An Efficient Real Time Moving Object Detection Method for Video Surveillance System

Pranab Kumar Dhar[*], Mohammad Ibrahim Khan[*], Ashoke Kumar Sen Gupta[**], D.M.H. Hasan[*] and Jong-Myon Kim[***1]

[*]*Department of Computer Science and Engineering, Chittagong University of Engineering and Technology, Chittagong-4349, Bangladesh*
[**]*Department of Electrical and Electronic Engineering, Chittagong University of Engineering and Technology, Chittagong-4349, Bangladesh*
[***]*School of Computer Engineering and Information Technology, University of Ulsan, San 29, Mugeo 2-Dong, Nam-Ku, Ulsan, 680-749, Korea*

{pranab_cse, muhammad_ikhancuet, simon_cuet}@yahoo.com, aksengupta@cuet.ac.bd, jmkim07@ulsan.ac.kr

Abstract

Moving object detection has been widely used in diverse discipline such as intelligent transportation systems, airport security systems, video monitoring systems, and so on. In this paper, we propose an efficient moving object detection method using enhanced edge localization mechanism and gradient directional masking for video surveillance system. In our proposed method, gradient map images are initially generated from the input and background images using a gradient operator. The gradient difference map is then calculated from gradient map images. The moving object is then detected by using appropriate directional masking and thresholding. Simulation results indicate that the proposed method consistently performs well under different illumination conditions including indoor, outdoor, sunny, and foggy cases. Moreover, it outperforms well known edge based method in terms of detecting moving objects and error rate. Moreover, the proposed method is computationally faster and it is applicable for detecting moving object in real-time.

Keywords: *Video Surveillance, Gradient Directional Masking, Edge Localization, and Object Detection*

1. Introduction

Moving object detection in real time is a challenging task in visual surveillance systems. It often acts as an initial step for further processing such as classification of the detected moving object. In order to perform more sophisticated operations such as classification, we need to first develop an efficient and accurate method for detecting moving objects. A typical moving object detection algorithm has the following features: (a) estimation of the stationary part of the observed scene (background) and obtaining its statistical characteristics (b) obtaining difference images of frames taken at different times and difference images of the sequence with the image of the stationary part of the scene (c) discrimination of regions belonging to objects, identification of these objects, determining the trajectories of motion of these objects, and their classification (d) adaptation of the stationary part of the background for changing detection conditions

[1] Corresponding author

and for changing the content of the scene (e) registration of situations and necessary messages. One of the simplest and popular method for moving object detection is a background subtraction method which often uses background modeling, but it takes long time to detect moving objects [1-6]. Temporal difference method is very simple and it can detect objects in real time, but it does not provide robustness against illumination change. The foreground extraction problem is dealt with the change detection techniques, which can be pixel based or region based. Simple differencing is the most intuitive by arguing that a change at a pixel location occurs when the intensity difference of the corresponding pixels in two images exceeds a certain threshold. However, it is sensitive to pixel variation resulting from noise and illumination changes, which frequently occur in complex natural environments. Texture based boundary evaluation methods are not reliable for real time moving object detection.

This paper proposes a new method to detect moving objects from a stationary background based on the improved edge localization mechanism and gradient directional masking for video surveillance systems. In the proposed method, gradient map images are generated from the input and background images using a gradient operator. The gradient difference map is then calculated from gradient map images. Finally, moving objects are detected by using appropriate directional masking and thresholding. Simulation results indicate that the proposed method provides better results than well-known edge based methods under different illumination conditions, including indoor, outdoor, sunny, and foggy cases for detecting moving objects. Moreover, it detects objects more accurately from input images than existing methods.

The rest of this paper is organized as follows. Section 2 provides a brief description of related research. Section 3 introduces the proposed method for detecting moving objects. Section 4 compares the performance of the proposed method with well-known edge based methods. Section 5 concludes the paper.

2. Related Research

A significant number of moving object detection methods have been reported in recent years. Mahbub et. al., [7] proposed a method using statistical background modeling. This method detects moving objects by matching every edge segment of current frame with every edge segment of background. However, this method fails to detect a moving edge segment which falls into a background edge segment. Islam and Lee [8] proposed a method for moving object tracking using a particle filter in which the shape similarity between a template and estimated regions in the video scene is measured by their normalized cross-correlation of distance transformed images. Dunne and Matuszewski [9] introduced an object detection scheme that utilizes a localized temporal difference change detector and a particle filter type likelihood detector to detect possible trackable objects, and to find a point within a detected object at which a particle filter tracker might be initialized. Shin and Hong [10] described a method in which a clear outline of an object and a loss part are restored by using edge information and a boosting factor according to adjustment corresponding to a change of an input image. The object is finally extracted by analyzing a shadow area of the object generated through the full process and removing the shadow. Hossain et. al., [11] presented a method that utilizes edge segments, but it requires a number of initial training frames for generating background image. Whether the method is pixel based or region based, thresholding of the difference image always presents itself as most challenging. In many cases, the threshold is selected empirically or by trial and error. Obviously, a threshold chosen in this way is ineffective for images with significantly different distributions. As a result, several adaptive threshold selection methods have been proposed. Some of these methods are based on histograms. For example, Otsu's method [12] calculates the best threshold by minimizing the ratio of intraclass and interclass variations, the isodata algorithm [13] searches for the best threshold by an iterative estimation of the mean values of the foreground and background pixels, the triangle algorithm [14] particularly deals with unimodal histograms, Kita [15] analyzes the characteristics of the ridges of clusters on the joint histogram, and Sen and Pal [16] select the threshold by using the fuzzy and rough set theories. Kim and Hwang [17] presented a scheme that adapts an edge difference method to compute current moving edges and temporal moving edges for detecting moving objects. This method fails to handle the changes of dynamic background, resulting in higher false alarm. Dewan et al. [18] proposed an edge segment based approach that utilizes three consecutive frames to compute difference images between each two consecutive frames. This method extracts edges from two difference images and finally detects moving objects using an edge segment matching algorithm. However, it fails to detect slow moving objects because it does not utilize any background modeling. Rosin and Ellis [19] perform thresholding with hysteresis, where the difference image is first thresholded by two levels, and regions in the intermediate range are not considered to be changed unless they are connected with regions generated by the higher threshold. This method scratches the surface of the problem and stops short of tackling the real issue. Kapur et. al., [20] selected thresholds by virtue of the entropy of the image. Gray [21] considered the Euler number, and O'Gorman [22] used image connectivity. Another set of approaches is to assume that the distributions of the changes and the noise of the difference image are Gaussian or Laplacian. For example, Bruzzone and Prieto [23] modeled the difference image as a mixture of two Gaussian distributions, representing changed and unchanged pixels. The means and variances of the class-conditional distributions are then estimated using an expectation maximization algorithm. Lu et. al., [24] introduced a method for extracting moving objects from their background based on multiple adaptive thresholds and boundary evaluation. The major shortcoming of the single-threshold (ST) approach is that it often ends up with a foreground that is either over segmented or under segmented. This comes as no surprise because the foreground and background pixels intertwine in the measurement space, which makes it impossible to have a global threshold that segments well. To alleviate the problem, it appears logical to consider multiple thresholds. Zhou and Hoang [25] proposed a bithreshold method in which pixels in the intermediate range of difference are further evaluated to decide whether they are shadow pixels or not. Reference [26] is similar to [25], except that a silhouette extraction technique is applied on the foreground extraction result to smooth the boundary. Dailey et. al., [27] devised a method that generates two edge maps from the edge difference images of three consecutive frames. Moving edges are extracted by applying logical AND operation between two edge maps. Due to random noise or small camera fluctuation, pixel position of edges may change in consecutive frames. Thus, edge matching is not sufficient for detecting moving objects. This method also fails to detect moving objects with slow motion, which indicates that it is not useful for real-time applications.

3. Background Information

3.1. Edge Detection

Edge detection is the first step to recover information from images. Edges are the significant local changes of intensity in an image. Edges typically occur on the boundary between two different regions in an image. Edge also can be defined as discontinuities in image intensity from one pixel to another. A typical edge detector has the following steps: (a) it suppresses noise as much as possible, without destroying the true edges; (b) it applies a filter to enhance the quality of the edges in the image, (c) it determines which edge pixels should be discarded as noise and which should be retained, (d) it determines the exact location of an edge. An optimal edge detector should satisfy the following criteria: (a) the optimal detector must minimize the probability of false positives (detecting spurious edges caused by noise) as well as that of false negatives (missing real edges), (b) the edges detected must be as close as possible to the true edges, (c) the detector must return one point only for each true edge point; that is, it minimizes the number of local maxima around the true edge created by noise.

3.2. Canny Edge Detector

The Canny edge detector is one of the most commonly used image processing tool to detect edges from image. It has the following steps:

3.2.1 Gray Scale Conversion

In photography and computing, a grayscale digital image is an image in which the value of each pixel is a single sample, that is, it carries only intensity information. Images of this sort, also known as black-and-white, are composed exclusively of shades of gray, varying from black at the weakest intensity to white at the strongest. To convert any color to a grayscale representation of its luminance, first one must obtain the values of its red, green, and blue (RGB) primaries in linear intensity encoding, by gamma expansion. Then, add together 30% of the red value, 59% of the green value and 11% of the blue value.

3.2.2 Noise Reduction

The Canny edge detector uses a filter based on the first derivative of a Gaussian, because it is susceptible to noise exists in raw unprocessed image data. Thus, at first the raw image is convolved with a Gaussian filter. The result is a slightly blurred version of the original which is not affected by a single noisy pixel to any significant degree.

3.2.3 Gradient Computation

An edge in an image may point in a variety of directions, thus the canny algorithm uses four filters to detect horizontal, vertical and diagonal edges in the blurred image. The edge detection operator returns a value for the first derivative in the horizontal direction (G_y) and the vertical direction (G_x). From this the edge gradient and direction can be determined by the following equations:

$$G = \sqrt{G_x^2 + G_y^2} \qquad \textbf{(1)}$$

$$\theta = \arctan\left(\frac{G_y}{G_x}\right) \quad \textbf{(2)}$$

The edge direction angle is rounded to one of four angles representing vertical, horizontal, and the two diagonals (for example: 0, 45, 90 and 135 degrees).

3.2.4 Non-maxima Suppression

Non-maxima suppression (NMS) can be positively formulated as a local maximum search, where a local maximum is greater than all its neighbors (excluding itself). For a given n, the neighborhood of any pixel consists in the one dimensional case of the n pixels to its left and right side (referred to as (2n+1) neighborhood and in the two dimensional case of the quadratic (2n+1) × (2n+1) region centered around the pixel. The same value may appear several times in an image, thus the question arises, which pixel should be suppressed in case of a tie. In practice, either all or all but one is suppressed according to some ordering.

3.2.5 Hysteresis Thresholding

Intensity gradients which have large values are more likely correspond to edges. In most cases, it is impossible to select a threshold at which a given intensity gradient switches from one edge to other edge. Therefore Canny edge detector uses thresholding with hysteresis.

Thresholding with hysteresis requires two threshold values: one is high value and another is low value. Consider that important edges are continuous in the image which allows to follow a section of a given line and to discard a few noisy pixels that do not constitute a line but produced large gradients. Therefore, initially a high threshold value is considered to calculate the genuine edges. Starting from these, using the directional information, edges can be traced through the image. While tracing an edge, a lower threshold value is applied which allows to trace faint sections of edges.

The key parameter in the thresholding process is the selection of the threshold value. Several different methods for selecting a threshold have been introduced. Users can manually select a threshold value, or a thresholding algorithm can compute a value automatically, which is known as automatic thresholding. A simple method would be to select the mean or median value of the pixels. If the object pixels are brighter than the background pixel, they should also be brighter than the average. In a noiseless image with uniform background and object values, the mean or median will work well as the threshold, however, in case of noisy image this does not provide good result. A more sophisticated approach is to create a histogram of the image pixel intensities and use the valley point as the threshold. The histogram based thresholding approach considers an average value for the background and object pixels, but the actual pixel values have some variation around the average values. However, this may be computationally expensive, and also image histograms do not have clearly defined valley points, resulting in difficulties to select an accurate threshold value.

2.4 Masking

In computer graphics, when a given image is intended to be placed over a background, the transparent areas can be specified through a binary mask. Thus, for each intended image there are actually two bitmaps: the actual image, in which the unused areas are assigned a pixel value with all bits set to 0's, an additional *mask*, in which the corresponding image areas are assigned a pixel value of all bits set to 0's and the surrounding areas are assigned a value of all bits set to 1's. For example, in an image, black pixels have all-zero bits and white pixels have all-one bits. At run time, to put the image on the screen over the background, the program first masks the screen pixel's bits with the image mask at the desired coordinates using the bitwise AND operation. This preserves the background pixels of the transparent areas while resets the bits of the pixels with zeros which will be obscured by the overlapped image. Then, the program renders the image pixel's bits by blending them with the background pixel's bits using the bitwise OR operation. In this way, the image pixels are appropriately placed while keeping the background pixels. The result is a perfect compound of the image over the background. This technique is widely used for pointing device cursors, for creating characters in typical two dimensional videogames, and other image mixing applications.

4. Proposed Moving Object Detection Method

Our proposed method aims to extract moving objects from an input image by utilizing their background. As depicted in Figure 1, the proposed method consists of four steps: (a) gradient map and edge map generation; (b) gradient difference image calculation; (c) gradient masking and threshold selection; (d) update background. Details of these steps are described in the following sections. Figures 2 (a) and 2(b) show the input and background image of an indoor video sequence.

4.1. Gradient Map and Edge Map Generation

Edges are local changes in the image and are important features for analyzing images. Most of the edge based methods used a conventional canny edge detector to generate edge map. This paper proposes a new method that detects edges directly from RGB images (i.e. there is no grayscale conversion which reduces computational cost). It computes gradient magnitude and direction by using the Euclidian distance and vector angle concept. Figure 3 shows two 3x3 orthogonal and diagonal masks that are used for gradient computation. For each pixel, let $\underline{V_1}$ be the vector from the origin of the RGB color space to that pixel and $\underline{V_2}$ be the vector along the main diagonal which represents gray scale line. Then, the angular distance of that pixel from the gray scale line through the origin can be determined as follows:

$$\gamma = \underline{V_1} . \underline{V_2} / (|\underline{V_1}|*|\underline{V_2}|) \qquad \textbf{(3)}$$

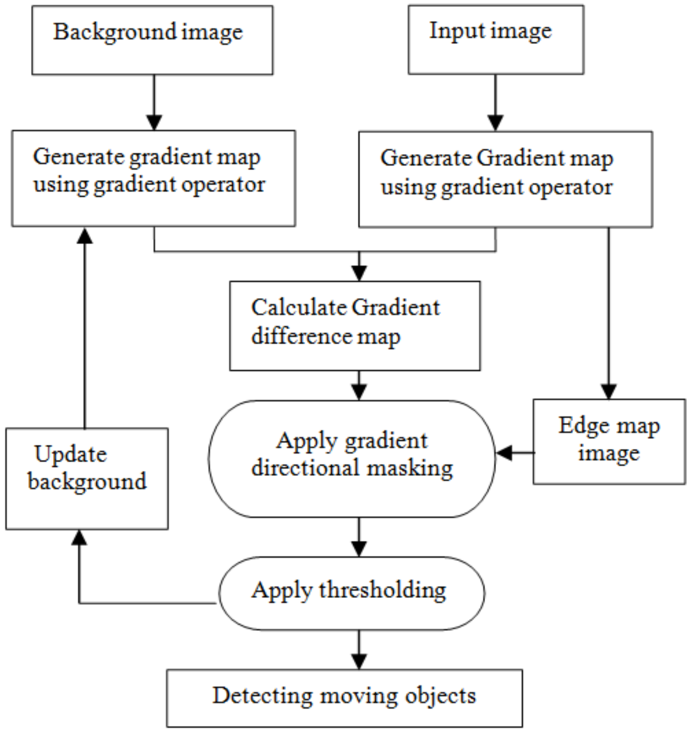


Figure 1. An Overview of the Proposed Method

(a) (b)

Figure 2. (a) Input Image; (b) Background Image

The 3x3 orthogonal mask is applied to each pixel of both input and background images to calculate the Euclidian distance between each two adjacent neighbors in horizontal and vertical direction. Let the distance in horizontal direction be E_H and the distance in vertical direction be E_V. Similarly, the 3x3 diagonal mask is applied to calculate the Euclidian distance between each two adjacent neighbors in two diagonal directions. Let the distance along one diagonal direction be E_{D1} and the distance along another direction be E_{D2}. Finally, the total distance between each two adjacent neighbors along four directions can be approximated in the gray scale line using the following equation:

$$\delta M_{(x,y)} = \sum E * \gamma, \quad \textbf{(4)}$$

where $\sum E$ is the total distance between each pair of adjacent neighbors in each four directions and $\delta M_{(x,y)}$ is the gradient map image. The gradient direction is perpendicular to edge direction. This direction specifies that the maximum changes occur in each direction. In the proposed method, the gradient direction is taken as the direction in which the Euclidian distance between two adjacent neighbors is maximum i.e.

$$\theta_{(x,y)}=(H,V,D1,D2) \mid MAX(E_H, E_V, E_{D1}, E_{D2}), \quad \textbf{(5)}$$

where *H*, *V*, *D1*, *D2* represent horizontal direction, vertical direction, 1st diagonal direction, and 2nd diagonal direction, respectively.

1	0	1
1	0	1
1	0	1

(a)

1	1	1
0	0	0
1	1	1

(b)

Figure 3. (a) Orthogonal Mask; (b) Diagonal Mask

The edge map image *EM* is extracted from the gradient map image by applying non-maxima suppression and thresholding. If $\delta M_{(x,y)}$ is greater than a certain threshold value, then $EM_{(x,y)}$ = True. Otherwise, $EM_{(x,y)}$ = False. Figs. 4(a) and 4(b) show the gradient map and edge map of the input image of Fig 2(a).

4.2. Gradient Difference Image Calculation

Gradient difference image shows the structural difference between two images. Gradient difference image is calculated from the input and background gradient maps. This gradient difference is significantly large when changes occur between the images. In the proposed method, the gradient difference image ΔM is computed by the following equation:

$$\Delta M = | \delta M_{I(x,y)} - \delta M_{B(x,y)} |, \quad \textbf{(6)}$$

where $\delta M_{I(x,y)}$ is the input gradient map and $\delta M_{B(x,y)}$ is the background gradient map.

4.3. Masking and Threshold Selection

Gradient difference map discriminates the structure of the input and background image. This gradient structure provides robustness against illumination change. Most existing edge based methods utilize proper edge segment matching criteria to differentiate between moving and background edges. However, when a moving edge falls just over a background edge, the edge segment matching algorithm detects the moving edge as a background edge. To overcome this problem, the proposed method utilizes a gradient map which holds the position of edges as well as the information of edges. This information can be used to differentiate between overlapped moving and background edges.

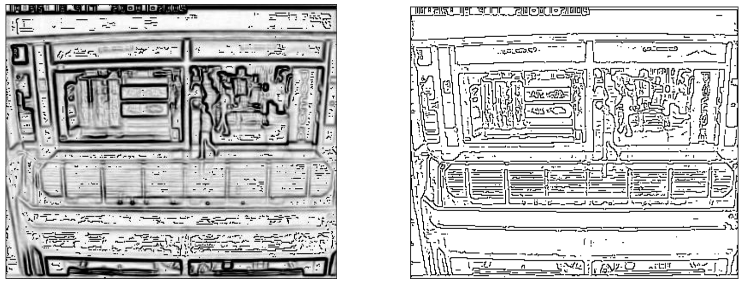


Figure 4. (a) Gradient Map Image; (b) Edge Map Image

Figure 5 shows the gradient difference image of the input image of Figure 2(a).

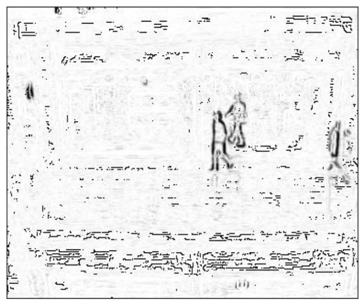


Figure 5. Gradient Difference Image

The proposed method introduces four directional masks such as Horizontal mask (M_H), Vertical mask (M_V), 1st Diagonal mask (M_{D1}), and 2nd Diagonal mask (M_{D2}) for four gradient directions as shown in Figure 6. Gradient direction is perpendicular to edge direction. The gradient structure of an edge of both background and input gradient maps exhibits almost similar pattern. Hence, the similarity between two gradient maps at a certain pixel is calculated by using directional masking on gradient difference map. If this similarity is less than a certain threshold value, the pixel is then considered as a background pixel. For simplicity, the number of identical gradient directions on each mask having non-zero value can also be considered.

For each edge pixel of the input edge map, if $EM_{(x,y)}$ is true, then the respective mask is selected according to the gradient direction for that pixel. If the direction $\theta_{(x,y)}$ = H, the horizontal mask M_H is selected. If the direction $\theta_{(x,y)}$ = V, the vertical mask M_V is selected. The 1st diagonal mask (M_{D1}) is selected when the direction $\theta_{(x,y)}$ = D1. If the direction $\theta_{(x,y)}$ = D2, the 2nd diagonal mask (M_{D2}) is selected. The selected mask is applied to the gradient difference map. The value $\Delta\delta_{(x,y)}$ at a position (x,y), is calculated by the following equation:

$$\Delta\delta_{(x,y)} = \sum \Delta M_{(x+m,y+n)} * Mask_{(m,n)} , \qquad \textbf{(7)}$$

where m and n vary from -1 to 1 and $Mask_{(m,n)}$ is the mask value at position (m,n).

The threshold value for the mask at any pixel can be determined as follows:

$$T_{(x,y)} = (\delta M_{I(x,y)} + \mu * S) / (S-M), \qquad \textbf{(8)}$$

where $\delta M_{I(x,y)}$ is the input gradient map, and $M \times M$ is the size of the mask.

S= $M*(M-1)$ and $\mu = \sum \delta M_{I\,(x,y)} / N$, where N is the number of pixels for which $EM_{(x,y)}$ is True. If $\Delta\delta_{(x,y)} < T_{(x,y)}$, the pixel is then considered as a background pixel. Otherwise, it is a moving pixel.

1	0	1
1	0	1
1	0	1

(a)

1	1	1
0	0	0
1	1	1

(b)

1	1	0
1	0	1
0	1	1

(c)

0	1	1
1	0	1
1	1	0

(d)

Figure 6. (a) Horizontal Mask M_H; (b) Vertical Mask M_V; (c) 1st Diagonal Mask M_{D1}; (d) 2nd Diagonal Mask M_{D2}

4.4. Update background

Updating background is very important to adapt the changes in the background. Most of existing methods require some reference background images to model the background. Random noise may affect any pixel in the background, but masking based detection alleviates this affect. The proposed method only updates the background pixels. As the gradient structure is very consistent, each background pixel *(x, y)* is updated as follows:

$$\delta M_{B(x,y)} = \beta * \delta M_{I(x,y)} + (1 - \beta) * \delta M_{B(x,y)}, \qquad \textbf{(9)}$$

where $\delta M_{B(x,y)}$ is the background gradient map, $\delta M_{I(x,y)}$ is the input gradient map, and β is the learning rate. The selected value for β is 0.5. Figure 7 shows the updated background image for the background image of Figure 2(b).

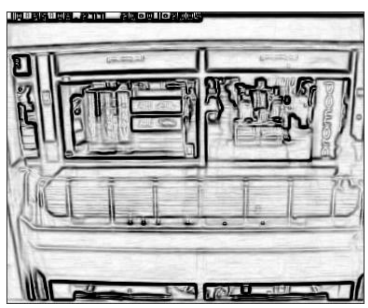
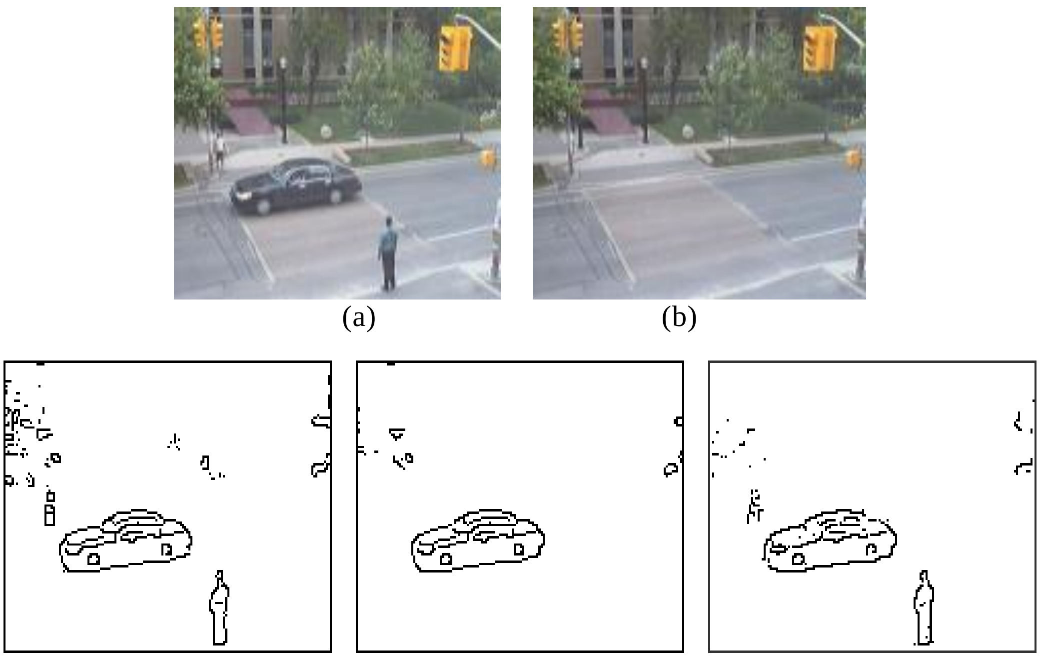


Figure 7. Updated Background Image

5. Simulation Results and Discussion

In this section, the performance of the proposed method is compared with the Kim & Hwang method [7] and Dewan & Chae method [8]. Experiments carried out with several video sequences captured from indoor, outdoor, sunny, and foggy video sequences to verify the effectiveness of the proposed method. We used the video format of size 640x520 and Intel Pentium IV 3.06 GHz processor.

Figure 8 shows experimental results for an outdoor video sequence. Figures 8(a) and 8(b) show the 952th input frame and the background image, respectively. Figures 8(c), 8(d), and 8(e) show the detection results obtained by the Kim & Hwang method, the Dewan & Chae method, and the proposed method, respectively. Figure 8(c) shows that the leaves of the blowing tree with small vibrations have strong effects on detecting objects. If the leaves of the tree change their position, a significant difference in each pixel at that location causes false detection. The Dewan & Chae method failed to detect moving objects which are suddenly stopped because it uses two consecutive difference images among three consecutive frames as shown in Figure 8(d). Moving object in first two consecutive frames causes significant difference in the first difference image. If any object suddenly stopped at the third frame, the second difference image contains no significant information. Thus, no edges are found in that image. Figure 8(e) shows that the proposed method detects moving objects except for the leaves of the blowing tree with small vibrations because of using gradient directional masking.



(a) (b)

(c) (d) (e)

Figure 8. (a) the 952th Input Frame (b) background Image; Detected Moving Object by (c) the Kim & Hwang Method; (d) the Dewan & Chae Method; (e) the Proposed Method

Figure 9 shows experimental results for a foggy video sequence. Figures 9(a) and 9(b) show the 126th input frame and the background image, respectively. Figures 9(c), 9(d), and 9(e) show the detection results obtained by the Kim & Hwang method, the Dewan & Chae method, and the proposed method, respectively. Figure 9(c) shows that the Kim and Hwang method failed to adapt the background changes because it does not uses background updating. In addition, this method also failed to detect slow motion objects effectively. The proposed method overcomes those problems. It properly detected the moving objects as shown in Figure 9(e).

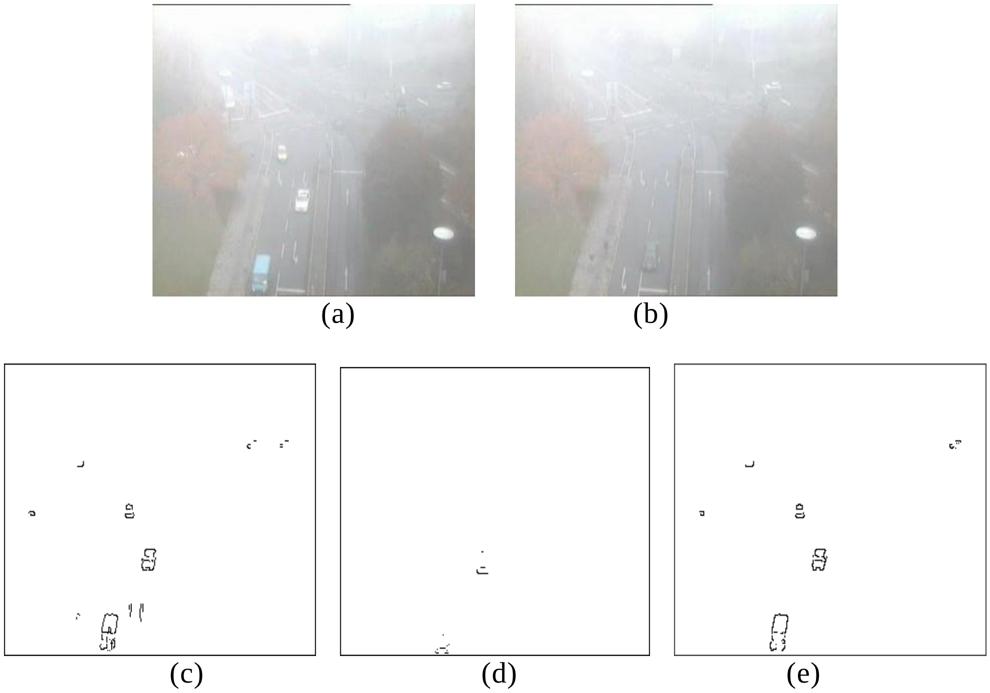


Figure 9. (a) the 126th Input Frame; (b) Background Image; Detected Moving Object by (c) the Kim & Hwang Method; (d) the Dewan & Chae Method; (e) the Proposed Method

Figure 10 shows experimental results for an indoor video sequence. Figures 10(a) and 10(b) show the 267th input frame and the background image, respectively. Figures 10(c), 10(d), and 10(e) show the detection results obtained by the Kim & Hwang method, the Dewan & Chae method, and the proposed method, respectively. Figure 10(c) shows that the Kim & Hwang method highly affected by shadow because of using conventional edge detection technique which is very sensitive to intensity change. Figure 10(d) shows that the Dewan & Chae method produces multiple responses for a single moving edge. This is because slow motion objects do not provide significant difference in consecutive frames. Figure 10(e) shows that the proposed method overcomes these problems.

Figure 11 shows the experimental results for a sunny video sequence. Figure 11(a) and Figure 11(b) show the 339th input frame and the background image respectively. Figures 11(c), 11(d), and 11(e) show the detection results obtained by Kim and Hwang method [7], Dewan and Chae method [8] and the proposed method. Figure 11(c) shows that due to illumination variation Kim and Hwang method detects a lot of scattered edge pixels. The area illuminated by sunlight changes their position with time. As this method does not utilize background updating mechanism, this change detects many false edges. Figure 11(d) shows that some moving objects with slow motions are absent in the detected region. Moreover, this method fails to detect properly the overlapped objects from consecutive frames. On the other hand, our proposed method effectively detected moving objects from the selected frame because of using edge localization mechanism and gradient directional masking.

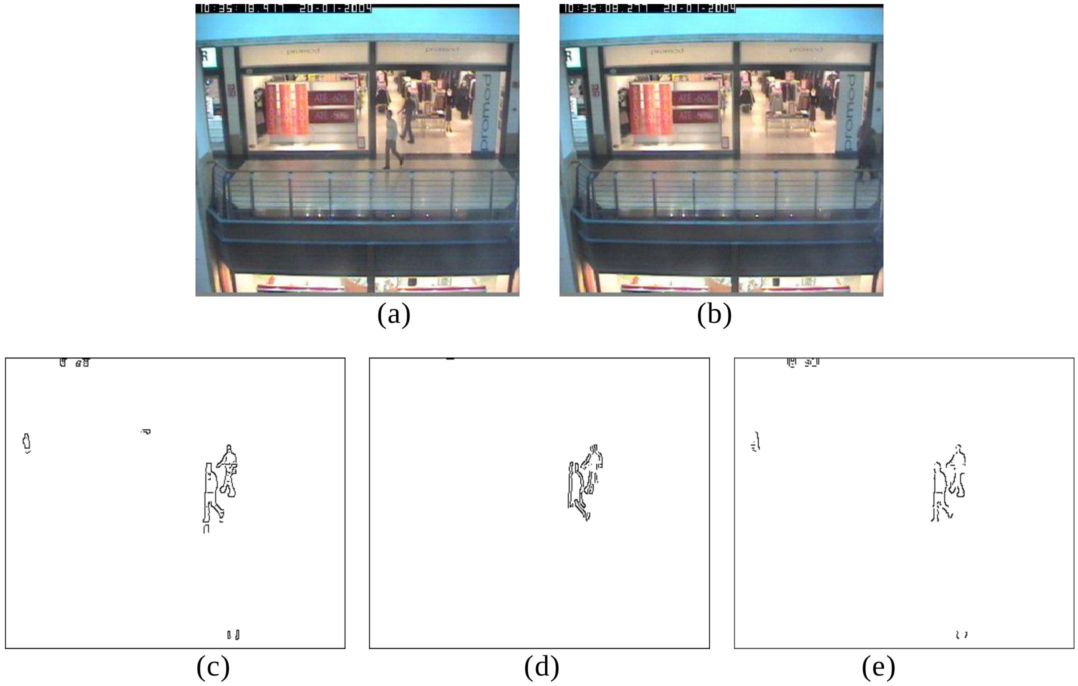
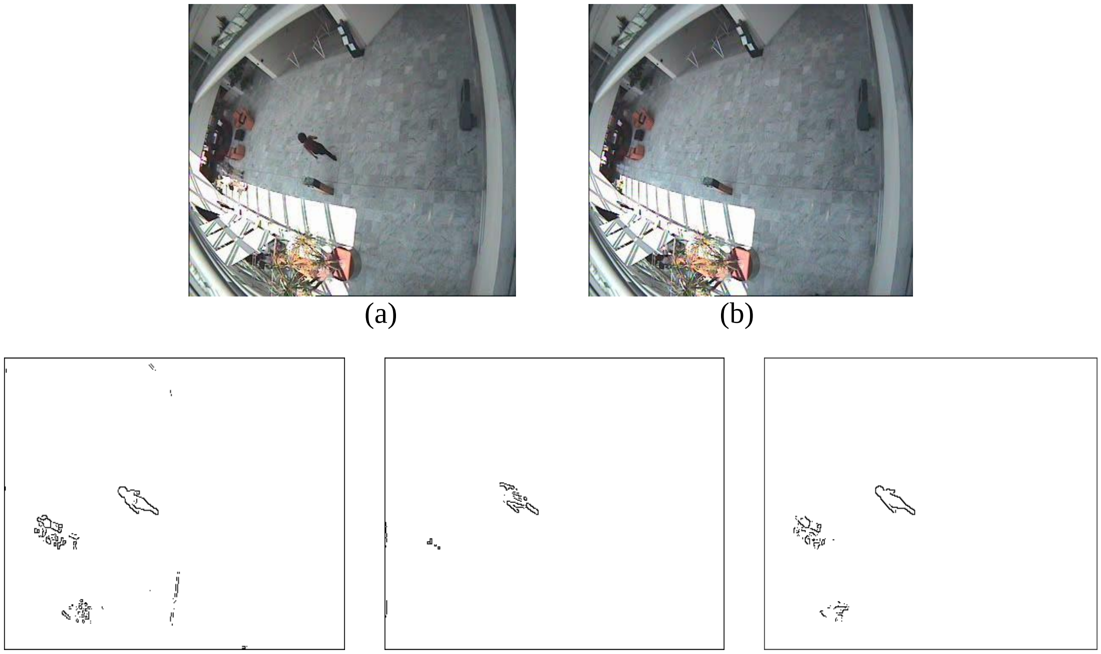


Figure 10. (a) the 267th Input Frame; (b) background Image; Detected Moving Object by (c) the Kim & Hwang Method; (d) the Dewan & Chae Method; (e) the Proposed Method

In addition, performance of the proposed method was compared with that of the existing methods in terms of error rate which is mainly used for qualitative evaluation and is calculated by using the following equation:

$$\text{Error Rate} = (FP + FN)/(TP + FP + FN) \times 100\%, \qquad \textbf{(10)}$$

where FP represents the number of no-change pixels detected incorrectly, FN represent the number of change pixels detected incorrectly, and TP represents the number of change pixels detected correctly. The calculation is performed on almost 25 frames by comparing the detection results against the ground truth images. These ground truth images were manually calculated in advance. Table 1 shows the comparison of mean processing time among the proposed and existing methods. It shows that the total times required for processing an image of size 640x520 using the Kim & Hwang method, the Dewan & Chae method and the proposed method are about 56ms, 74ms, and 64ms, respectively. Therefore, the proposed method can process about 15 frames per second which is relatively good for real-time detection. Table 2 shows the comparison of error rate among the proposed and existing methods. Experiments show that the proposed method provides better results than well-known existing methods in terms of error rate.



(a) (b)

(c) (d) (e)

Figure 11. (a) the 339th Input Frame; (b) Background Image; Detected Moving Object by (c) Kim and Hwang Method; (d) Dewan and Chae Method; (e) Proposed Method

Table 1. Comparison of Mean Processing Time (in milliseconds) among the Proposed and Existing Methods

Processing steps	Kim & Hwang method	Dewan & Chae method	Proposed method
Difference images calculation	5	5	×
Edge map generation	39	39	×
Performing OR operation	12	×	×
Gradient map & edge map generation	×	×	30
Gradient difference map calculation	×	×	4
Distance Transform (DT) image generation	×	11	×
Matching confidence and moving edge detection	×	19	×
Masking & threshold selection	×	×	26
Updating background	×	×	4
Total time	56	74	64

Table 2. Comparison of Error Rate (in percentage) among the Proposed and Existing Methods

Cases	Kim & Hwang method	Dewan & Chae method	Proposed method
Indoor images	2.74	5.37	1.78
Outdoor images	5.57	8.93	4.85
Foggy images	21.51	34.03	15.52
Sunny images	8.98	19.73	6.54

Overall, the proposed method outperforms well known existing methods under different illumination conditions, including indoor, outdoor, sunny, and foggy cases for detecting moving object in real time.

6. Conclusion

In this paper, real time moving object detection method based on the improved edge localization and gradient directional masking was presented. Experimental results indicate that the proposed method provides better results than well-known edge based methods. This is because it works on most recent successive frames and utilizes edge localization for detecting moving object. In addition, it is robust against different illumination changes. Moreover, the proposed method detects objects more accurately from the input images than existing methods. These results demonstrate that the proposed method can be a suitable candidate for moving object detection in real time video surveillance system.

Acknowledgements

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MEST) (No. 2012-0004962) and the Network-based Automation Research Center (NARC) at the University of Ulsan.

References

[1] A. D. Sappa and F. Dornaika, "An Edge-Based Approach to Motion Detection", V. N. Alexandrov, G. D. van Albada, P. M. A. Sloot and J. J. Dongarra, (eds.) LNCS, vol. 3991, **(2006)**, pp. 563–570. Springer, Heidelberg.

[2] Y. Chen, S. Yu, W. Sun and H. Li, "Objects Detecting Based on Adaptive Background Models and Multiple Cues", ISECS International Colloquium on Computing, Communication, Control, and Management, vol. 1, Issue 3-4, **(2008)**, pp. 285 – 289.

[3] J. Tao and C. Yu, "Real- Time Detection and Tracking of Moving Object", International Symposium on Intelligent Information Technology Application, **(2008)**.

[4] G. Li Y. Wang and W. Shu, "Real-Time Moving Object Detection for Video Monitoring Systems", International Symposium on Intelligent Information Technology Application, **(2008)**.

[5] L. Maddalena and A. Petrosino, "Moving Object Detection for Real-Time Applications", International Conference on Image Analysis and Processing, **(2007)**.

[6] E. Rivlin, M. Rudzsky, R. Goldenberg, U. Bogomolov and S. Lepchev, "A Real-Time System for Classification of Moving Objects", International Conference on Pattern Recognition, **(2002)**.

[7] M. Murshed, A. Ramirez, O. Chae, "Statistical Background Modeling: An Edge Segment Based Moving Object Detection Approach", Proc. of IEEE International Conf. on Advanced Video and Signal Based Surveillance, **(2010)**, pp. 300-305.

[8] Md. Z. Islam, C.M. Oh and C. W. Lee, "Video Based Moving Object Tracking by Particle Filter", International Journal of Signal Processing, Image Processing and Pattern, vol. 2, no. 1, **(2009)**, pp. 119–132.

[9] P. Dunne and B. J. Matuszewski, "Histogram Based Detection of Moving Objects for Tracker Initialization in Surveillance Video", International Journal of Grid and Distributed Computing, vol. 4, no. 3, **(2011)**, pp. 71-78.
[10] N. R. Shin and K. Hong, "A Study on the Development of Multiple Objects Extraction Method Using Difference Image Edge Information", International Journal of Energy, Information and Comm., vol. 2, no. 2, **(2011)**, pp. 108–120.
[11] M. J. Hossain, M. A. A. Dewan and O. Chae, "Moving Object Detection for Real Time Video Surveillance: An Edge Based Approach", IEICE Transaction on Communication, vol. E90-B, no. 12, **(2007)**, pp. 3654–3664.
[12] N. Otsu, "A Threshold Selection Method from Gray-level Histograms", IEEE Trans. Syst., Man, Cybern., vol. SMC-9, no. 1, **(1979)** January, pp. 62–66.
[13] T. W. Ridler and S. Calvard, "Picture Thresholding Using an Iterative Selection Method", IEEE Trans. Syst., Man, Cybern., vol. SMC-8, no. 8, **(1978)** August, pp. 630–632.
[14] G. W. Zack, "Automatic Measurement of Sister Chromatid Exchange Frequency", J. Histochem. Cytochem., vol. 25, no. 7, **(1977)**, pp. 741–753.
[15] Y. Kita, "Change Detection Using Joint Intensity Histogram", in Proc. Int. Conf. Pattern Recog., Hong Kong, **(2006)**, pp. 351–356.
[16] D. Sen and S. K. Pal, "Histogram Thresholding Using Fuzzing and Rough Measures of Associated Error", IEEE Trans. Image Process., vol. 18, no. 4, **(2009)** April, pp. 879–888.
[17] C. Kim and J. -N. Hwang, "Fast and Automatic Video Object Segmentation and Tracking for Content-based Applications", IEEE Transactions on Circuits and Methods for Video Technology, vol. 12, no. 2, **(2002)**.
[18] M. A. A. Dewan, M. J. Hossain and O. Chae, "Reference Independent Moving Object Detection: An Edge Segment Based Approach", International Conference on Knowledge-Based Intelligent Information and Engineering Method, part I, vol. LNAI 4692, **(2007)**, pp. 501-509.
[19] P. L. Rosin and T. Ellis, "Image Difference Threshold Strategies and Shadow Detection", in Proc. Brit. Mach. Vis. Conf., Birmingham, U.K., **(1995)**, pp. 347–356.
[20] J. N. Kapur, P. K. Sahoo and A. K. C. Wong, "A New Method for Graylevel Picture Thresholding Using the Entropy of the Histogram", Comput. Vis. Graph. Image Process., vol. 29, no. 3, **(1985)**, pp. 273–285.
[21] S. B. Gray, "Local Properties of Binary Images in Two Dimensions", IEEE Trans. Comput., vol. C-20, no. 5, **(1971)** May, pp. 551–561.
[22] L. O'Gorman, "Binarization and Multi-thresholding of Document Images Using Connectivity", in Proc. Symp. Document Anal. Inf. Retrieval, Las Vegas, NV, **(1994)**, pp. 237–252.
[23] L. Bruzzone and D. F. Prieto, "Automatic Analysis of the Difference Image for Unsupervised Change Detection", IEEE Trans. Geosci. Remote Sens., vol. 38, no. 3, **(2000)** May, pp. 1171–1182.
[24] L. Wang and N. H. C. Yung, "Extraction of Moving Objects from Their Background Based on Multiple Adaptive Thresholds and Boundary Evaluation", IEEE Transactions on Intelligent Transportation Methods, vol. 11, no. 1, **(2010)**, pp. 40-51.
[25] J. Zhou and J. Hoang, "Real Time Robust Human Detection and Tracking System", in Proc. IEEE Comput. Vis. Pattern Recog. Workshop, San Diego, CA, **(2005)**, pp. 149.
[26] H. Kim, R. Sakamoto, I. Kitahara, T. Toriyama and K. Kogure, "Robust Foreground Extraction Technique Using Background Subtraction with Multiple Thresholds", Opt. Eng., vol. 46, no. 9, **(2007)**, pp. 097 004-1–097 004-12.
[27] D. J. Dailey, F. W. Cathey and S. Pumrin, "An Algorithm to Estimate Mean Traffic Speed using Uncalibrated cameras", IEEE Transactions on Intelligent Transportation Methods, vol. 1, no. 2, **(2000)**, pp. 98–107.

Authors

Pranab Kumar Dhar received the Bachelor of Science (B.Sc.) degree in Computer Science and Engineering from Chittagong University of Engineering and Technology (CUET), Chittagong, Bangladesh in 2004. He received Master of Science (M.Sc.) degree from the School of Computer Engineering and Information Technology at University of Ulsan, South Korea in 2010. In 2005, he joined as a Lecturer in the Department of Computer Science and Engineering at Chittagong University of Engineering and Technology (CUET), Chittagong, Bangladesh where he is currently serving as an Assistant Professor. His research interest includes Digital Signal Processing, Multimedia Security, Digital Watermarking, Sound Synthesis, and Digital Image Processing.

Mohammad Ibrahim Khan received the Bachelor of Science (B.Sc.) degree in Electrical and Electronic Engineering from Bangladesh University of Engineering and Technology (BUET), Dhaka, Bangladesh in 1999. He received Master of Science (M.Sc.) degree in Computer Science and Engineering from the same University in 2002. He received his Ph.D degree in Computer Science and Engineering from Jahangirnagar University in 2010. In 1999, he joined as a Lecturer in the Department of Computer Science and Engineering at Chittagong University of Engineering and Technology (CUET), Chittagong, Bangladesh where he is currently serving as an Associate Professor. His research interest includes Digital Image Processing, Graph Theory, Cryptography, Digital Watermarking, Multimedia Systems, and Digital Signal Processing.

Ashoke Kumar Sen Gupta received the Bachelor of Science (B.Sc.) degree in Electrical and Electronic Engineering from Chittagong University of Engineering and Technology (CUET), Chittagong, Bangladesh in 1998. He received Master of Science (M.Sc.) degree in Electric Power Engineering from Chalmers University of Technology, Goteborg, Sweden in 2007. In 1999, he joined as a Lecturer in the Department of Electrical and Electronic Engineering at Chittagong University of Engineering and Technology (CUET), Chittagong, Bangladesh where he is currently serving as an Associate Professor. His research interests include Wind Energy, Power System Analysis, Digital Signal Processing, Smart Grid, Electrical Machines, and High Voltage Technology.

D.M.H. Hasan received the Bachelor of Science (B.Sc.) degree in Computer Science and Engineering from Chittagong University of Engineering and Technology (CUET), Chittagong, Bangladesh in 2011. His research interest includes Digital Image Processing, Computer Vision, and, Multimedia Systems.

Jong-Myon Kim received the Bachelor of Science (B.Sc.) degree in Electrical Engineering from Myongji University, South Korea in 1995. He received Master of Science (M.Sc.) degree from the School of Electrical and Computer Engineering at University of Florida, Gainesville, FL, USA in 2000. He received his Ph.D. degree from the School of Electrical and Computer Engineering at Georgia Tech, Atlanta, GA in 2005. Currently he is a faculty member in the School of Computer Engineering and Information Technology at University of Ulsan, South

Korea. His research interest includes Computer Architectures, SOC Design, Digital Watermarking, Sound Synthesis, Multimedia Systems, and Digital Signal Processing. Prior to joining academia, he was a processor architect in both multi-format CODEC chip design and prototyping team and reconfigurable processor team at Samsung Advanced Institute of Technology (SAIT), South Korea. He is a member of different professional research societies, like IEEE, KIPS, KIISE, KSCI, and EASCO.