International Journal of Engineering Research and Technology.
ISSN 0974-3154 Volume 11, Number 1 (2018), pp. 29-39

http://www.irphouse.com

An Efficient Moving Human Detection Algorithm for Intelligent CCTV Systems

Nuwan Sanjeewa and Won Ho Kim*

Division of Electrical, Electronic & Control Engineering,
Kongju National University, Republic of Korea

**Corresponding Author*

Abstract

This paper presents an efficient moving human detection algorithm based on simplified background subtraction, and human-body decision with 1D template correlation for intelligent CCTV system. In general, moving human detection algorithm consisted of object segmentation and human decision parts. In the proposed algorithm, firstly simplified background subtraction is performed to get coarse moving object regions by using simplified background modeling and updating processing. Then noise elimination using optimized adaptive threshold is done to get refined object regions. At the final step, it distinguishes the human body objects and non-human body objects by using one-dimensional template correlation. The proposed algorithm is simulated with outdoor test videos, and simulation results show the proposed algorithm is suitable for real-time video surveillance system by over 98% of correct detection rate.

Keywords - Image signal processing, Moving human detection, Computer vision, Video surveillance system.

I. INTRODUCTION

Analyzing human activities from CCTV images is a very important task in computer vision in the past few years. There are various application areas analyzing human activities of videos, and the most famous and important field is video surveillance systems [1]-[13]. It needs an autonomous method to extract human bodies from a video sequence before understanding the various human activities. Then applying a further process, the activities of a human can be understood.

Human detection techniques from a video sequence can be divided in to mainly four categories. They are human detection based on frame difference method, background subtraction method, optical flow method and infrared (IR) image based detection. The frame difference method and background subtraction method can be called as non-direct method since they need a further process to detect human bodies, while optical flow and infrared method are direct method. Frame difference method is very fast and easy for implementation since it has low complexity. But it cannot be used for object tracking and shape detection in direct [5]-[8]. In method of optical flow, it is reliable than other methods since it uses trained data and well defined vector sets for detection. Also it is capable of detecting and recognizing objects as well as measuring moving speed and directions. But the big problem is the computational load of the algorithms that used to generate feature vectors [3].

IR image based object detection method is capable of object tracking and object's shape recognition. Also this method does not need to apply further step in order to detect objects since it gives the only heated object's detail as well as the shapes of objects. However this method cannot be used for object recognition since it gives only intensity values of object and there is no color information. Also it cannot detect those objects less than required minimum temperature. The performance of this method is mainly depends on the IR sensor [4].

Background subtraction method [2], [9] is useful for human-body detection, since it gives details of object's shape. Also it is possible to track human objects as well as counting human objects. However it is required low complexity and reliable method for background modelling and updating as well as for human recognition in order to make this method success in real time detection. In addition, it is required adaptive threshold method for noise elimination since this method is very sensitive to noise such as Gaussian-noise, sudden illumination, light intensity changes. As a result, human recognition, background modeling and updating tasks should have low complexity to get efficient human detection algorithm.

This paper proposes efficient moving human detection algorithm included fast human body classification, and low complexity background subtraction method which is verified in previous work. In the next section, we describe in detail the proposed human detection algorithm consisted of simplified background modeling/updating method, optimized noise elimination with adaptive threshold, and human-body decision based on one-dimensional template correlation. Finally, we present simulation results to confirm practicality of the proposed algorithm.

II. MOVING HUMAN DETECTION ALGORITHM

The functional diagram of the proposed efficient moving human detection algorithm is shown in figure 1 and it consists of three processing steps. Firstly background subtraction is performed to get coarse moving regions by using proposed simplified background modeling and updating processing. Then noise elimination using optimized adaptive threshold is done to get clear candidate object regions. At the final step, it distinguishes the human-body and non- human body by one-dimensional template correlation. Also, it uses gray-level value of image to reduce the complexity to get efficient human detection algorithm. Each steps of the proposed algorithm are described in the followings.

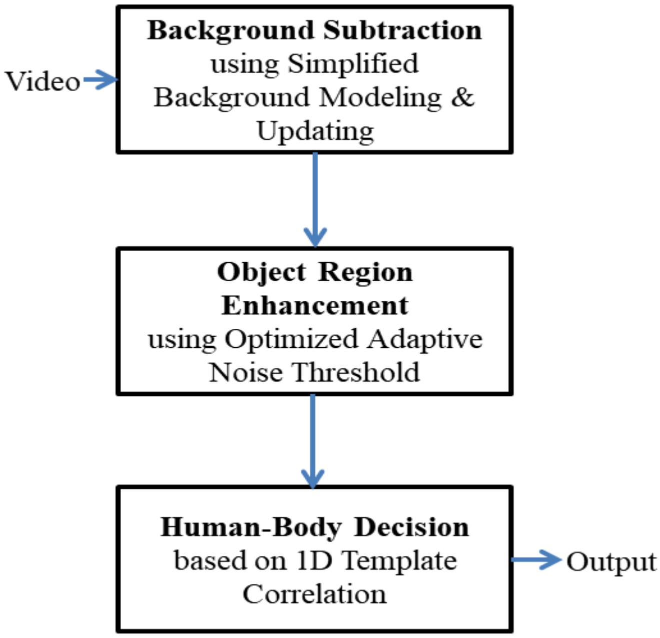

Figure 1. Block diagram of proposed algorithm

III. BACKGROUND SUBTRACTION

Background subtraction is the most common method for foreground detection in fields of image processing and computer vision. Generally interested regions such as human, car, etc. are considered to be as foreground while static objects are considered as background such as buildings, trees, etc. Therefore background subtraction is a widely used technique for detecting moving objects in videos. Furthermore the background subtraction is done by differencing of the current image and the background image.

In order to get a reliable background, the background image has to be updated in real time according to the situation in the surveillance field. This paper uses simplified Gaussian model for background updating based on probability of a pixel sequence over certain time [1]. It is based on pixel based processing, and detailed steps of the background modeling and updating is as bellows.

STEP1: Quantize the gray level of each image those pixel values varying 0-255 into 26 linear steps as shown in figure 2. The system noise is assumed as white Gaussian noise and any pixel's value can be changed by 0~10 due to the noise. Therefore the tolerance of a single bin is set to be 10.

STEP2: Start to fill those bins by putting current frame's pixel values to the proper bins and repeat filling until n^{th} frame. After processing of n^{th} frame, replace the most weighted bin's average value (μ_{ω}) to pixel value in background image (B_{im}) as shown in figure 2.

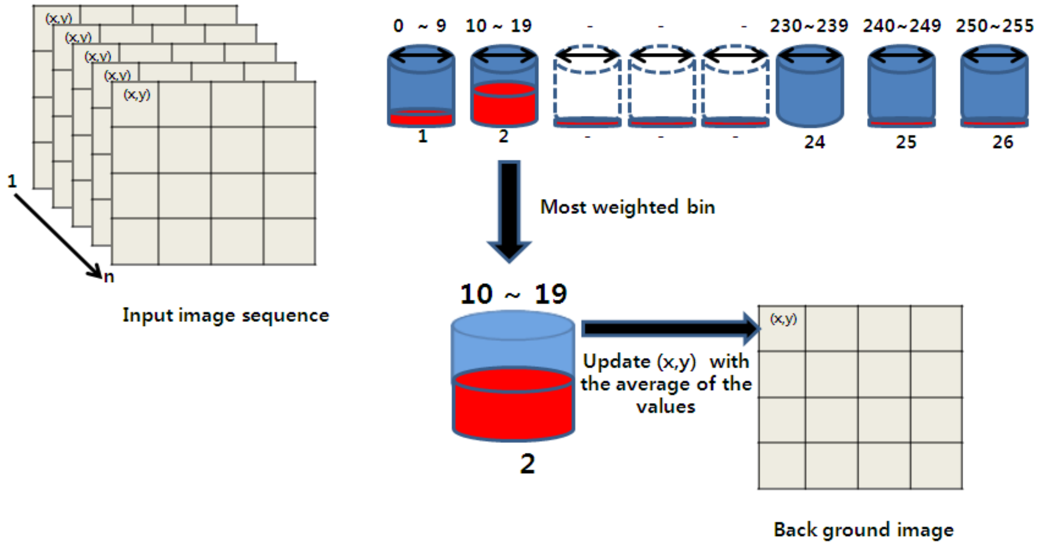

Figure 2. Diagram of background modeling and updating processing

The operations described above are done for each pixels in the input image sequence over a certain period and finally background can be obtained by taking the average of most weighted bin, since the values in the most weighted bin are the values those most repeated values during the certain period. Therefore, the average value of those most repeated values is updated as background pixel's value (B_{im}) at the regarding coordinates.

$$B_{im}(x, y) = \mu_{\omega}(x, y) \quad (1)$$

After background image is generated, background subtraction processing is done to get the moving object areas. The resulted image has not only real moving regions, but also the noise regions due to system noise, shadow, texture etc. Therefore, further processing is required, and next step is to enhance object regions as removing noise pixels.

IV. OBJECT REGION ENHANCEMENT

Most of the conventional methods used fixed threshold to remove noises of difference image obtained by background subtraction. But it can be deleted considerable number of pixels of real objects in the difference image obtained from various surveillance environments. Therefore, it needs adaptive threshold approach to remove only noise pixels related the various video surveillance environments and systems. The popular adaptive threshold approach uses the standard deviation and mean of the background subtracted difference image.

This paper uses method which is fully adaptive to the real-time situations in the surveillance fields and there is no any parameter that should be set by user. In this method, all the parameters and constant will be calculated automatically and the used adaptive threshold is calculated as a combination of mean value of difference image μ[k] and number of zero valued pixels (N_{ZP}) [1].

$$T_p[k] = \frac{N}{N_{zp}[k]} \times \frac{\mu[k]}{(N - N_{zp}[k])} \qquad (2)$$

The μ[k] is mean value and N_{ZP} is number of zero pixels in the difference image which is counted by scanning the difference image. As using this adaptive threshold, the noises are effectively eliminated depending on variable field environments of system. In this step, the moving regions are segmented by threshold processing of the background difference image (D_{im}). The pixel values of the difference image higher than the adaptive threshold are consider to be moving pixels and the lower values than threshold are considered as noise pixels.

$$M_{region} = \begin{cases} 1, & if\ D_{im} \geq T_p \\ 0, & else \end{cases} \qquad (3)$$

According to the equation 3, the binary difference image including moving regions M_{region} is generated. The binary moving object image has usually a number of closely spaced scattered small regions. It needs further processing for segmenting those disorder pixels in order to obtain the final moving objects. Generally, the eight-connected component labeling [14] is applied to the resulted image after processing of background subtraction and adaptive noise elimination. The labels those have small number of pixels is removed by considering as scatters and noises.

V. HUMAN BODY DECISION BASED ON 1D TEMPLATE CORRELATION

There are template correlation methods are available for human body detection, but all those methods are based on 2D image correlation. However, this paper proposes simple one-dimensional correlation method which transforms 2D image into 1D data, and it is used to calculate correlation with pre-defined 1D template. Different point with the conventional algorithms, the proposed method used only one reference template.

It is very difficult task to detect moving human object since the movements of human are varying in a large range. It causes body shape to be more complex over time. First of all, it has to find unique features of human's body shape when they move. There are many shapes respect to the movements. But this paper consider only about some unique features when a human moves in position of front-up, right-up, left-up, back-up. When human moves in all those positions, the upper part of body still remain almost same shape and this shape can be used to classify human body. Based on the above concept, this paper proposes a simple template correlation method by applying 1D template correlation. The steps of the moving human decision are described as bellows.

STEP1: Extract one third of rectangle including human body region. It contains the upper part of the human body such as shoulder and head, since the human is standing in any direction.

STEP2: Resize the extract part in to 10 by 10 block in order to match with the reference template, since the proposed method used fixed 10 data points. This block is 2D image that's size of 10×10.

STEP3: Get the sum of each column pixels. This step is to make the 1D candidate template that has 10 data points.

STEP4: Calculate normalized cross correlation (NCC) to measure the difference between reference template and candidate template. If the correlation value is larger than 0.8, the candidate objet is classified as moving human object. This threshold value is selected by experimental method.

$$r_{xy} = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2 \sum_{i=1}^{n}(y_i-\bar{y})^2}} \qquad (4)$$

r_{xy} is the normalized cross correlation value and x is the input candidate signal and y is reference signal while $\bar{x}, \bar{y}$ represent respectively mean value of the input signal and reference signal.

In detail, if correlation value is 0.9034, it matches over 80% of decision threshold. Therefore the candidate object is classified as a human body. Therefore the proposed 1D correlation method is able to distinguish moving human body from other objects and it also reduced the computational load comparatively to the 2D correlation method.

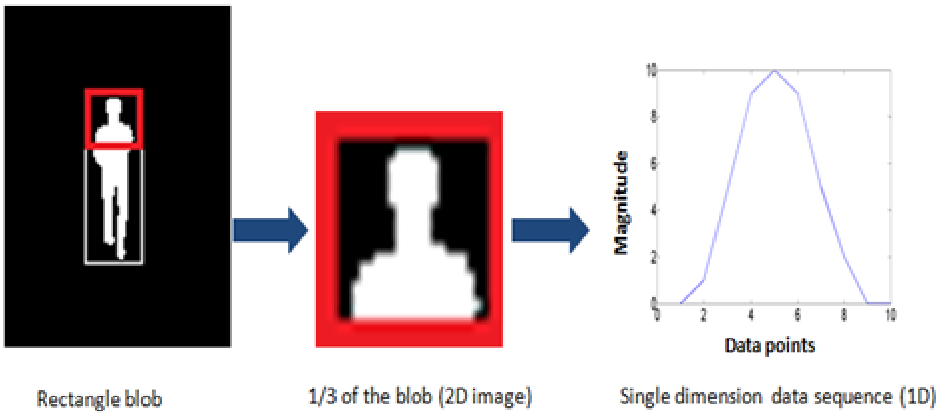

Figure 3. Calculation of 1D candidate template from 2D human body's rectangle image

In addition, the proposed 1D correlation method (blue line) gives higher correlation values than 2D correlation (red line) for same candidate blocks as shown in figure 4. Because there could be many data points that are miss-matched in 2D image and it cause correlation coefficient to have negative values. Those negative values drive NCC value to be small. Therefore the proposed 1D template correlation method gives better result than the 2D correlation method and it is much more suitable for human object detection and the complexity of the algorithm is reduced since it used only one reference template.

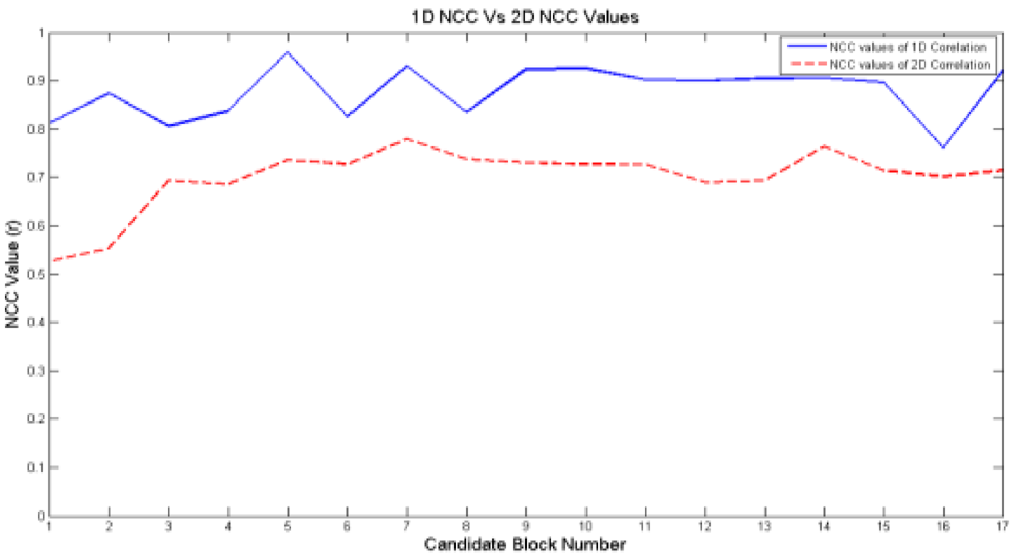

Figure 4. NCC comparison of 1D template method and conventional 2D template method

VI. SIMULATION RESULTS

The simulation has been done by a computer and 720 × 480 test videos are used. The output images of the simulation are re-formatted in RGB-type images and the detection results in outdoor surveillance video at various places are shown in figure 5. Moving human objects are indicated by green rectangle and other moving objects are indicated by red rectangle.

Figure 5. Simulation result images for some test videos

Table 1 shows the summary of detection performance for outdoor video sequences. The simulation results show 98.4% of average correct detection rate while average false detection rate is bounded to 0.28%. The average miss detection rate is 1.61%, and miss detection occurred when the background's color and the object's color is similar. It is occurred due to very small difference values obtained by the background subtraction step, and they are eliminated as noise pixels.

Table 1. Summary of simulation results

Video	Decision points (Total)	Frame of human	True detection	Missed detection	False detection	True rate	Missed rate	False rate
Out 1	400	350	345	5	0	98.57%	1.43%	0.00%
Out 2	250	230	230	0	0	100.00%	0.00%	0.00%
Out 3	100	100	96	4	4	96.00%	4.00%	4.00%
Out 4	300	150	148	2	0	98.67%	1.33%	0.00%
Out 5	300	200	198	2	0	99.00%	1.00%	0.00%
Out 6	300	250	247	3	0	98.80%	1.20%	0.00%
Out 7	200	198	198	0	0	100.00%	0.00%	0.00%
Out 8	300	260	250	10	0	96.15%	3.85%	0.00%
Out 9	200	200	195	5	0	97.50%	2.50%	0.00%
Out 10	250	198	198	0	0	100.00%	0.00%	0.00%
Out 11	300	200	196	4	6	98.00%	2.00%	2.00%
Out 12	450	410	398	12	0	97.07%	2.93%	0.00%
Out 13	275	170	170	0	0	100.00%	0.00%	0.00%
Average	3625	2916	2869	47	10	98.39%	1.61%	0.28%

VII. CONCLUSIONS

This paper proposed a method to detect moving human based on simple background subtraction, object region enhancement and human recognition based on 1D template correlation. First it does background subtraction to extract moving object and then adaptive threshold method to remove noise from the background subtraction image. Once the moving objects are extracted after removing noise using adaptive threshold, the algorithm distinguish the moving human bodies from other objects based on proposed 1D template correlation method. The proposed 1D correlation method gives comparatively better correlation between reference template and candidate templates than conventional 2D correlation method. Also the proposed method has low computational load then 2D correlation method. The simulation result shows average correct detection rate of 98.4% for moving human object of outdoor videos. The Simulation result shows the proposed method is able to perform well in real-time intelligent CCTV video application comparatively to the conventional methods.

ACKNOWLEDGEMENTS

This work was supported by the research grant of the Kongju National University. (Project No. 2015-0735-01)

REFERENCES

[1] Nuwan Sanjeewa and W.H. Kim, "Low complexity background subtraction and adaptive noise elimination for moving object segmentation", International Journal of Applied Engineering Research, Volume 12, Number 4, pp. 509-512, 2017.

[2] Sen, Das, Chowdhury, "Saving electrical power in a surveillance environment", Seventh IEEE International Conference on Advances in Pattern Recognition, pp.274-277, 2009.

[3] Sidenbladh, Hedvig, "Detecting human motion with support vector machines", IEEE International Conference on Pattern Recognition, Volume 2, pp.188-191, 2004.

[4] Jianchao Zeng, Sayedelahl, Chouikha, Gilmore, Frazier, "Human detection in non-urban environment using infrared images", IEEE International Conference on Information, Communications & Signal Processing, pp.1-4, 2007.

[5] Zhan Chaohui, Duan Xiaohui, Xu Shuoyu, Song Zheng, Luo Min, "An improved moving object detection algorithm based on frame difference and edge detection", IEEE Conference on Image and Graphics, pp.519-523, 2007.

[6] Lian Xiaofeng, Zhang Tao, Liu Zaiwen, "A novel method on moving-object detection based on background subtraction and three frame differencing", International Conference on Measuring Technology and Mechatronics Automation, Volume 1, pp.252-256, 2010.

[7] Wei Li, Kaipeng Yang, Jiaxin Chen, Qingtao Wu, Ming Chuan Zhang, "A fast moving object detection method via local neighborhood similarity", International Conference on Mechatronics and Automation, pp.4500-4505, 2009.

[8] Cucchiara, Grana, Piccardi, Prati, "Detecting moving objects, ghosts, and shadows in video streams", IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume 28, pp.1337-1342, 2013.

[9] Heikkila, Pietikainen, "A texture-based method for modeling the background and detecting moving objects", IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume 28, pp.657-662, 2006.

[10] Murshed, Ramirez, Chae, "Statistical background modeling: an edge segment based moving object detection approach", Seventh IEEE International Conference on Advanced Video and Signal Based Surveillance, pp.300-306, 2010.

[11] Zheng Yi, Fan Liangzhong, "Moving object detection based on running average background and temporal difference", International Conference on Intelligent Systems and Knowledge Engineering, pp.270-272, 2010.

[12] Jabid, Mohammad, Ahsan, Abdullah-Al-Wadud, "An edge-texture based

moving object detection for video content based application", 14th International Conference on Computer and Information Technology, pp. 112-116, 2011.

[13] Bo-Hao Chen, Shih-Chia Huang, "Advanced moving object detection algorithm for automatic traffic monitoring in real-world limited bandwidth networks", IEEE Transactions on Multimedia, Vol.16, pp. 837-847, 2014.

[14] Rafael C. Gonzalez and Richard E. Woods, Digital Image Processing. Prentice Hall, 2nd edition, USA, 2002.