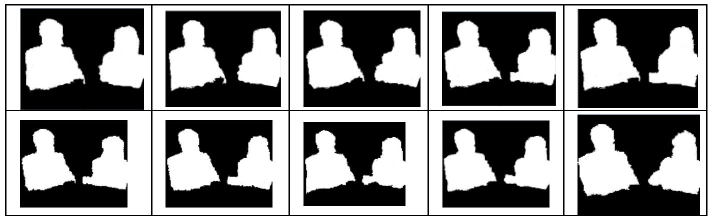


Fig 2.8 Ground truth

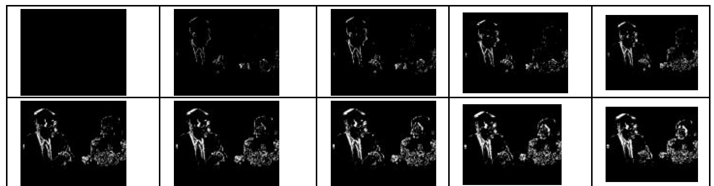


Fig 2.9 Mixture of Gaussian [29]

Fig 2.10 SGMR [31]

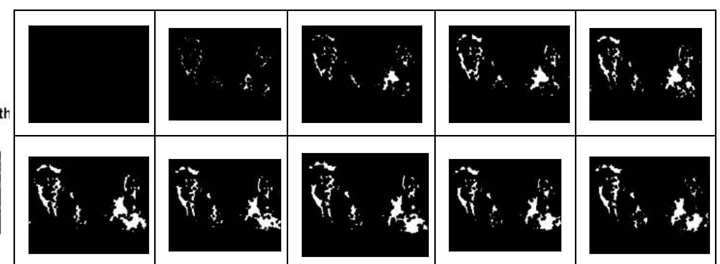


Fig 2.11 Soo Wan Kim algorithm [32]

Fig 2.12 Optical flow [30]

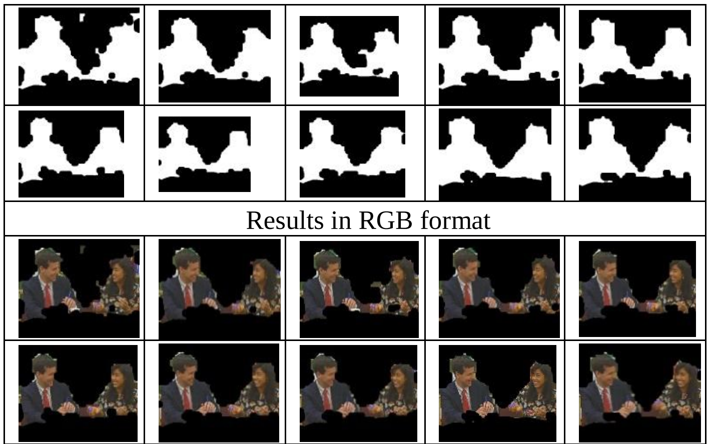


Fig 2.13 Proposed algorithm

IV. PERFORMANCE MEASURE RESULTS

The performance of any foreground detection algorithm can be judged via qualitative or quantitative methods. The qualitative method is applied by a human who judges the visual quality of results based on human visual perception. However, most of the researchers opt for the quantitative method as an accurate tool for performance measurement. Although quantitative evaluation is a difficult and time consuming job, [41, 45], in terms of generating valid ground truth, ground truth is the correct representation that is expected from the proposed algorithm. A second issue is that, ground truth being generated by humans, each human observer can segment differently for the same data at different timings. Another issue is to describe the relative importance of the different types of errors as there are various quantitative methods to compare ground truths with respective candidate binary mask. There are different standard procedures for comparing the ground truth to a candidate binary change mask. In general, the following parameters are involved while calculating different performance measures:

- True Positive (*tp*) refers to the number of foreground pixels correctly detected.
- False Positive (*fp*) refers to the number of background pixels incorrectly detected as foreground or, in other words, the average of false alarms per frame.
- False Negative (*fn*) refers to the number of foreground pixels incorrectly detected as background, or we could say, the average of false misses.
- True Negative (*tn*) refers to the number of background pixels correctly detected.

The above parameters can be seen in factorial form in the Figure (4.1), describing *tp*, *fp*, *fn*, and *tn*, respectively. In this figure, the detected foreground mean result obtained from the proposed algorithm and ground truth foreground is considered to be the perfect result based on human segmented result.

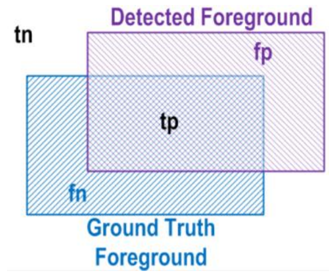


Figure 4.1 Confusion matric variables

In Table 4.1, C1 represents first column elements *tp* and *fp* and C2 represents second column elements *fn* and *tn* of the confusion matrix.

Table 4.1 Confusion matric binary values

Resultant		Ground Truth	Resultant image
C1	*tp*	0	0
	fp	1	0
C2	*fn*	0	1
	tn	1	1

Confusion matrix for binary classification and corresponding array representation

Table 4.2 Confusion matric classifiers

Data Class	Classified as *positive/ detected*	Classified as *negative/not detected*
positive *(pos)/actual object*	true positive (*tp*)	false negative (*fn*)
negative *(neg)/non-object*	false positive (*fp*)	true negative (*tn*)

From Table (4.2), we can derive its mathematical form as given in Equation (6.2).

$$Confusion\ matrix = [C1 \quad C2] \qquad 4.1$$

or

$$Confusion\ matrix = \begin{bmatrix} tp & fn \\ fp & tn \end{bmatrix} \qquad 4.2$$

To quantitatively compare the proposed method, the desired pixels for the foreground objects in the test images were manually labeled and taken as the ground truth. Then the true positive rate (*tpr*) and false positive rate (*fpr*) pixels were computed for the segmentation results. The *tpr* is defined as the ratio of the number of correctly classified object pixels to the number of total object pixels in the ground truth. The *fpr* is defined as the ratio of the number of background pixels but classified as object pixels to the number of background pixels in the ground truth. Obviously, the higher the *tpr* and the lower the *fpr*, the better is the proposed method performance [11].
It is to be noted that in the performance measurements below all values are converted into percentages for more clarity.

Performance Measures and Results Comparison

There are 11 different performance measurements: precision, recall, F-score, specificity, area under the cure, BER%, accuracy, geometric mean of sensitivity and specificity, similarity and false positive rate. With the help of these measures we will also compare our results with state of the art algorithms such as: optical flow [30], Soo Wan Kim approach [32], Mixture of Gaussian (MoG) [29], and the SGM-R algorithm, [31].

1. Precision

To quantify how well the proposed algorithm matches the ground truth some researchers use precision and recall [42, 43]. Precision is also known as Positive Predictive Value (PPV).
Precision is defined by Equation (4.3), and is the measure of how well we have identified the ground truth foreground without misidentifying the background.
fp is the area of miscalculated foreground in the resultant segmented image. The lower its value the greater is the value of precision.
Our precision value is 93.60% as shown in Table (4.1), meaning that we have been able to identify more of the ground truth (intended region foreground) than other techniques, while the ideal value of precision is 100.
The second highest value is that of the SGM-R algorithm which is 73.51%, while the Optical Flow method performs poorly with a value of only 65.75%.

$$P = \frac{tp}{tp + fp} \qquad 4.3$$

2. Recall or Sensitivity or True Positive Rate(*tpr*)

As stated earlier, Recall is another measure used to quantify how the proposed algorithm matches the ground truth. Recall, or Sensitivity, or equivalently True Positive Rate (TPR) is defined by Equation (4.4) and is a measure of how well we have identified the ground truth foreground without misidentifying the foreground [5].
As shown in Figure (4.2) and in Table (4.3), there was as much false identification of regions with the proposed method as with the other techniques.
fn is the area of foreground over calculated in the resultant image. The lower its value, the greater the value of Recall.
The ideal value of Recall is 100. The proposed algorithm has achieved 93.44%.
The overall highest value is that of the Soo Wan Kim algorithm which is 97.86%, while the Optical Flow method performs worse with a value of 90.81%.

$$R_S = \frac{tp}{tp + fn} \qquad 4.4$$

3. F-score of Precision and Recall

F-score is the weighted percentage average of precision and recall. F-score of Precision and Recall (i.e., harmonic mean) is defined in Equation (4.5). F-score measures the proposed methods accuracy.
The ideal value of F-score is 100%, and the proposed algorithm has achieved 93.46%, which is the highest value among the other four algorithms.
The second highest value is that of the SGM-R algorithm, which is 82.65%, while the Optical Flow method performs worse with a value of 75.88%.

$$\%F_{scorePR} = \frac{100 \times 2 \times (recall \times precision)}{recall + precision} \qquad 4.5$$

4. Specificity or True Negative Rate

This measure describes the ratio of detected foreground pixels that are true positives. If the value of specificity is 100%, this shows that the segmentation process recognizes all actual negatives, or in other words, 100% specificity shows no positives are incorrectly tagged.
Specificity is defined by Equation (4.6), and is a measure of how well we have been able to identify the ground truth background without misidentifying the ground truth foreground.
It is the opposite of precision; the lower the value of fp, the greater the value of specificity.
The ideal value of specificity is 100%, and the proposed algorithm has achieved 88.23%, which is the highest value among the other four algorithms.
The second highest value is that of the SGM-R algorithm, which is 39.24%, while the Optical Flow method performs worse with a value of 17.68%.

$$Spec = \frac{tn}{tn + fp} \qquad 4.6$$

5. Balance Classification Rate or Area Under the Curve

This statistical tool is also called *Yule Coefficient* (YC). Balance Classification Rate (BCR) or Area Under the Curve is defined by Equation (4.7), and is the overall measure of how well we have been able to identify the ground truth foreground and background. The greater the area under the curve, the better is the performance.
The ideal value of BCR or area under the curve is 100%, and the proposed algorithm has achieved 90.84% which is the highest value among the other four algorithms.
The second highest value is that of the SGM-R algorithm which is 66.84%, while the Optical Flow method performs worse with a value of 54.25%.

$$BCR\ or\ AUC = \frac{1}{2}\left(\frac{tp}{tp + fn} + \frac{tn}{tn + fp}\right) \qquad 4.7$$

or

$$BCR\ or\ AUC = \frac{1}{2}(Recall + Specificity)$$

6. Geometric Mean of Sensitivity and Specificity

Geometric mean of sensitivity and specificity is defined by Equation (4.8), and is an overall measure of how well we have been able to identify the ground truth foregrounds and backgrounds.

The ideal value of the geometric mean of sensitivity and specificity is 100%, and the proposed algorithm has achieved 90.65% which is the highest value among the other four algorithms.
The second highest value is that of SGM-R which is 60.67%, while Optical Flow performs worse with a value of 38.79%.

$$G = \sqrt{recall \times specificity} \qquad 4.8$$

7. F-Score of Sensitivity and Specificity

F-score of sensitivity and specificity (i.e., harmonic mean) is defined by Equation (4.9) and is an overall measure of how well we have been able to identify the ground truth foregrounds and backgrounds.
The ideal value of F-score of sensitivity and specificity is 100%, and the proposed algorithm has achieved 90.48%. The second highest value is that of SGM-R which is 55.16%, while Optical Flow performs worse with a value of 28.35%.

$$\%F_{scoreSS} = \frac{100 \times 2 \times (recall \times specificity)}{recall + specificity} \qquad 4.9$$

8. %Balance Error Rate

Percentage Balance Error Rate is defined by Equation (6.10), and is the overall measure of how much we have misidentified the ground truth foreground and background.
The ideal value of %Balance Error Rate is 0, and the proposed algorithm has achieved 9.16% which is the best value. The second best value is that of SGM-R which is 33.16%, while Optical Flow performs poorly with the value of 45.75%.

$$\mathrm{BER}\,\% = 100 \times \left[1 - \frac{1}{2} \times \left\{\left(\frac{\mathrm{tp}}{\mathrm{tp} + \mathrm{fn}} + \frac{\mathrm{tn}}{\mathrm{tn} + \mathrm{fp}}\right)\right\}\right] \qquad 4.10$$

Or

$$\mathrm{BER}\,\% = 100 \times \left[1 - \frac{1}{2} \times \{(\mathrm{Recall} + \mathrm{Specificity})\}\right]$$

9. Similarity

Similarity is defined by Equation (6.11), also called Jaccard coefficient, which is a statistic tool used for comparing the similarity and diversity of sample sets.
It is a measure of how similar the segmented foreground is to the ground truth foreground with 1 being most similar and anything less than 1 being increasingly less similar.
The lower the value of $(fn + fp)$, the greater is the value of similarity.
The ideal value of similarity is 100%, and the proposed algorithm has achieved 87.78% which is the highest value. The second highest value is that of SGM-R which is 70.44%, while Optical Flow performs poorly with the value of 61.91%.

$$Sim = \frac{tp}{tp + fn + fp} \qquad 4.11$$

10. Accuracy

Accuracy is also known as percentage correct classification. This statistical measure describes how well the proposed segmentation process excludes or identifies foreground pixels. 100% accuracy means that the values obtained from the proposed algorithm are exactly the same as the values in the ground truth.
Accuracy is defined by Equation (6.12), and is a measure of how well we have identified the foreground and background ground truths without misidentifying the foregrounds and backgrounds.
The ideal value of accuracy is 100%, and the proposed algorithm has achieved 91.58% which is the highest value from other four algorithms. The second highest value is that of SGM-R which is 74.59%, while Optical Flow performs poorly with the value of 64.51%.

$$A = \frac{tp + tn}{tp + fn + fp + tn} \qquad 4.12$$

11. False Positive Rate

This measure is used to calculate the background pixels misclassified as foreground.
False Positive Rate is defined by Equation (6.13), and is the fraction of the ground truth background that has been misidentified as foreground. The greater the value of tn, the lesser the value of the false positive rate.
The ideal value of false positive rate is 0, and the proposed algorithm has achieved 11.76% which is the best value from other four algorithms. The second highest value is that of SGM-R which is 60.76%, while Optical Flow perform poorly with the value of 82.32%.

$$FPR = \frac{fp}{fp + tn} \qquad 4.13$$

Table 4.3 Rank of proposed algorithm with the state of the art algorithms

Results												
Algo No	Ideal or Perfect match value	100	100	100	100	100	0.0	100	100	100	100	0.0
	Frames	Precision	Recall or Sensitivity or True Positive Rate	%F-score of Precision and Recall	Specificity	AUC/BCR (Balanced Classification Rate)	BER (%)	%F-score of Sensitivity and Specificity	Geometric mean of sensitivity and specificity	Accuracy: mean of sensitivity and specificity	Similarity	False Positive Rate
1	Our Approach	93.5979↑	93.4452↓	93.4652↑	88.2348↑	90.8399↑	9.1601↑	90.4771↑	90.6552↑	91.5770↑	87.7849↑	11.7652↑
2	Optical Flow	65.7502	90.8131	75.8838	17.6818	54.2474	45.7526	28.3530	38.7906	64.5063	61.9064	82.3182
3	Soo Wan Kim algo	69.1276	97.8641↑	81.0086	22.0438	59.9540	40.0460	35.6124	45.9462	70.5913	68.0895	77.9562
4	MoG	70.8468	95.3105	81.1824	29.4351	62.3728	37.6273	43.0487	49.5021	70.6347	68.3454	70.5649
5	SGM-R	73.5124	94.4399	82.6509	39.2389	66.8393	33.1607	55.1559	60.6727	74.5871	70.4394	60.7611
Diff between No 1st & 5th *		20.0855	-4.4189	10.8143	48.9959	24.0006	24.0006	35.3212	29.9825	16.9899	17.3455	48.9959
Average												24.7376

↑ **Best value**

↓ **Lower value**

* In case of Recall the difference is between algorithm 1st and 3rd as our obtained value is lesser than the Soo Wan Kim approach

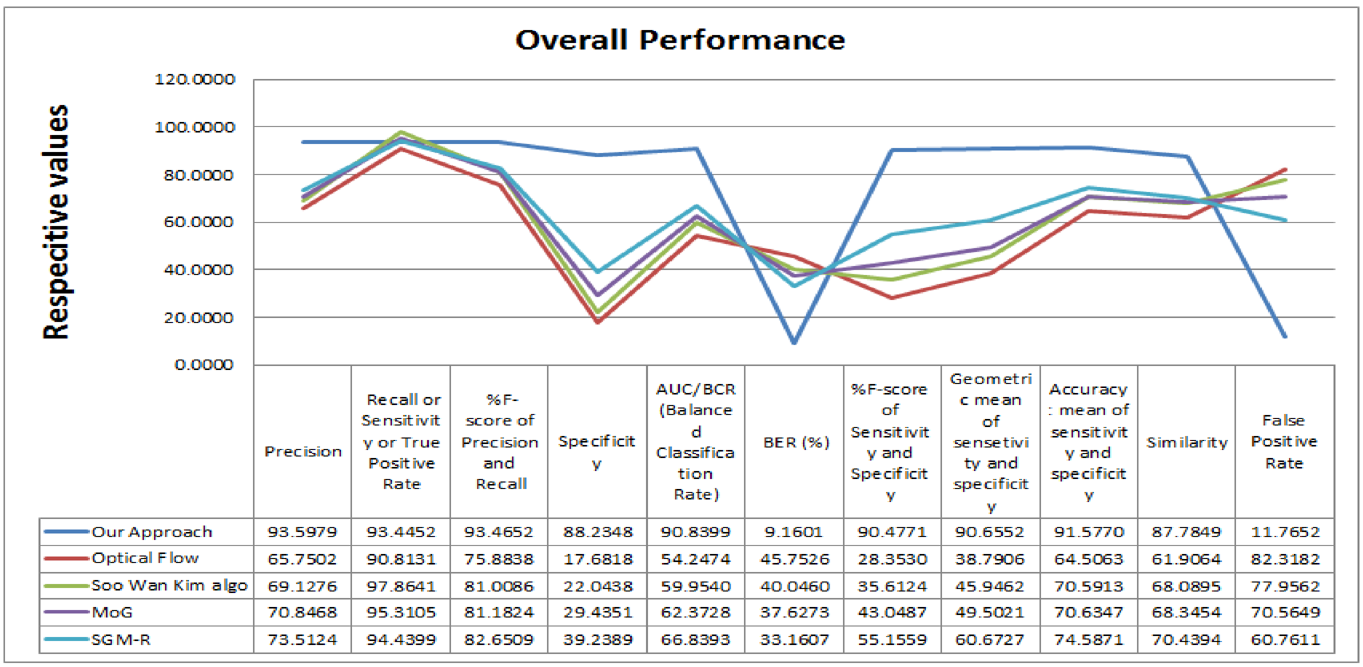


Figure 4.2 Overall Performances

Technical Evaluation

One of the main reasons of this big difference in results is that apart from optical flow all other methods use background subtraction and requires reference image which is free of foreground object(s). And in real world videos like the one used in the proposed algorithm it is not possible to have reference image in advance, which is free from foreground.

Soo Wan Kim, MoG and SGMR uses Mixture of Gaussian, are among most recent methods, proposed for foreground detection. These methods produce good results in outdoor scenes. In Mixture of Gaussian, the colours of the background objects' pixels are represented by multiple Gaussian distributions. Many researchers have reported that more than two Gaussians can badly degrade the foreground object extraction [5, 20]. The main disadvantage of Mixture of Gaussian is that it is computationally complex method and the fact that the variables require careful setting. Thus, the method requires more time in processing. Also, Mixture of Gaussian is very sensitive to sudden changes in global illumination and thus produces inaccurate results. Consequently, when the scene is still for a long time, a rapid change in global illumination may turn the whole frame into foreground [20, 46].

The comparison results are shown in the Table (4.3) and Figure (4.2). It is obvious that the proposed algorithm clearly outperforms the other four methods. SGM-R is the second best approach. MoG being the quite similar technique to SGM-R was found the third best method, while, the Soo Wan Kim algorithm was found the fourth best algorithm, based on performance measure results. Overall, the performance of the Optical Flow technique was found non satisfactory.

The recall value of the proposed method is lower than Soo Wan Kim algorithm by 4.42%. The recall or true positive rate (*trp*) and precision quantify how well an algorithm matches the ground truth [42, 43], but the proposed algorithm outperforms in precision and %F-score of precision and recall over the rest of the four methods by 20.09% and 10.81%, respectively. It is also important to know that only recall is not sufficient to compare different methods and is generally used in conjunction with precision, that gives the percentage of detected true positive as compared to the total number of items detected [44]. It is clearly shown from the results obtained, that the proposed algorithm performs much better than the second best algorithm SGM-R, on average by 24.74%.

V. Conclusion

This paper presents a simple and effective algorithm to obtain sufficient precise foreground from background using motion estimation, maxima and minima inside the foreground object. The previous works [25, 26, 29-35] on foreground detection shows that our final result has produced better foreground mask based in terms of quantitatively and qualitatively.

For quick and accurate execution of block motion estimation we have used Adaptive Rood Pattern Search algorithm.

In order to obtained precise mask of the foreground we used opening-and-closing operation.

From the performance measures it is shown that our algorithm is relatively more accurate in terms of precision, %F-score of precision, recall, sensitivity, and specificity, specificity, area under the curve, accuracy and similarity.

Acknowledgment

We are grateful to the Institute of Management Sciences Peshawar (http://imsciences.edu.pk/) through Higher Education Commission Islamabad, Pakistan (http://hec.gov.pk/) who financially supported this project.

I am also thankful to Thierry Bouwmans, Associate Professor at MIA Lab-University La Rochelle, France; Benalia Mohcene, University of Djelfa, (Djelfa) Algeria; and Soo Wan Kim, Seoul National University Korea who helped me in conducting this research.

References

[1] K. Karmann and A. V. Brandt. Moving object recognition using an adaptive background memory. *Time-Varing Image Processing and Moving Object Recognition*, 2:289–296, 1990.

[2] D. Koller, J. Weber, T. Huang, J. Malik, G. Ogasawara, B. Rao, and S. Russel. Toward robust automatic traffic scene analysis in real-time. In *Proceedings Int'l Conf. Pattern Recognition*, pages 126–131, 1994.

[3] C. Wren, A. Azarbaygaui, T. Darrell, and A. Pentland. Pfinder: Real-time tracking of the human body. IEEE Trans. *Pattern Analysis and Machine Intelligence*, 19(7):780–785, July 1997.

[4] T. E. Boult, R. Micheals, X. Gao, P. Lewis, C. Power, W. Yin, and A. Erkan. Frame-rate omnidirectional surveillance & tracking of camouflaged and occluded targets. In *Proceedings IEEE Workshop on Visual Surveillance*, pages 48–55. IEEE Computer Society, 1999.

[5] X. Gao, T. Boult, F. Coetzee, and V. Ramesh. Error analysis of background adaption. In *Proceedings of IEEE Computer Vision and Pattern Recognition*, pages 503–510. IEEE Computer Society, 2000.

[6] A. Iketani, A. Nagai, Y. Kuno, and Y. Shirai. Detecting persons on changing background. In *Proceedings of International Conference on Pattern Recognition*, pages 74–76, 1998.

[7] L. Li and M. K. H. Leung, "Integrating intensity and texture differences for robust change detection," *Image Processing, IEEE Transactions on*, vol. 11, no. 2, pp. 105–112, 2002.

[8] T. Aach, A. Kaup, and R. Mester, "Statistical model-based change detection in moving video," *Signal Processing*, vol. 31, no. 2, pp. 165–180, 1993.

[9] F. Archetti, C. Manfredotti, V. Messina, and D. Sorrenti, "Foreground-to-ghost discrimination in single-difference pre-processing," in *Proceedings of the International Conference on Advanced Concepts for Intelligent Vision Systems*, 2006.

[10] J. Xia, J. Wu, H. Zhai, and Z. Cui, "Moving vehicle tracking based on double difference and camshift," in *Proceedings of the International Symposium on Information Processing*, 2009.

[11] R. Collins, A. Lipton, T. Kanade, et al., "A system for video surveillance and monitoring," Tech. Rep. CMU-RI-TR-00-12, Robotics Institute, Pittsburgh, Pa, USA, May 2000.

[12] B. K. P. Horn and B. G. Schunck, "Determining optical flow," *Artificial Intelligence*, vol. 17, no. 1–3, pp. 185–204, 1981.

[13] Piccardi, M.; , "Background subtraction techniques: a review," *Systems, Man and Cybernetics, 2004 IEEE International Conference on* , vol.4, no., pp. 3099- 3104 vol.4, 10-13 Oct. 2004.

[14] Haritaoglu, I.; Harwood, D.; Davis, L.S.; , "W^4: real-time surveillance of people and their activities," *Pattern Analysis and Machine Intelligence, IEEE Transactions on* , vol.22, no.8, pp.809-830, Aug 2000

[15] Tao Zhao; Nevatia, R.; , "Tracking multiple humans in complex situations," *Pattern Analysis and Machine Intelligence, IEEE Transactions on* , vol.26, no.9, pp.1208-1221, Sept. 2004

[16] Haibing Ren, Guangyou Xu, *Human action recognition in smart classroom,* Proceedings of Fifth IEEE International Conference on Automatic Face and Gesture Recognition, 2002, 20-21 May 2002 Page(s): 399–404.

[17] Gutchess, D.; Trajkovics, M.; Cohen-Solal, E.; Lyons, D.; Jain, A.K.; , "*A background model initialization algorithm for video surveillance* ," Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on , vol.1, no., pp.733-740 vol.1, 2001.

[18] C. Stauffer and W. E. L. Grimson, “*Learning patterns of activity using real-time tracking,” Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 22, no. 8, pp. 747–757, 2000.

[19] P. Kaewtrakulpong and R. Bowden, “*An improved adaptive background mixture model for real time tracking with shadow detection,*” in Proceedings of the 2nd European Workshop on Advanced Video Based Surveillance Systems (AVBS ’01), 2001.

[20] Sen-ching S. Cheung and Chandrika Kamath, "*Robust techniques for background subtraction in urban traffic video*", Proc. SPIE 5308, 881 (2004).

[21] J. Lin, J.H. Xu, W. Cong, L.L. Zhou and H. Yu, “*Research on real-time detection of moving target using gradient optical flow*,” IEEE International Conference on Mechatronics and Automation, 2005, pp. 1796-1801.

[22] E. Trucco, T. Tommasini and V, Roberto, “*Near-recursive optical flow from weighted image differences*,” *Systems, Man, and Cybernetics, Part B Cybernetics, IEEE Transactions on*:, vol. 35, n 1, February, 2005, pp. 124-129.

[23] H. Ishiyama, T. Okatani and K. Dequchi, “*High-speed and high-precision optical flow detection for real-time motion segmentation*, ” Proceedings of the SICE Annual Conference, 2004, pp. 751-754.

[24] L. Wixson, “*Detecting salient motion by accumulating directionally-consistent flow*,” *Pattern Analysis and Machine Intelligence, IEEE Transactions on* vol. 22, issue 8, Aug. 2000, pp. 774-780.

[25] P. Kelly, J. Diego, P.-M. Agapito, C. O. Conaire, D. Connaghan, J. Kuklyte,and N. E. O’Connor., “Performance analysis and visualisation in tennis using a low-cost camera network,” in Multimedia Grand Challenge Track at ACM Multimedia, 2010.

[26] Nawaz, M.; Cosmas, J.; Adnan, A.; Ali, M.; , "Inter-intra frame segmentation using colour and motion for region of interest coding of video," Broadband Multimedia Systems and Broadcasting (BMSB), 2011 IEEE International Symposium on , vol., no., pp.1-4, 8-10 June 2011.

[27] A. Barjatya, "Block Matching Algorithms for Motion Estimation", in Technical Report, Utah State University, 2004.

[28] P. Kaewtrakulpong, R. Bowden, An Improved Adaptive Background Mixture Model for Real time Tracking with Shadow Detection, In Proc. 2nd European Workshop on Advanced Video Based Surveillance Systems, AVBS01, VIDEO BASED SURVEILLANCE SYSTEMS: Computer Vision and Distributed Processing (September 2001)

[29] Stauffer, C. and Grimson, W.E.L, Adaptive Background Mixture Models for Real-Time Tracking, Computer Vision and Pattern Recognition, IEEE Computer Society Conference on, Vol. 2 (06 August 1999), pp. 2246-252 Vol. 2.

[30] Barron, J.L., D.J. Fleet, S.S. Beauchemin, and T.A. Burkitt. Performance of optical flow techniques. CVPR, 1992.

[31] Olson, T., Brill, F.: Moving object detection and event recognition algorithms for smart cameras. In: Proceedings DARPA Image Understanding, Workshop, pp. 159–175 (1997)

[32] Soo Wan Kim, Kimin Yun, Kwang Moo Yi, Sun Jung Kim and Jin Young Choi “Detection of moving objects with a moving camera using non-panoramic background model” Machine Vision and Applications, Springer Berlin / Heidelberg, pp. 1 – 14, 5 October 2012.

[33] Andr´eKaup and TilAach, “Efficient Prediction of Uncovered Background in Interframe Coding Using Spatial Extrapolation”, in: “IEEE International Conference on Acoustics, Speech and Signal Processing, Volume 5: I Image and Multidimensional Signal Processing”, Adelaide, South Australia, pp.V501-V504, 1994.

[34] Patwardhan, K.A.; Sapiro, G.; Morellas, V.; , "Robust Foreground Detection In Video Using Pixel Layers," *Pattern Analysis and Machine Intelligence, IEEE Transactions on* , vol.30, no.4, pp.746-751, April 2008.

[35] Elgammal, A.; Duraiswami, R.; Harwood, D.; Davis, L.S.; , "Background and foreground modeling using nonparametric kernel density estimation for visual surveillance," Proceedings of the IEEE , vol.90, no.7, pp. 1151- 1163, Jul 2002.

[36] Yao Nie; Kai-Kuang Ma; , "Adaptive rood pattern search for fast block-matching motion estimation," *Image Processing, IEEE Transactions on* , vol.11, no.12, pp. 1442- 1449, Dec 2002

[37] Y. Nie and K. K. Ma, "Adaptive rood pattern search for fast block-matching motion estimation," *Image Processing, IEEE Transactions on,* vol. 11, pp. 1442-1449, 2002.

[38] S Zhu and K-K Ma “A New Star Search Algorithm For Fast Block Matching Motion Estimation”, in Proc. *Workshop On Very Low Bitrates Coding (VLVB)* Oct 8-9, 1998, pp. 173 -176.

[39] S. Zhu and K.-K. Ma, “A new diamond search algorithm for fast block-matching motion estimation,” in *Proc. 1997 Int. Conf. Information, Communications and Signal Processing (ICICS)*, vol. 1, Sept. 9–12, 1997, pp. 292–296.

[40] A. Barjatya, "Block matching algorithms for motion estimation," *Evolution Computation, IEEE Transactions* vol. 8, pp. 225-239, 2004.

[41] Hu, J., Kahsi, R., Lopresti, D., Nagy, G., Wilfong, G.: Why table ground-truthing is hard. In: Sixth International Conference on Document Analysis and ecognition, pp. 129–133. Seattle, WA, USA (2001)

[42] Cheung, S., Kamath, C.: Robust techniques for background subtraction in urban traffic video. Electronic Imaging: Video Communications and Image Processing 5308(1), 881–892 (2004)

[43] Benezeth, Y., Jodoin, P., Emile, B., Laurent, H., Rosenberger, C.: Review and evaluation of commonly-implemented background subtraction algorithms. In: 19th International Conference on Pattern Recognition (ICPR), pp. 1–4. Tampa, Florida (2008)

[44] Maddalena, L., Petrosino, A.; , "A Self-Organizing Approach to Background Subtraction for Visual Surveillance Applications," *Image Processing, IEEE Transactions on* , vol.17, no.7, pp.1168-1177, July 2008

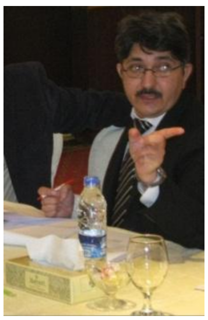
[45] MARTNEZ-MARTN, E. (2012) "Robust motion detection in real-life scenarios", . ++++

[46] T. Bouwmans, F. El Baf, B. Vachon, “Background Modeling using Mixture of Gaussians for Foreground Detection: A Survey”, Recent Patents on Computer Science, Volume 1, No 3, pages 219-237, November 2008.

[47] F. Cheng, S. Huang, and S. Ruan, “Illumination-Sensitive Background Modeling Approach for Accurate Moving Object Detection,” IEEE Transactions on Broadcasting, vol. 57, issue 4, Dec. 2011, pp. 794-801.

[48] I. Daribo, H. Saito, “A Novel Inpainting-Based Layered Depth Video for 3DTV,” IEEE Transanctions on Broadcasting, vol. 57, issue 2, Jun. 2011, pp. 533 – 541.

[49] Y. Zhao, C. Zhu, Z. Chen, D. Tian, L. Yu, “Boundary Artifact Reduction in View Synthesis of 3D Video: From Perspective of Texture-Depth Alignment,” IEEE Transactions on Broadcasting, vol 57, issue 2, pp. 510 – 522, Jun. 2011



Muhammad Nawaz received his MSc (Computer Science) and MS in Information Technology from University of Peshawar, Pakistan. He worked as a lecturer at the University of Peshawar, Pakistan; followed by working as a computer programmer at Khyber Teaching Hospital, Peshawar; and then appointed as Assistant Professor in Multimedia at the Institute of Management Sciences, Peshawar - a position he still holds. Currently he is a PhD candidate at Brunel University London researching foreground detection in images.

John Cosmas (M’86) obtained a B.Eng honors degree in Electronic Engineering at Liverpool University in 1978 and a PhD in Image Processing and Pattern Recognition at Imperial College, University of London in 1987. He is a Professor of Multimedia Systems and became a Member (M) of IEEE in 1987 and a Member of IEE in 1977. His research interests are concerned with the design, delivery and management of new TV and telecommunications services and networks, multimedia content and databases, and video/image processing. He has contributed towards eight EEC research projects and has published over 80 papers in refereed conference proceedings and journals. He leads the Wireless Networks and Communications Centre within the School of Engineering and Design at Brunel University.

Pavlos Lazaridis was born in Larissa, Greece. He received the electrical engineering degree from Aristotle University of Thessaloniki, Greece, in 1990, the DEA degree in electronics from the Université Pierre & Marie Curie, Paris, France, in 1992, and the Ph.D degree in electronics and telecommunications from the Ecole Nationale Supérieure des Télécommunications (ENST), Paris, in 1996.From 1989 to 1990, he was with the Greek Public Power Corporation, Athens, Greece, where he worked on network reliability evaluation algorithms. From 1991 to 1996, he was involved with research on semiconductor lasers, wave propagation, and nonlinear phenomena in optical fibers for the Centre National d’Etudes des Télécommunications (CNET) and teaching at the ENST. In 1997, he became Head of the Antennas and Propagation Laboratory, TDF Metz (Télédiffusion de France/France Télécom research center), where he was involved with research in antenna coupling and radio coverage prediction algorithms for cellular mobile systems, DAB, and DVB-T broadcasting. From 1998 to 2002 he was with European Patent Office (EPO) Rijswijk,The Netherlands, as a senior examiner in the field of Electronics-Telecommunications. He is teaching as a senior lecturer at the Alexander Technological Educational Institute of Thessaloniki, Greece, and at the postgraduate level, at the jointly organized MSc course with Brunel University West London.

Zaharias D. Zaharis received the B.Sc. degree in physics in 1987, the M.Sc. degree in electronics in 1994 and the Ph.D. degree in 2000 from the Aristotle University of Thessaloniki, Greece. Also, in 2011 he obtained the Diploma degree in electrical and computer engineering from the same university. Since 2002 he has been working in the administration of the telecommunications network at the Aristotle University of Thessaloniki.

His research interests include design and optimization of antenna arrays, design and optimization of microwave circuits and systems, mobile communications, RF measurements, radiowave propagation and electromagnetic scattering.

Dr. Zaharis is a member of the Greek Physics Society and the Technical Chamber of Greece.

Hamdullah Mohib (M’11) obtained a B.Eng honors degree in Computer Systems Engineering at Brunel University in 2008. He is a PhD candidate at the Centre for Multimedia Communications Research, School of Engineering and Design, Brunel University. His research topic is: Quality Optimisation in 3D Video Communication over Heterogeneous Networks.

Yue Zhang obtained his B.E and M.E degree in 2001 and 2004 respectively at Beijing University of Post and Telecommunications, China. In 2008 he received his PhD degree in Brunel University, UK, where he also worked as a Research Engineer for the EU IST FP6 project- PLUTO. He was responsible for the radio propagation modeling and measurement, transmitter/receiver diversity design and measurement for DVB-T/H systems and RF/Digital/DSP design for the on-channel repeater for DVB-T/H systems. From 2008, he was a Signal Processing Design Engineer in Microwave Measurement Division-Europe, Anritsu Corp. He was responsible for the RF/IF, digital and DSP design for the measurement instruments for various wireless and broadcasting systems. From 2010, he joined Department of Computer Science and Technology at the University of Bedfordshire, Luton, UK as a Senior Lecturer. He also worked as a Royal Academy of Engineering, UK, Industrial Secondment with Aeroflex Ltd. His research interests are signal processing, wireless and broadcasting systems, MIMO-OFDM systems, radio propagation model and multimedia and wireless networks. Dr. Zhang currently serves as an Associate Editor for IEEE Transactions on Broadcasting.