

Research Article

Enhancement of Background Subtraction Techniques Using a Second Derivative in Gradient Direction Filter

Farah Yasmin Abdul Rahman,¹ Aini Hussain,¹ Wan Mimi Diyana Wan Zaki,¹ Halimah Badioze Zaman,² and Nooritawati Md Tahir³

¹ Smart Engineering Systems Laboratory, Department of Electrical, Electronic & Systems Engineering, Faculty of Engineering & Built Environment, Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Selangor, Malaysia

² Institute of Visual Informatics, Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Selangor, Malaysia

³ Faculty of Electrical Engineering, Universiti Teknologi MARA, 40450 Shah Alam, Selangor, Malaysia

Correspondence should be addressed to Aini Hussain; aini@eng.ukm.my

Received 27 March 2013; Revised 5 September 2013; Accepted 24 September 2013

Academic Editor: Vijayan K. Asari

Copyright © 2013 Farah Yasmin Abdul Rahman et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A new approach was proposed to improve traditional background subtraction (BGS) techniques by integrating a gradient-based edge detector called a second derivative in gradient direction (SDGD) filter with the BGS output. The four fundamental BGS techniques, namely, frame difference (FD), approximate median (AM), running average (RA), and running Gaussian average (RGA), showed imperfect foreground pixels generated specifically at the boundary. The pixel intensity was lesser than the preset threshold value, and the blob size was smaller. The SDGD filter was introduced to enhance edge detection upon the completion of each basic BGS technique as well as to complement the missing pixels. The results proved that fusing the SDGD filter with each elementary BGS increased segmentation performance and suited postrecording video applications. Evidently, the analysis using *F*-score and average accuracy percentage proved this, and, as such, it can be concluded that this new hybrid BGS technique improved upon existing techniques.

1. Introduction

Object extraction is a technique used in suppressing the background of a video scene to detect subjects that appear in the frame. The technique involves comparing or subtracting the current frames from the background frame and treating the remaining pixels as foreground [1]. Prior research on background subtraction (BGS) used several parametric BGS techniques, such as running average [2–4], running Gaussian average [5–7], approximate median filter [7, 8], and Gaussian Mixture Model [9–11]. These parametric techniques determine the foreground and update the subsequent background based on the distribution of intensity value [12]. Aside from these techniques, other studies have introduced nonparametric models that detect foreground and background based on the intensity of statistical properties [13]. Other nonparametric models include a kernel density estimator [14] and mean shift estimation [15].

This work focuses on basic BGS techniques, such as frame differencing, approximate median, running average, and running Gaussian average. The motivation of this work lies in the fact that most edge pixels are undetected after performing object extraction techniques based on FD, AM, RA, and RGA. In this study, however, we have overcome this limitation by detecting all edge pixels; hence a perfect blob can be retrieved through morphological procedures.

This is done by applying an SDGD filter on the results of background suppression and combining the foreground pixels generated by BGS techniques with the detected edge as our extracted object. The edge pixels are expected to fill in the boundary gap, which creates better connections among pixels in the boundary. This leads to better foreground frame detection.

This paper is organized as follows. Section 2 presents an overview of several basic BGS techniques and the SDGD

filter. Section 3 describes the methodology. Section 4 discusses the results. Finally, Section 5 concludes our paper.

2. Literature Review

This section provides a review of the literature on the four BGS techniques evaluated in this study, namely, frame differencing, approximate median, running average, and running Gaussian average. SDGD filter studies are also presented.

2.1. Frame Differencing. Frame differencing (FD) is the most fundamental technique in BGS. FD involves finding the absolute difference between the current frame and a previous or background frame [1]. The absolute difference is then compared with an appropriate threshold value, A , to detect the object as shown in (1), where F_i is the current frame intensity value, B_i is the background intensity value, and F_g is the foreground intensity value. This technique uses the same background frame for all video sequences:

$$F_g(x, y) = \begin{cases} 1 & |F_i(x, y) - B_i(x, y)| > A \\ 0 & \text{otherwise.} \end{cases} \quad (1)$$

2.2. Approximate Median Filter. The approximate median (AM) algorithm is adaptive, dynamic, nonprobabilistic, and intuitive [8]. AM is obtained by calculating the difference between two video frames and using this difference in determining the perfect method for updating the background. AM is considered one of the most acceptable methods because it provides the most accurate pixel identification.

Several studies have evaluated the efficacy of the AM algorithm. He et al. [7] tested the effectiveness of the AM algorithm as part of their optimized algorithm for vehicle detection in an embedded system. Their approach yields highly accurate information with less computational time when detecting and tracking vehicles in a traffic scene. Equation (2) presents AM updates in the reference frame of every video sequence. The succeeding background frame, B_{i+1} , is dependent on the intensity value of both the present frame, F_i , and background frame, B_i :

$$B_{i+1}(x, y) = \begin{cases} F_i(x, y) + 1 & F_i(x, y) > B_i(x, y) \\ F_i(x, y) - 1 & \text{otherwise.} \end{cases} \quad (2)$$

2.3. Running Average. Running average (RA) is another technique for updating a background image. A pixel is classified as background when the pixel value belongs to the corresponding distribution of the background model, and if otherwise, the mean of the distribution is updated [4]. The updated image is then used in the changing scene. The computational effect in an RA background is lesser because only the weighted sums of the two images are computed, and thus, low computational and space complexities are produced [3]. Moreover, several researchers have utilized this method to detect moving objects in video captured by a static camera.

Several studies were conducted to enhance the efficiency of BGS based on the RA method [2–4]. The outcome by Park et al. in [4] showed that application of a hierarchical data

structure significantly increased the processing speed with accurate motion detection. This outcome can be attributed to the updating of the background frame by the RA method. Equation (3) shows a specified learning rate based on the previous background frame, where α is the learning rate and A is the threshold value:

$$B_{i+1}(x, y) = \begin{cases} B_i(x, y) & F_i(x, y) > A \\ \alpha * F_i(x, y) + (1 - \alpha) * B_i(x, y) & \text{otherwise.} \end{cases} \quad (3)$$

2.4. Running Gaussian Average. This method combines both the Gaussian function and RA. Overall, the running Gaussian average (RGA) method has a significant advantage over other approaches because it requires less processing time and utilizes less memory compared with nonrecursive methods such as mixture of Gaussian (MoG) and kernel density estimation (KDE) [9]. Equation (4) shows how the reference frame that is represented by the mean μ is updated in each video sequence using this method. Unlike AM and RA, this method uses $2\sigma_i$ as a threshold value:

$$\begin{aligned} \mu_{i+1} &= \alpha F_i + (1 - \alpha) \mu_i, \\ \sigma_{i+1}^2 &= \alpha (F_i - \mu_i)^2 + (1 - \alpha) \sigma_i^2. \end{aligned} \quad (4)$$

2.5. Second Derivative in a Gradient Direction (SDGD) Filter. In image processing studies, researchers use first- and second-order derivatives to detect the edge of an object based on its gradient. By using the first derivative, the edge location is defined at the maximum position of the steep and the descent [16]. Traditional edge detection methods, such as those by Prewitt, Sobel, and Roberts, convolved the image with a specific kernel [16, 17]. However, these techniques were reported to be sensitive to noise and inaccurate [17]. In 1986, the Canny edge detector was introduced, which represented an improvement over the traditional methods [17, 18]. The detector applied Gaussian smoothing to reduce noise, unwanted details, textures, nonmaximum suppression, and hysteresis thresholding to find the edges [19].

The second-order derivative approach defines the edge pixels based on changes in brightness or zero crossing in the image area [19, 20]. SDGD is a nonlinear operator that can be expressed in the first and second derivatives. Additionally, similar to Canny, SDGD is combined with a Gaussian low pass filter for smoothing purposes [21]. Moreover, a Laplace operator is used to simplify the SDGD operation [22].

The Laplacian is defined as

$$\nabla^2 = \frac{\partial^2 a}{\partial x^2} + \frac{\partial^2 a}{\partial y^2} = (h_{2x} \otimes a) + (h_{2y} \otimes a), \quad (5)$$

where h_{2x} and h_{2y} are the second derivative filters.

The basic versions in the second derivative filters are given by

$$\begin{aligned} [h_{2x}] &= [h_{2y}] = [1 \quad -2 \quad 1], \\ [h_2] &= \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}. \end{aligned} \quad (6)$$

Associating (5) with the Gaussian filter yields

$$b = g_\sigma \otimes (h_2 \otimes a) = (g_\sigma \otimes h_2) \otimes a, \quad (7)$$

where g_σ is the Gaussian low pass filter.

Five partial derivatives are used in the SDGD filter, as follows:

$$\begin{aligned} A_{xx} &= \frac{\partial^2 a}{\partial x^2}, & A_{xy} &= \frac{\partial^2 a}{\partial x \partial y}, & A_x &= \frac{\partial a}{\partial x}, \\ A_{yy} &= \frac{\partial^2 a}{\partial y^2}, & A_y &= \frac{\partial a}{\partial y}. \end{aligned} \quad (8)$$

Therefore,

$$\text{SDGD} = \frac{A_{xx}A_x^2 + 2A_{xy}A_xA_y + A_{yy}A_y^2}{A_x^2 + A_y^2}. \quad (9)$$

A detailed explanation of SDGD can be found in [20, 21, 23].

In [19, 24], SDGD was presented as a filter used in finding edges and measuring objects. Several studies utilize the SDGD filter. For example, Aarnick et al. [25] used this filter in analyzing the ultrasound images of male kidneys and prostate in their study on preprocessing algorithms for edge detection at multiple resolution scales. Their study [25] reported that detecting the contour of objects in grey medical images could be improved by applying an adaptive filter size in SDGD. Nader El-Glaly [24] used SDGD as part of her work in a digital inpainting algorithm. Hagara and Moravcik [23] introduced the PLUS operator, a combination of the SDGD filter and Laplace operator for edge detection. Similar findings were obtained by using PLUS and SDGD filters based on a kernel size of nine or lower, and PLUS yielded better results when the kernel size was greater than nine and was suitable in locating the edges of small objects. Similarly, Verbeek and Van Vliet [22] compared Laplace, SGD, and PLUS derived from 2D and 3D images. Their research confirmed the findings in [23].

The idea of combining two methods in one algorithm was inspired by a study made by Zheng and Fan [3], where RA and temporal difference were combined to detect the moving object. Another example of hybrid research in BGS was conducted by Lu and Wang [26]. They crossbred optical flow and double background filtering to detect the moving object. Zaki et al. [27] combined frame differencing with a scale invariant feature detector to detect the moving object in various environments.

Based on the study of Persoon et al. [19], the SDGD filter gave better surface localization, especially in highly-curved areas, compared with the Canny edge detection

technique. Thus, we adopted this filter in our present work. In addition, Persoon et al. showed that SDGD guaranteed minimal detail smoothing that led to better visualization of polyps in computed tomography (CT) scan data. This finding is aligned with our results reported in [28]. Further, the study by Nader El-Glaly [24] used the SGD filter in developing an enhanced partial-differential equation-based digital inpainting algorithm to find the missing data in digital images.

To the best of our knowledge, this study is new because no prior work exists that integrates the SDGD filter with the BGS technique. Although Al-Garni and Abdennour used edge detection and the FD technique to find moving vehicles, no information was provided on the edge detection technique they utilized [29]. We used an SDGD filter to enhance the performance of the existing background subtraction technique by combining foreground pixels generated by BGS techniques with the detected edge as our extracted object. The edge pixels are expected to fill in the boundary gap, which will create better connections among pixels in the boundary. This research is an extension of our previous work published in [28] which uses more data from a variety of data sources and more detailed analysis is being done.

3. Methodology

This section discusses the database used and the proposed method.

3.1. Dataset. This study utilized datasets that were acquired from selected prerecorded video collections of several online databases.

- (a) *Smart Engineering System Research Group (SESERG) UKM Collections.* This video collection consists of various human actions and activities recorded by students involved in SESERG studies on smart surveillance systems. Besides humans, this database also has a collection of moving cars that are used as nonhuman samples for classification that will be explained further in Section 3.4.
- (b) *CMU Graphic Lab Motion Capture (MoCap) Database [30].* This database, which is owned by Carnegie Mellon University, contains 2,506 trials in 6 categories and 23 subcategories. The videos were recorded in an indoor environment.
- (c) *CMU Motion of Body (MoBo) Database [31].* This database, which is also owned by Carnegie Mellon University, consists of videos showing six different angles of the subject who is walking on a treadmill.
- (d) *Multicamera Human Action Video Data (MuHAVi) [32].* This database is owned by the Digital Imaging Research Centre at Kingston University. It presents 17 action classes with 8 camera views.
- (e) *Human Motion Database (HMDB51) [33].* This database consists of collections of edited videos from digitized movies and YouTube. The collection contains 51 action categories with 7,000 manually annotated clips.

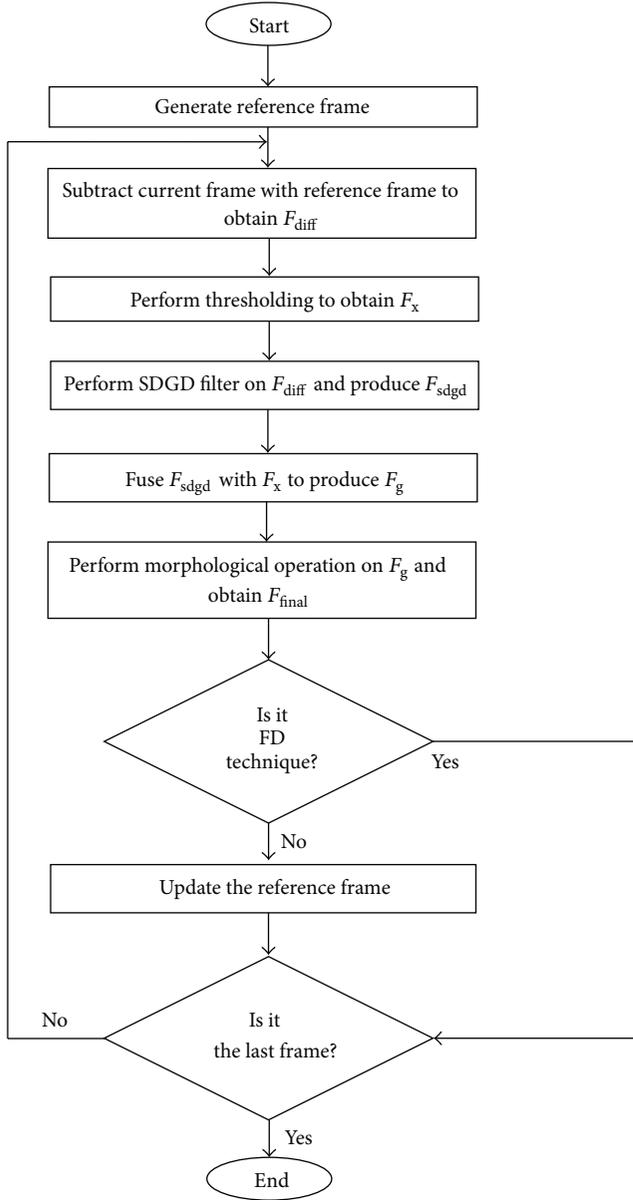


FIGURE 1: Flowchart of the proposed algorithm.

The initial background for videos taken from databases (a), (b), (d), and (e) was modeled using the median value from a selected frame interval. We used five frames from the video sequences with 10 intervals between the frames, that is F_1 , F_{10} , F_{20} , F_{30} , and F_{40} . Next, the median value of these selected frames was used as the reference image. Details of this technique are presented in [34]. Meanwhile, database (c) provided the background reference frame. All datasets except the MuHAVi videos were manually segmented to obtain the ground truth.

3.2. The Proposed Algorithm. The methodology of our proposed techniques is as described in the flowchart shown in Figure 1.

First, the dataset was tested using the following basic parametric BGS techniques: FD, RA, AM, and RGA. Next, the SDGD filter was fused with each technique by combining the output of a background technique with the SDGD filter output. The SDGD filter was selected as a segmentation tool because it produced better results compared to other edge detection techniques (Sobel, Canny, and Roberts) [19, 28].

3.3. Evaluation Method. To evaluate the performance of each BGS technique, we calculated the average of recall, precision, F score, and accuracy.

- (a) Recall (Rcl) refers to the detection rate, which is calculated by comparing the total number of detected true positive pixels with the total number of true positive pixels in the ground truth frame [35]. This is also known as sensitivity. The following equation shows how recall is calculated:

$$\text{recall} = \frac{\Sigma TP}{\Sigma TP + \Sigma FN}, \quad (10)$$

where FP is the false positive and FN is false negative.

- (b) Precision (Prdsn) is the ratio between the detected true positive pixels and the total number of positive pixels detected by the method [35, 36]. This is also known as specificity:

$$\text{precision} = \frac{\Sigma TP}{\Sigma TP + \Sigma FP}, \quad (11)$$

where FP is false positive, FN is false negative, TP is true positive, and TN is true negative.

- (c) F -measure or balance F -score is the weighted harmonic mean of recall and precision. It is used as a single measurement for comparing different methods [35, 36]:

$$F = \frac{2 \cdot \text{recall} \cdot \text{precision}}{\text{recall} + \text{precision}}. \quad (12)$$

- (d) Accuracy is the percentage of correct data retrieval. It is calculated by dividing the number of pixels with true positive plus pixels with true negative pixels over the total number of pixels in the frame. The following equation displays the calculation of accuracy [36]:

$$\text{accuracy} = \frac{\Sigma TP + \Sigma TN}{\Sigma TP + \Sigma TN + \Sigma FP + \Sigma FN} \times 100. \quad (13)$$

This study utilized videos with a multiple number of frames. Hence, we presented a comparison of the average F score and average accuracy percentage as the overall performance benchmark in each BGS technique with and without SDGD filter. The following equation is used to calculate the percentage of improvement:

$$\begin{aligned} \%im(F \text{ score}) &= (F \text{ score}_{\text{with}} - F \text{ score}_{\text{without}}) \times 100\%, \\ \%im(\text{accuracy}) &= \text{Accuracy}_{\text{with}} - \text{Accuracy}_{\text{without}} \end{aligned} \quad (14)$$

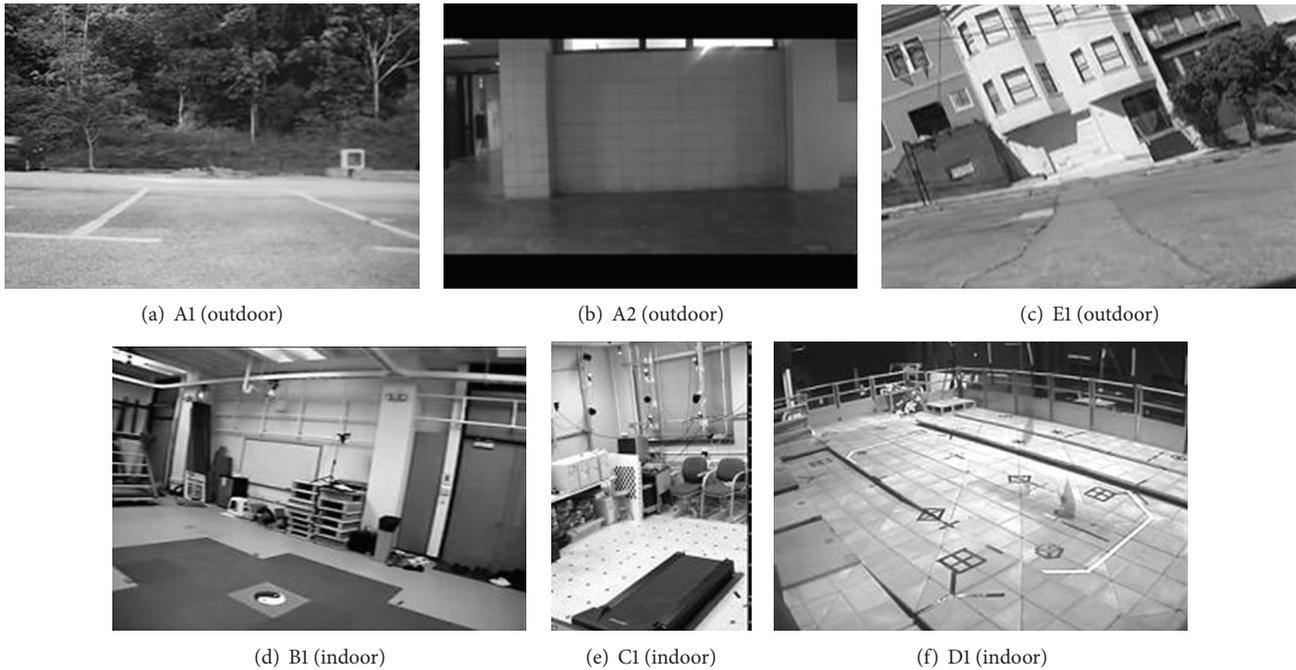


FIGURE 2: Background image of data in this study.

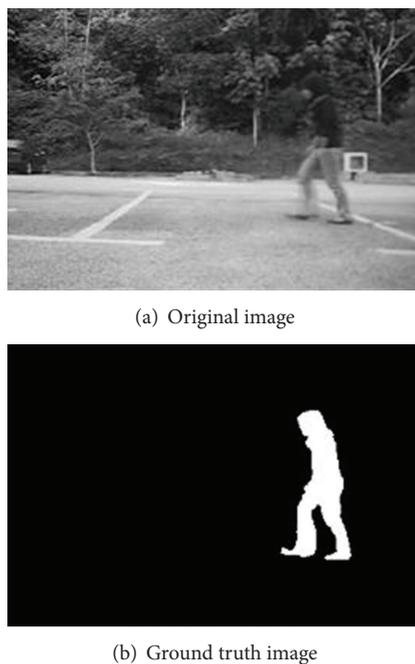


FIGURE 3: Original image and its ground truth of a frame in A1 for an outdoor environment.

3.4. Classification. Using the artificial neural network (ANN) as classifier, the segmented image from the proposed technique was subjected to classification testing. Training input of the ANN was extracted from 1,500 randomly chosen segmented frames/images. 750 human blob images represented

human samples, and another 750 car blob images represented non-human samples.

A scale conjugate gradient and the back propagation rules were chosen to train the classifier. The ANN was designed with one hidden layer containing ten hidden neurons and an output layer containing two hidden neurons. Both layers used sigmoid as the activation function and the mean squared error value as the performance function. Images were classified as either human or non-human. Next, another set of 1,000 frames/images were chosen for testing. We applied leave-one-out cross validation in our study. Ten experiments and the average classification rate were used to evaluate whether the recognition performance was human or not.

The evaluation was based on the segmented frames/images generated from the enhanced FD technique, with the SDGD filter added to the algorithm. In this experiment, we only used the FD instead of other BGS techniques because it has the fastest processing time. We also performed statistical tests such as recall, precision, and F -score on the obtained classification result by taking positive identification on the classification of human and negative for the classification of non-human.

4. Results and Discussion

This section discusses the robustness of the proposed technique based on videos taken from five different databases. To confirm the robustness of the proposed algorithm, we tested using video that presented a different environment or camera angle. Specifically, the video sequences were recorded in multiple environments (indoors and outdoors) and multiple views.

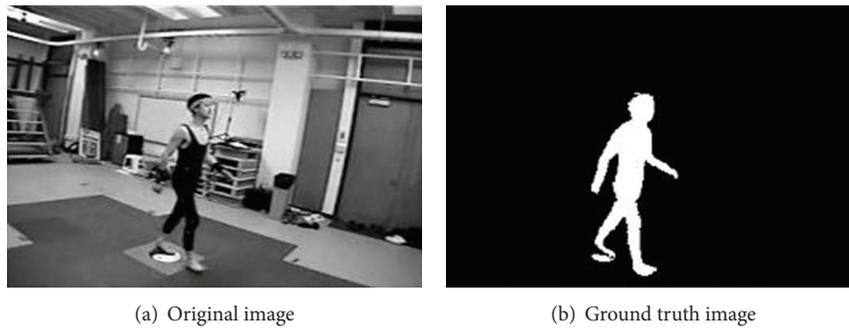


FIGURE 4: Original image and its ground truth of a frame in B1 for an indoor environment.

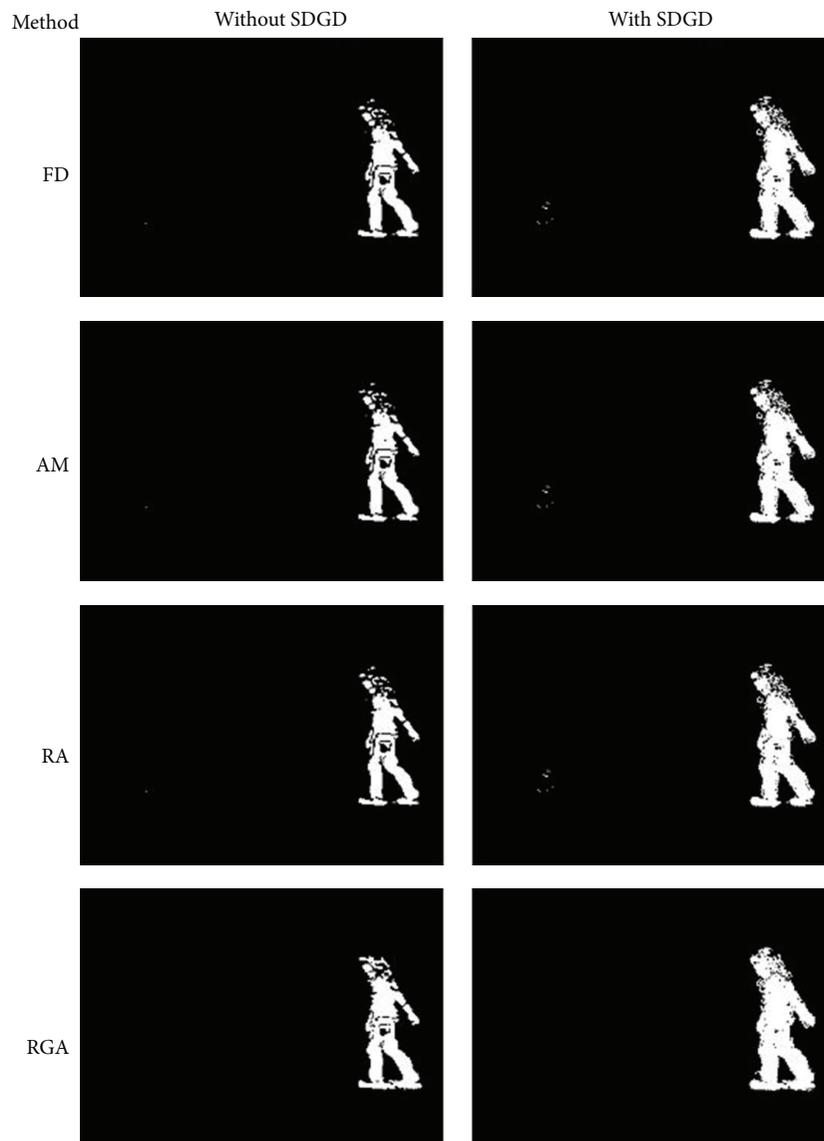


FIGURE 5: Segmentation performance using various background subtraction techniques for an outdoor environment.

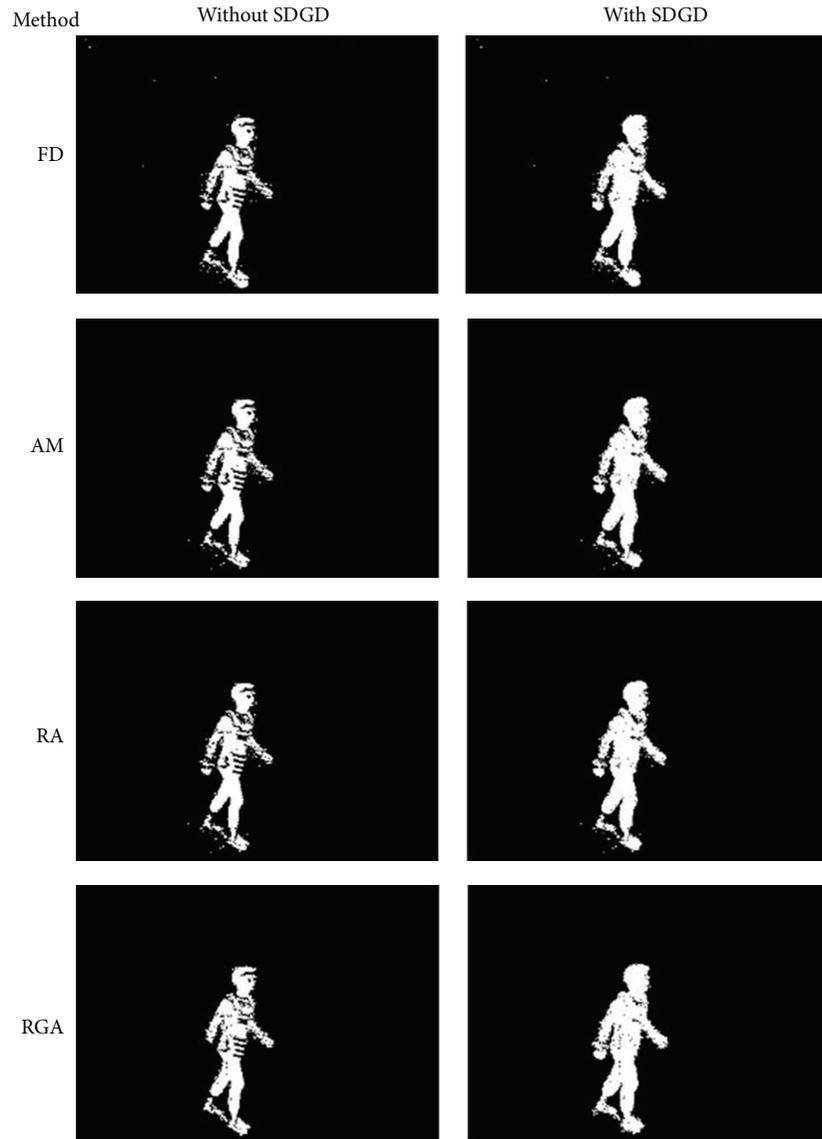


FIGURE 6: Segmentation performance using various background subtraction techniques for an indoor environment.

Figures 2(a)–2(f) show some of the background frames used in this research. These background frames were generated by using the median value of selected frames in the video sequences, except for the videos obtained from the MoBo database. Rather than stating the filename, we assigned letters, from A to E, to identify the videos representing the data: A refers to SESRG, B to MoCap, C to MoBo, D to MuHAVi, and E to HMDB51. The words in brackets indicate whether the video environment was either indoor or outdoor.

Next, we present the subjective results of our object extraction. Since this study involves video data with a multiple number of frames, we only depict the results obtained for frame number 10 in data A1 and frame number 29 in data B1 to represent the outdoor and indoor scene samples. Figures 3(a) and 4(a) illustrate the original images of a frame in videos A1 and B1, respectively. Figures 3(b) and 4(b) depict the ground truth images.

Figures 5 and 6 present the extracted subjects in both indoor and outdoor environments using FD, AM, RA, and RGA, with and without the SDGD filter.

The first column of Figures 5 and 6 show that all the basic BGS techniques were capable of detecting the object in the scene of interest based on the tested videos. However, many pixels were missing and resulted in a smaller blob compared with the ground truth image. Our proposed technique solves the problem of missing pixels and reduced blob size because it combines the SDGD filter with FD, AM, RA, and RGA, as shown in column 2 of Figures 5 and 6. The pixel size of the extracted object is slightly enlarged and becomes more compound. Foreground detection showed significant improvement after the proposed technique was applied to all datasets.

Tables 1 and 2 show the number of TP, TN, FP, and FN for the selected indoor and outdoor samples with and without the

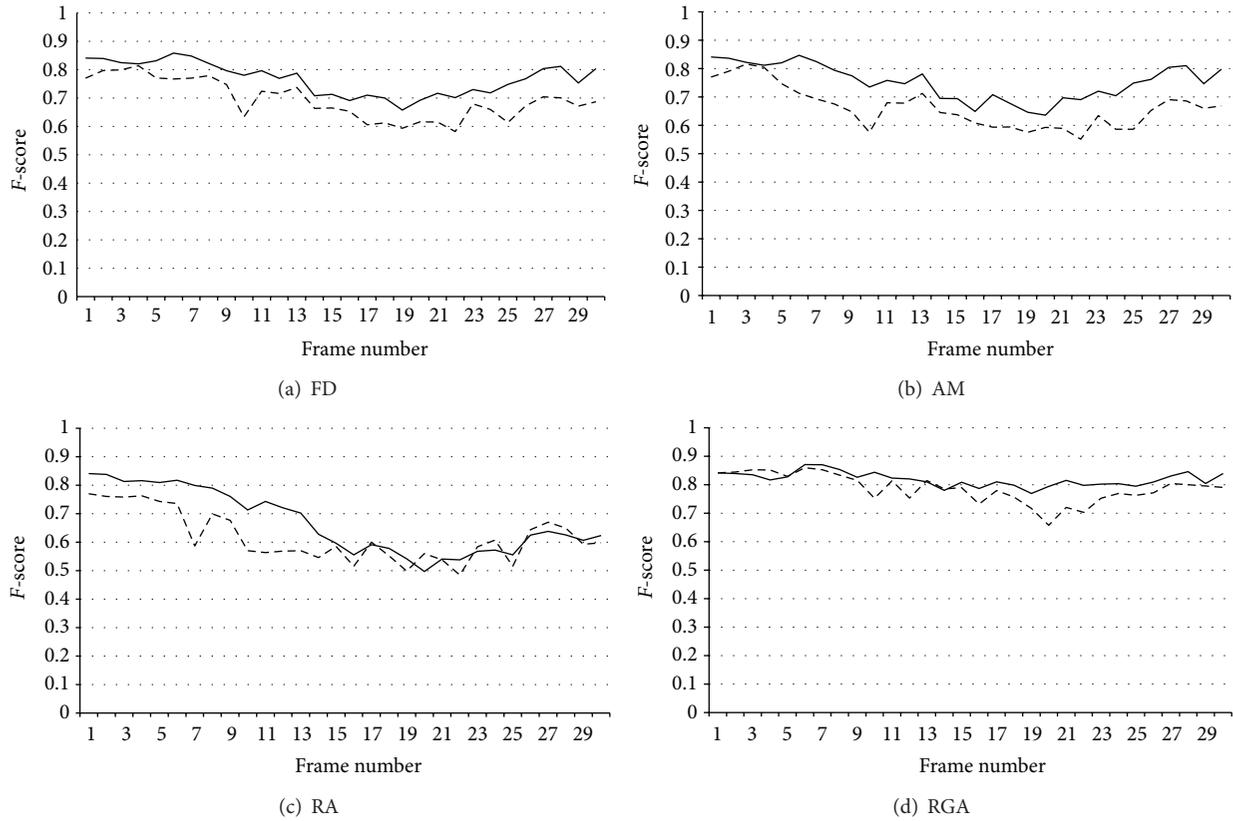
FIGURE 7: Graph of F -score versus frame number for outdoor sample.

TABLE 1: Performance for frame 10 in outdoor sample with and without SDGD filter.

	Method	TP	FP	FN	TN	Recall	Precision	F -Score
FD	without SDGD	2046	1279	1091	80064	0.65	0.62	0.63
	With SDGD	2521	804	616	80539	0.80	0.76	0.78
AM	without SDGD	1986	1785	1151	79558	0.63	0.53	0.57
	With SDGD	2374	951	763	80392	0.76	0.71	0.73
RG	without SDGD	1959	1781	1178	79562	0.62	0.52	0.57
	With SDGD	2303	1022	834	80321	0.73	0.69	0.71
RGA	without SDGD	2441	912	696	80431	0.78	0.73	0.75
	With SDGD	2654	499	483	80844	0.85	0.84	0.84

TABLE 2: Performance for frame 29 in indoor sample with and without SDGD filter.

	Method	TP	FP	FN	TN	Recall	Precision	F -Score
FD	without SDGD	2336	1372	809	79963	0.74	0.63	0.68
	With SDGD	2788	920	1165	79607	0.71	0.75	0.73
AM	without SDGD	2771	1437	808	79964	0.74	0.61	0.67
	With SDGD	2726	982	1077	79695	0.72	0.74	0.73
RG	without SDGD	2289	1419	858	79914	0.73	0.62	0.67
	With SDGD	2723	985	1106	79666	0.71	0.73	0.72
RGA	without SDGD	2564	1092	1175	79649	0.69	0.70	0.69
	With SDGD	2979	677	1535	79289	0.67	0.81	0.73

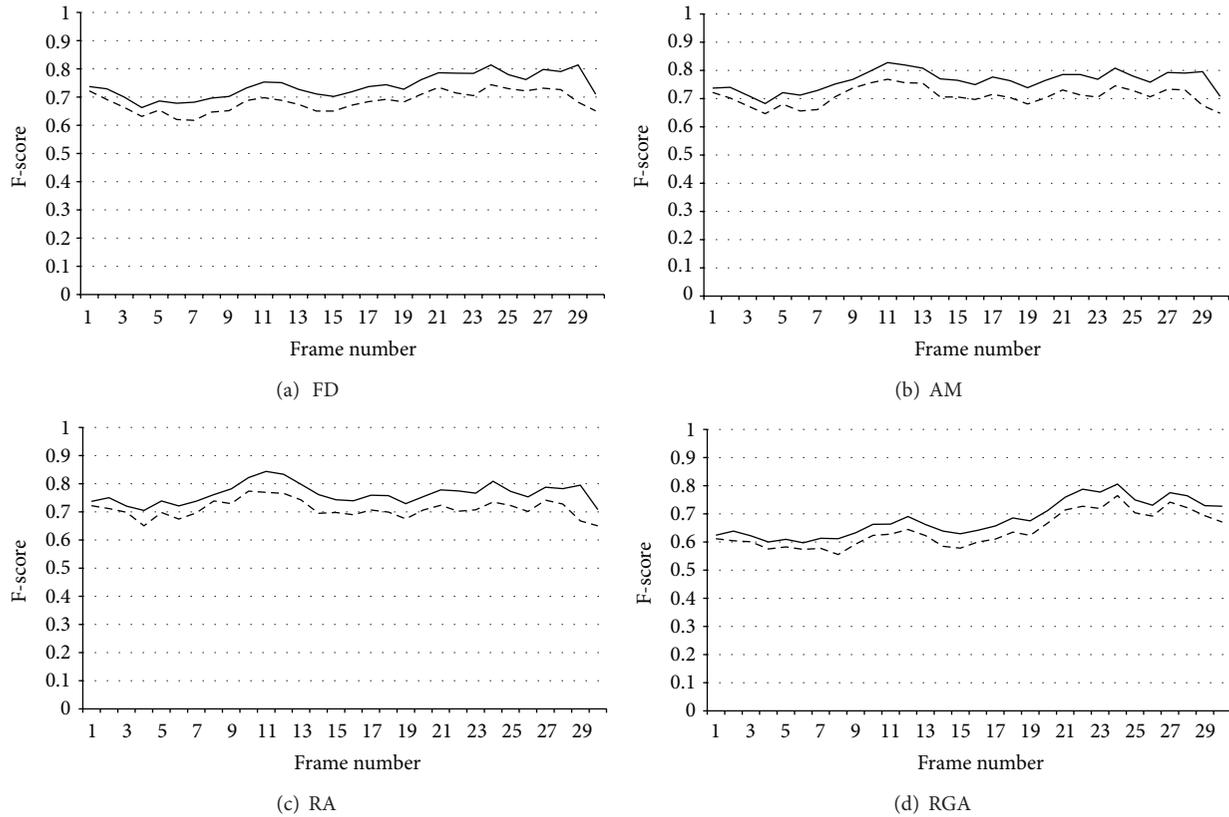


FIGURE 8: Graph of F -score versus frame number for indoor sample.

addition of the SDGD filter in the BGS techniques. Then the values of recall, precision, and F -score were calculated based on mathematical equations (10)–(12).

Based on the findings shown in Tables 1 and 2, we can see that the number of TPs has increased significantly, which proves that our technique is able to detect more of the compound blob than the original methods. This finding is also in line with the increment of precision values. Tables 1 and 2 also show that the F -score increased for both samples when we add the SDGD filter.

The graphs in Figures 7 and 8 show the F -score trends in both A1 and B1 for the four BGS techniques, namely, FD, AM, RA, and RGA. The solid lines represent the F -score results using our proposed technique, that is, with the SDGD filter, whereas the dashed lines represent the F -score results without using the SDGD filter. Based on Figures 7 and 8, higher F -score values were noted for the A1 and B1 videos when using the proposed hybrid technique compared with those when using the basic BGS techniques. Thus, our proposed technique improves upon traditional BGS techniques.

To confirm the effectiveness of the proposed technique, we tested the algorithm in six different videos with six different backgrounds, as obtained from the five different databases. Table 3 shows the performance of FD, AM, RA, and RGA with and without the SDGD filter, in terms of the F -score and average accuracy percentage of all six video samples. Columns 5 and 8 show the percentage of improvement for both F -score and average accuracy

percentage. Based on Table 3, the use of an SDGD filter improved the average F -score values for all data compared with the values produced by the methods without SDGD. Column 5 of Table 3 shows that the F -score values improved by 1% to 9%. Therefore, the proposed technique, compared with existing techniques, enhances object extraction.

Additionally, Table 3 depicts an increment in average accuracy percentage for each tested technique except for videos A2 and E1. A2 had poor video quality because the footage was taken in a corridor without proper lighting. Because of the poor video quality and bad lighting condition, the SDGD filter was unable to segment the foreground subjects properly and produced an unwanted shadow in the foreground. Nevertheless, our proposed technique is capable of detecting foreground pixels with over 90% accuracy on all videos tested.

Meanwhile, Table 4 exhibits the results of classifying human and non-human recognition based on the segmented frame generated by the proposed technique. In Table 4, ANN successfully recognized human and non-human images from the generated frames, using the improved FD technique. Incorrect classifications were minimal in all ten experiments.

Figure 9 presents a matrix describing the overall results of classification testing. The average recognition rate for human and non-human categories was 98.78% and 98.72%, respectively. The rate is higher for both categories as our algorithm provided good human and non-human blob images, which

TABLE 3: Performance of each technique.

Name	Method	Evaluation performance based on average value					
		<i>F</i> -score			Accuracy (%)		
		Without SDGD	With SDGD	%im*	Without SDGD	With SDGD	%im*
A1	FD	0.75	0.78	3	98.45	98.47	0.02
	AM	0.69	0.75	6	98.16	98.29	0.13
	RA	0.74	0.77	3	98.42	98.44	0.02
	RGA	0.79	0.81	2	98.42	98.47	0.03
A2	FD	0.78	0.79	1	96.74	96.65	-0.09
	AM	0.80	0.81	1	97.31	97.14	-0.17
	RA	0.76	0.78	2	96.91	96.83	-0.08
	RGA	0.76	0.77	1	95.75	95.65	-0.10
B1	FD	0.68	0.74	6	97.99	98.12	0.13
	AM	0.71	0.76	5	98.22	98.37	0.15
	RA	0.71	0.76	5	98.25	98.38	0.13
	RGA	0.64	0.68	4	97.47	97.51	0.04
C1	FD	0.84	0.85	1	95.64	95.69	0.05
	AM	0.79	0.80	1	94.87	95.03	0.16
	RA	0.57	0.65	8	91.69	92.17	0.48
	RGA	0.8	0.81	1	94.95	95.07	0.12
D1	FD	0.68	0.71	3	97.84	98.07	0.23
	AM	0.71	0.75	4	98.13	98.37	0.22
	RA	0.71	0.74	3	98.14	98.36	0.22
	RGA	0.70	0.74	4	98.04	98.25	0.21
E1	FD	0.80	0.81	1	96.66	96.69	0.03
	AM	0.80	0.81	1	96.70	96.71	0.01
	RA	0.80	0.81	1	96.69	96.70	0.01
	RGA	0.68	0.77	9	96.45	96.50	0.05

*im: percent improvement.

		Target class		(%)
		Human	Nonhuman	
Output class	Human	4939	64	98.78
	Non-human	61	4936	98.72
		98.78	98.72	98.75

FIGURE 9: Confusion matrix of the overall performance.

allowed ANN to distinguish between the two categories. Further, the overall performance in both classes was 98.75%. The findings confirmed that the proposed algorithm, which combines FD with the SDGD filter, could generate better silhouette images, thereby facilitating recognition of human and non-human images in the segmented images/frames.

From the matrix, the information of TP, TN, FP, and FN can be extracted; hence Table 5 shows results of the statistical analysis done on the classification findings by ANN.

5. Conclusion

We presented a new hybrid approach that incorporates the SDGD filter with four basic BGS techniques, namely, FD, AM, RA, and RGA. This hybrid technique enhanced segmentation performance, as indicated by the *F*-score values and average accuracy percentages. The technique was tested on six different videos from five different databases, and each video was taken either indoors or outdoors and showed different scenes. An ANN classifier was used to classify human and nonhuman images appearing in the segmented images generated by our algorithm. As the algorithm was capable of providing good blob images, ANN recognition of human and non-human images in the silhouette images was facilitated.

Although computational time increased, this aspect is acceptable considering the enhancement and the characteristics of second-order derivatives. Therefore, this study is

TABLE 4: Classification results.

Exp #		Human	Non-human	%
1	Correct	498	492	99
	Missed	2	8	1.0
2	Correct	499	491	99.0
	Missed	1	9	1.0
3	Correct	498	495	99.3
	Missed	2	5	0.7
4	Correct	495	500	99.5
	Missed	5	0	0.5
5	Correct	481	492	97.3
	Missed	19	8	2.7
6	Correct	497	490	98.7
	Missed	3	10	1.3
7	Correct	484	496	98.0
	Missed	16	4	2.0
8	Correct	493	498	99.1
	Missed	7	2	0.9
9	Correct	497	496	99.3
	Missed	3	4	0.7
10	Correct	497	486	98.3
	Missed	3	14	1.7

TABLE 5: Performance for ANN classification.

TP	TN	FP	FN	Recall	Precision	F-Score
4939	4936	61	64	0.987	0.988	0.987

valid and suitable for implementation in non-real-time applications. The proposed hybrid technique can improve upon traditional BGS techniques as indicated by the improved F -score, accuracy, and ANN recognition values after testing various data sources and data environments. The technique can also be considered in detecting moving objects in non-real-time applications, such as investigations of human actions or traffic conditions.

Acknowledgments

This work was supported by Universiti Kebangsaan Malaysia (UKM) research Grant DPP-2013-003 and Ministry of Higher Education (MoHE) research Grant LRGS/TD/2011/ICT/04/02.

References

- [1] A. M. McIvor, "Background subtraction techniques," in *Proceedings of the Image and Vision Computing Conference*, Auckland, New Zealand, 2000.
- [2] M. H. Sigari, N. Mozayani, and H. M. Pourreza, "Fuzzy running average and fuzzy background subtraction: concepts and application," *International Journal of Computer Science and Network Security*, vol. 8, pp. 138–143, 2008.
- [3] Y. Zheng and L. Fan, "Moving object detection based on running average background and temporal difference," in *Proceedings of the IEEE International Conference on Intelligent Systems and Knowledge Engineering (ISKE '10)*, pp. 270–272, November 2010.
- [4] J. Park, A. Tabb, and A. C. Kak, "Hierarchical data structure for real-time background subtraction," in *Proceedings of the IEEE International Conference on Image Processing (ICIP '06)*, pp. 1849–1852, October 2006.
- [5] C. R. Wren, A. Azarbayeni, T. Darrel, and A. P. Petland, "Pfinder: real-time tracking of the human body," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 780–785, 1997.
- [6] Z. Tang, Z. Miao, and Y. Wan, "Background subtraction using running Gaussian average and frame difference," in *Proceedings of the International Conference Entertainment Computing (ICEC '07)*, vol. 4740 of *Lecture Notes in Computer Science*, pp. 411–414, 2007.
- [7] Z. He, Y. Liu, H. Yu, and X. Ye, "Optimized algorithm for traffic information collection in an embedded system," in *Proceedings of the IEEE Congress on Image and Signal Processing*, May 2008.
- [8] A. Singh, S. Sawan, M. Hanmandlu, V. K. Madasu, and B. C. Lovell, "An abandoned object detection system based on dual background segmentation," in *Proceedings of the 6th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS '09)*, pp. 352–357, IEEE Press, September 2009.
- [9] S. Su and Y. Chen, "Moving object segmentation using improved running Gaussian average background model," in *Proceedings of the International conference on Digital Image Computing: Techniques and Applications*, 2008.
- [10] H. Kim, R. Sakamoto, I. Kitahara, T. Toriyama, and K. Kogure, "Background subtraction using generalised Gaussian family model," *Electronics Letters*, vol. 44, no. 3, pp. 189–190, 2008.
- [11] Z. H. Huang and K. W. Chau, "A new image thresholding method based on Gaussian mixture model," *Journal Applied Mathematics and Computation*, vol. 205, pp. 899–907, 2008.
- [12] Y. Liu, H. Yao, W. Gao, X. Chen, and D. Zhao, "Nonparametric background generation," in *Proceedings of the 18th International Conference on Pattern Recognition (ICPR '06)*, pp. 916–919, August 2006.
- [13] P. Vu, V. Phong, T. H. Vu, and H. B. Le, "GPU implementation of Extended Gaussian mixture model for background subtraction," in *Proceedings of the 8th IEEE-RIVF International Conference on Computing and Communication Technologies: Research, Innovation and Vision for the Future (RIVF '10)*, November 2010.
- [14] A. Mittal and N. Paragios, "Motion-based background subtraction using adaptive kernel density estimation," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '04)*, pp. II302–II309, July 2004.
- [15] F. Porikli and O. Tuzel, "Human body tracking by adaptive background models and mean shift analysis," in *Proceedings of the IEEE International Workshop on Performance Evaluation of Tracking and Surveillance (PETS '03)*, July 2003.
- [16] I. J. Young, J. J. Gerbrands, and L. J. van Vliet, "Fundamentals of Image Processing," Version 2. 3, 1995–2007.
- [17] V. G. Narendra and K. S. Hareesh, "Study and comparison of various image edge detection techniques," *International Journal of Image Processing*, vol. 4, no. 2, article 83, 2009.
- [18] J. F. Canny, "A computational approach to edge detection," *IEEE Transaction Pattern Analysis and Machine Intelligent*, vol. 8, no. 6, pp. 679–697, 1986.

- [19] M. P. Persoon, I. W. O. Serlie, F. H. Post, R. Truyen, and F. M. Vos, "Visualization of noisy and biased volume data using first and second order derivative techniques," in *Proceedings of the 14th IEEE Visualization Conference*, pp. 379–385, October 2003.
- [20] I. T. Young, "Generalized convolutional filtering," in *Proceedings of the 19th CERN School of Computing*, pp. 51–65, 1996.
- [21] "Image processing fundamentals: derivative-based operation," 2011, <http://www.mif.vu.lt/atpazinimas/dip/FIP/fip-Derivati.html>.
- [22] P. W. Verbeek and L. J. van Vliet, "Location error of curved edges in low-pass filtered 2-D and 3-D images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 16, no. 7, pp. 726–733, 1994.
- [23] M. Hagara and J. Moravcik, "PLUS operator for edge detection in digital images," in *Proceedings of the International Conference of Radioelektronika*, pp. 467–470, 2002.
- [24] Y. Nader El-Glaly, *Development of PDE-Based Digital Inpainting Algorithm Applied to Missing Data in Digital Images [M.S. thesis]*, Ain Shams University, 2007.
- [25] R. G. Aarnick, J. de la Rosette, W. Feitz, F. Debruyne, and H. Wijkstra, "A preprocessing algorithm for edge detection with multiple scales of resolution," *European Journal of Ultrasound*, vol. 5, pp. 113–126, 1997.
- [26] N. Lu and J. Wang, "Motion detection based on accumulative optical flow and double background filtering," in *Proceedings of the World Congress on Engineering*, vol. 1, 2007.
- [27] Zaki, W. M. D. W, A. Hussain, and M. Hedayati, "Moving object detection using Keypoints reference model," *EURASIP Journal on Image and Video Processing*, vol. 2011, 13 pages, 2011.
- [28] F. Y. A. Rahman, A. Hussain, N. M. Tahir, S. A. Samad, and M. H. M. Saad, "Hybrid background subtraction techniques with second derivative on gradient direction filter," in *Proceedings of the International Workshop on Advanced Image Technology (IWAIT '12)*, 2012.
- [29] S. M. Al-Garni and A. A. Abdennour, "Moving vehicle detection using automatic background extraction," in *proceedings of the World Academy of Science, Engineering and Technology*, vol. 24, pp. 82–86, Sydney, Australia, December 2006.
- [30] "CMU graphic lab motion capture database," 2010, <http://mocap.cs.cmu.edu/>.
- [31] R. Gross and J. Shi, "The CMU motion of body (MoBo) database," Technical Report CMU-RI-TR-01-18, Robotics Institute, Carnegie Mellon University, 2001.
- [32] S. Singh, S. A. Velastin, and H. Ragheb, "MuHAVi: a multicamera human action video dataset for the evaluation of action recognition methods," in *Proceedings of the 7th IEEE International Conference on Advanced Video and Signal Based (AVSS '10)*, pp. 48–55, Boston, Mass, USA, September 2010.
- [33] H. Kuehne, H. Jhuang, E. Garrote, T. Poggio, and T. Serre, "HMDB: a large video database for human motion recognition," in *Proceedings of the IEEE International Conference on Computer Vision (ICCV '11)*, pp. 2556–2563, November 2011.
- [34] F. Y. A. Rahman, A. Hussain, N. M. Tahir, and W. M. D. Zaki, "Modeling of initial reference frame for background subtraction," in *Proceedings of the 6th International Colloquium on Signal Processing and Its Applications (CSPA '10)*, pp. 125–128, 2010.
- [35] E. Fauske, L. M. Eliassen, and R. H. Bakken, "A comparison of learning based background subtraction," in *Proceedings of the Norwegian Artificial Intelligens Symposium (NAIS)*, pp. 181–192, 2009.
- [36] H. Kim, B. Ku, D. K. Han, S. Kang, and H. Ko, "Adaptive selection model in block-based background subtraction," *Electronics Letter*, vol. 48, no. 8, 2012.

