Proceedings of the 17th World Congress
The International Federation of Automatic Control
Seoul, Korea, July 6-11, 2008

Robust Salient Moving Object Detection with Light-Computational Load

Jeong Hwan Choi * Young Min Baek * Jin Young Choi *

** Electrical Engineering Department, Automation and System Research Institute(ASRI)*
Seoul National University, Seoul, Korea
(e-mail: jhchoi,ymbaek@neuro.snu.ac.kr, jychoi@snu.ac.kr)

Abstract: In this paper, we propose a method which detects salient moving objects with light-computational load. Though Gaussian Mixture Model is widely used for object detection, it is computationally heavy. On the other hand, basic methods like temporal difference are simple and fast but they have constraints as hole or ghost problems. We have combined these algorithms to overcome each one's weakness. We use background modeling and subtraction method which are similar to adaptive threshold with foreground map. Foreground map is generated by Modified Temporal Difference to speed up the process. Using adaptive threshold, we have improved the performance, when there is slightly moving background like branches in the wind. So we can eliminate meaningless objects with light-computational load. Experimental results show efficiency and robustness of our algorithm in several outdoor scenes.

1. INTRODUCTION

Object Detection is the first step in surveillance system. Performance of object detection has a major effect on classification, behavioral understanding, and other parts of surveillance system. If shapes of detected objects have distortion, it is difficult to expect a good classification result. In order to extract object's exact shapes, several morphological operations are performed after detection. However, problem is that morphological operations could also append distortions to images.

Many methods have been proposed for moving object detection. There are three approaches that are most widely used in surveillance systems; temporal difference method, background subtraction and an optical flow.

Temporal difference is a very simple and fast method that is suitable for dynamic environments. This method finds moving pixels by comparing a pre-defined threshold value to difference values calculated by subtraction of the previous and current gray-scale images. It is not only computationally light but also very effective for finding moving pixels. However, this measurement with respect to fixed threshold value is not suitable for noisy environments. If threshold is increased for reducing noise, moving pixels could also be reduced. Temporal difference scheme also suffers from 'Hole' and 'Ghost' problem because of overlapping of moving objects between previous and current images. To solve this problem, Cheng-Chi Chang et al. (2005) proposed Modified Temporal Difference method which is very effective to eliminate Hole and Ghost effects. By keeping the foreground map of previously detected objects, Hole and Ghost problems are solved. However, other problems (e.g. Trail effects) could occur if first image has objects. So, this algorithm shows good results only in indoor scenes, which is not noisy and has no swaying background.

Background subtraction scheme maintains the background model using some statistical modeling and detects objects in motion that do not fit this already-known model. This scheme is generally sensitive to changes due to lighting and events. Events contain sudden changes like parking car which was not present before, and remains there for a long period of time. For this reason, adaptive background subtraction techniques are used that provide a solution to the problem of a continually changing background environment. Chris Stauffer et al. (1999) have proposed the object detection algorithm which is based on a mixture of Gaussians models to store several background scenes. In their method, the representative background is modeled to k-Gaussian Distributions with mean and variance respectively. Gaussian Mixture Model (GMM) shows robust results in various environments even if there are noisy backgrounds. However, GMM has weakness to sudden light change and shadow removal. To solve this problem, Ying-Li Tian et al. (2005) used Gaussian mixture model along with multiple cues like intensity and texture information. Although background modeling method has a good merit which adapts to background change, it needs many Gaussian distributions to get an effective background model. Therefore, it takes more computational time compared with GMM.

Optical flow finds moving pixels by calculating motion vector which has similar color information between previous and current images. Normally, moving pixel has a distinct motion vector compared to surrounding pixels. Although this method has good performance in small camera movement, it is difficult to get real-time performance because of computational complexity.

Our algorithm is based on simple background modeling which is similar to adaptive threshold in gray-scale. Each pixel has two parameters; mean of intensity and variance. Background pixels are modeled only when foreground map indicates if those pixels are unchanged. Foreground map is generated by Modified Temporal Difference (Cheng-Chi

Chang et al. 2005). By combining two algorithms, we can choose background pixels without additional operations.

This paper is organized as follows. Section 2 explains our proposed method, and section 3 shows the results of our experiments. Finally, we bring to conclusion and mention future work in section 4.

2. PROPOSED METHOD

Our method consists of two parts; 1) Modified Temporal Difference with Stationary Pixel Removal to find background MASK, 2) Background Modeling with adaptive threshold. Figure 1 illustrates a flowchart of our method.

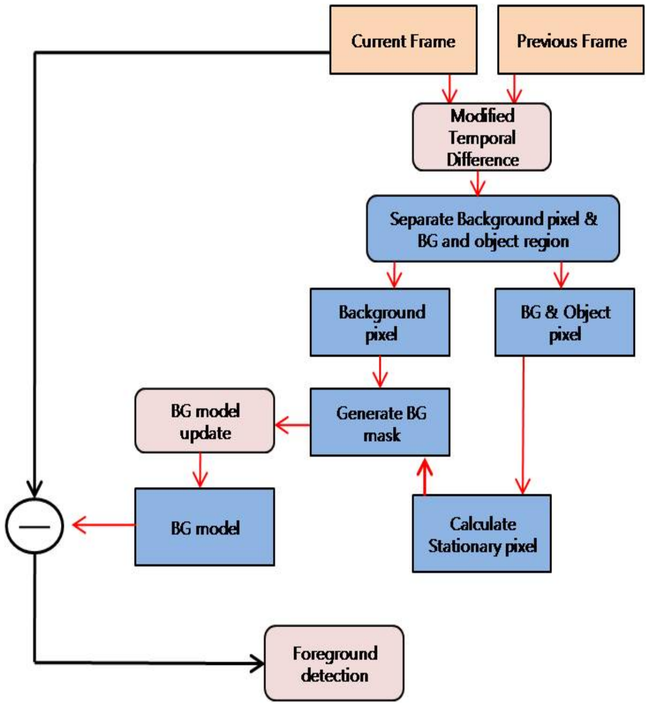


Fig. 1. A flowchart of our detection method

Modified Temporal Difference (MTD): In traditional temporal difference method, it is very difficult to find object shape because of Hole and Ghost problems. So, we store previous object regions and use this information to find filled-in moving objects. This method is very effective for detection of full-object region. However, trail effect could occur. We combined background modeling method to solve this problem. MTD is used only for generating certain background mask in our method.

Stationary Pixel Removal: A disadvantage of MTD is that objects are detected as foregrounds continuously if they were detected once before. To solve this problem, we regard objects, which are fixed on the same place for a certain period of time, as parts of the background.

Background Modeling(BGM): After finding certain background regions using MTD, we update the background image using foreground map. Updated background image is used to decide if detected moving object through MTD is foreground or background.

2.1 Modified Temporal Difference (MTD)

Generally, absolute difference pixel value of overlapping part which is generated by difference operation of two consecutive images is very small. If we define threshold to find changing pixels without considering these small values, the entire overlapping part pixels are rejected. Also, if we define threshold which considers overlapping part pixel values, then noises are also detected with these pixels. To solve this problem, MTD used previously detected object information. Detail of Modified Temporal Difference process is as follows,

1) Get a 'P-image' using subtraction of two consecutive gray scale images.
2) Get a 'R-image' using subtraction of 'Template image' and 'P-image'.
3) Store the 'R-image'.
4) At next image sequence, recursively use 'R-image' as 'Template image'

$$P_{(t)} = \bigcup_{g \in I_s} p_s(g) = \cup\left[i_s(g) - i_{s-1}(g)\right] \quad (1)$$

$$g \in \begin{Bmatrix} nonchangepixel, & if\ -\varepsilon \le i_s(g) - i_{s-1}(g) \le \varepsilon \\ positive-changepixel, & if\ \ \varepsilon \le i_s(g) - i_{s-1}(g) \\ negative-changepixel, & if\ \ i_s(g) - i_{s-1}(g) < -\varepsilon \end{Bmatrix} \quad (2)$$

Where P(t) is partial change image which occurs as intensity changes at two consecutive images, ε is a predefined threshold value.

$$R_{(s)} = \bigcup_{g \in M_{s-1_s}} r_s(g) = \bigcup_{g \in M_{s-1}} \left[m_{s-1}(g) - p_s(g)\right] \quad (3)$$

$$g \in \begin{Bmatrix} change-object, & \left|m_{s-1}(g) - p_s(g)\right| > \varepsilon \\ background, & \left|m_{s-1}(g) - p_s(g)\right| \le \varepsilon \end{Bmatrix} \quad (4)$$

M is Template image and all pixel values of first template image are zero. R is result of detected object, and is used next time as template image. Figure 2 shows a diagram of Modified Temporal Difference.

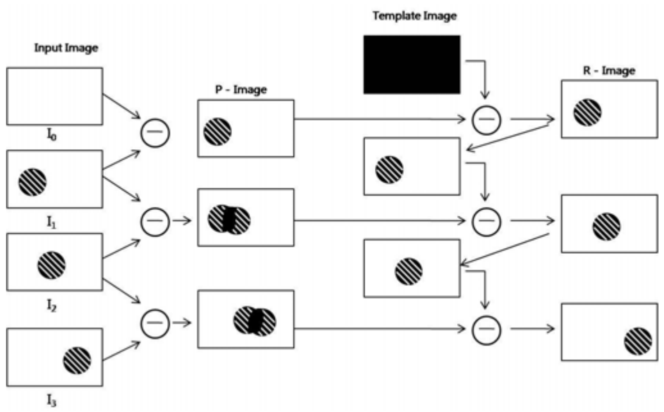


Fig. 2. Modified Temporal Difference to find foreground map.

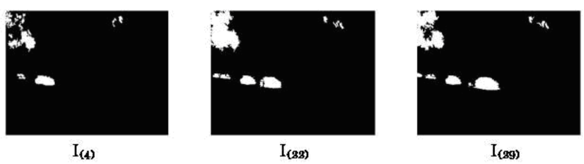
2.1.1 Stationary Pixel Removal

Modified temporal difference has restriction that the first image should not have any object. If the first image has moving object, there are constantly remaining trails. To solve this problem, we used Stationary Pixel Removal. Although this method recognizes previously moving objects during initial period, but after a certain period of time, it regards object which appear in the first image as part of the background.

Stationary Pixel Remove process is as follows,

1) Find pixels which change intensity in threshold.
2) Increase the pixel cumulative numbers.
3) If cumulated value over the threshold, regard it as background.

(a) Modified Temporal Difference with Stationary Pixel Remove



(b) Modified Temporal Difference

Fig. 3. Detection result comparing MTD with Stationary Pixel Removal and only MTD.

Figure 3 is a comparison of results between our method and MTD. $I_{(4)}$ has almost same detection result, but as time goes by, object which appears in the first image gradually disappears at $I_{(22)}$ and $I_{(29)}$. Stationary Pixel Equation is defined below.

$$S_{t+1}(x) = \begin{cases} S_t(x)+1 & ;\min(x) < I(x) < \max(x) \\ S_t(x) & ;otherwise \end{cases} \quad (5)$$

$$if \;\; S_t(x) > S_{\max} \quad \longrightarrow \quad remove \quad (6)$$

Where $I(x)$ is intensity value of pixel, S_{max} is threshold for number of cumulated stationary pixels. For more fast adaptation to an abnormal condition (e.g tremble and moving camera), we adjust S_{max} flexibly. If a number of detected objects increase unusually, S_{max} is adjusted to be small. As a result, abnormal detected objects which occur in the background can be removed quickly, and then certain moving objects are found.

2.2 Adaptive Background Modeling

Background model is generated for extracting distinguished motion. We simplified modeling method using the intensity value of background pixel. Variance has been adjusted according to differences of previous and current images' intensity values.

$$\mu_t = (1-\alpha)\,\mu_{t-1} + \alpha I_t \quad (7)$$
$$\sigma^2_t = (1-\alpha)\,\sigma^2_{t-1} + \alpha(I_t - \mu_t)^2 \quad (8)$$

Where α is learning rate. μ is the mean and σ^2 is the variance. Background mask plays an important role for modeling. We need to model the regions which are surely background pixels. For fast and effective background modeling, we use background mask through MTD with Stationary Pixel Remove. We only update pixels which are decided as changing pixels in background by the background mask. This method could reduce computational complexity comparing with method which updates whole pixels. In addition, a possibility that foreground is incorrectly modelled could be decreased because MTD finds larger foreground map than actual foreground regions. Certain background pixels can be modelled using both MTD with Stationary Pixel Remove and Adaptive Background Modeling.

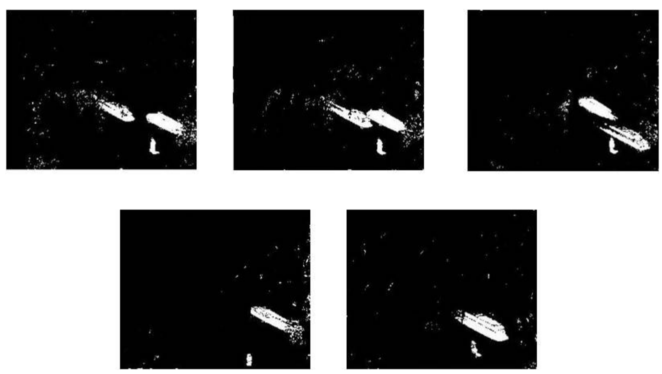
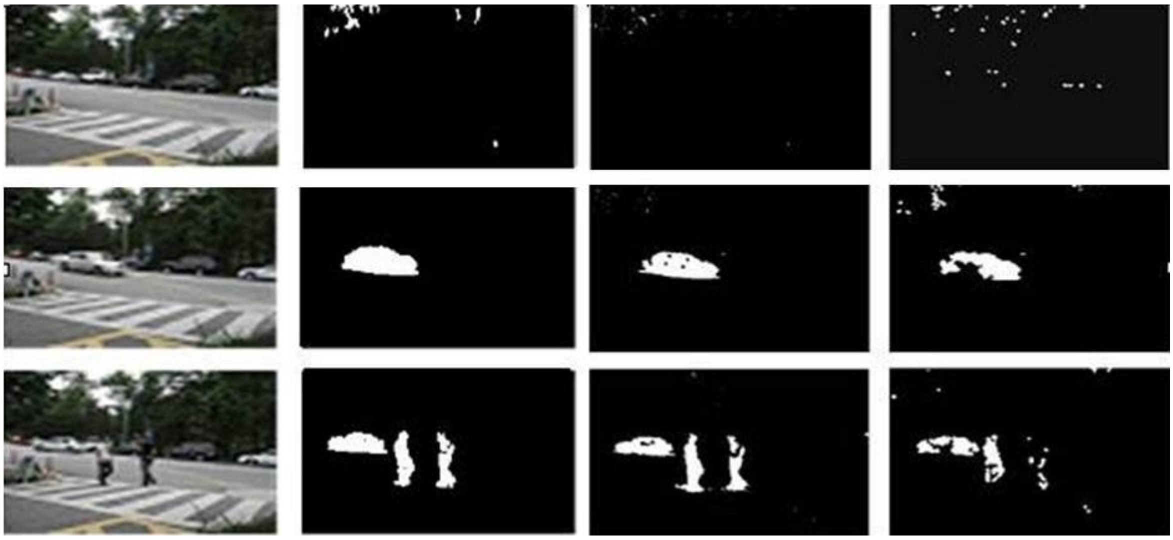
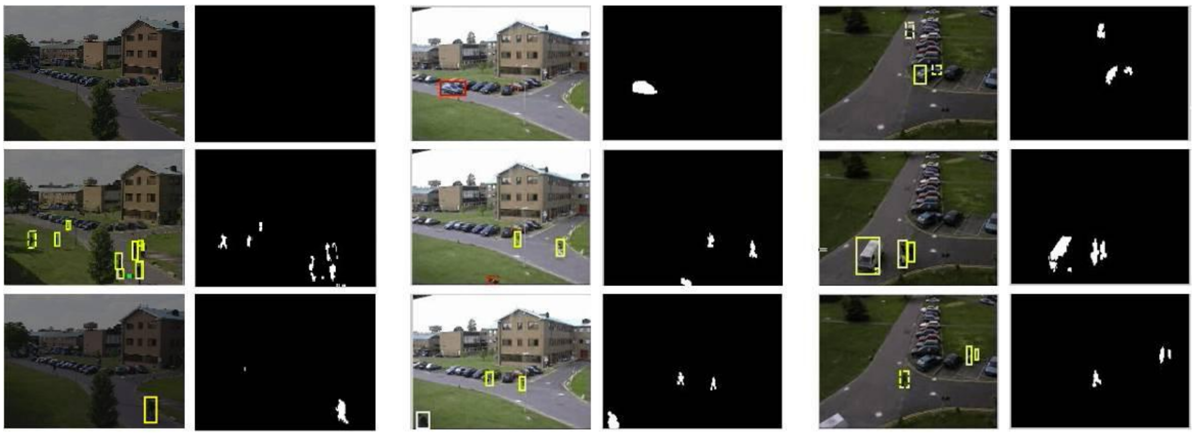


Fig. 4. Influence of Trail effect. As influence of previous image, there occurs 'Trail', which makes it difficult for recognition.



(a) (b) (c) (d)

Fig. 5. Result of moving object detection in our video resource (a) Input image (b) Proposed algorithm (c) Background modeling with GMM(RGB color space) (d) 3-frame temporal difference method



(a) (b) (c)

Fig. 6. Result of moving object detection in outdoor scenes. All videos are used from PETS database.

3. EXPERIMENTAL RESULTS

To verify proposed algorithm, we had experiments in two ways. First, we tested our algorithm in our video resources which are recorded in outdoor scenes. Second, we tested surveillance test movies from PETS. We compared the processing time, detection performance and robustness to noise. Compared algorithms are 3-frames temporal difference and Background modeling with 3-Gaussian mixture model in RGB color space. Our experiment was performed on 2GHz Pentium IV with 320 x 240 video resolution.

Table 1. Computational Time

3-frame difference	BGM with RGB space	Proposed method
25frame/sec	15frame/sec	19frame/sec

As can be seen in table 1, our algorithm is slower than 3-frame temporal difference, and is faster than method of using GMM. However, Figure 5 shows the performance of proposed algorithm comparing with other methods using our video sources. 3-frame difference detects too many noises including swaying branches in the wind, and also does not represent detected object shape exactly, whereas our algorithm and GMM method certainly reduce noise than 3-frames, and also represent object shape well. Even though the results show similar performance with GMM, computational load of our method is less than GMM method.

Figure 6 shows outdoor detection result of using PETS. Proposed method detects even small objects. In spite of showing outdoor environment, it almost does not detect noises. Especially, result 6-(a), testing same video in Ying-Li Tian et al. (2005) experiment, shows robustness of our algorithm against quick lighting changes.

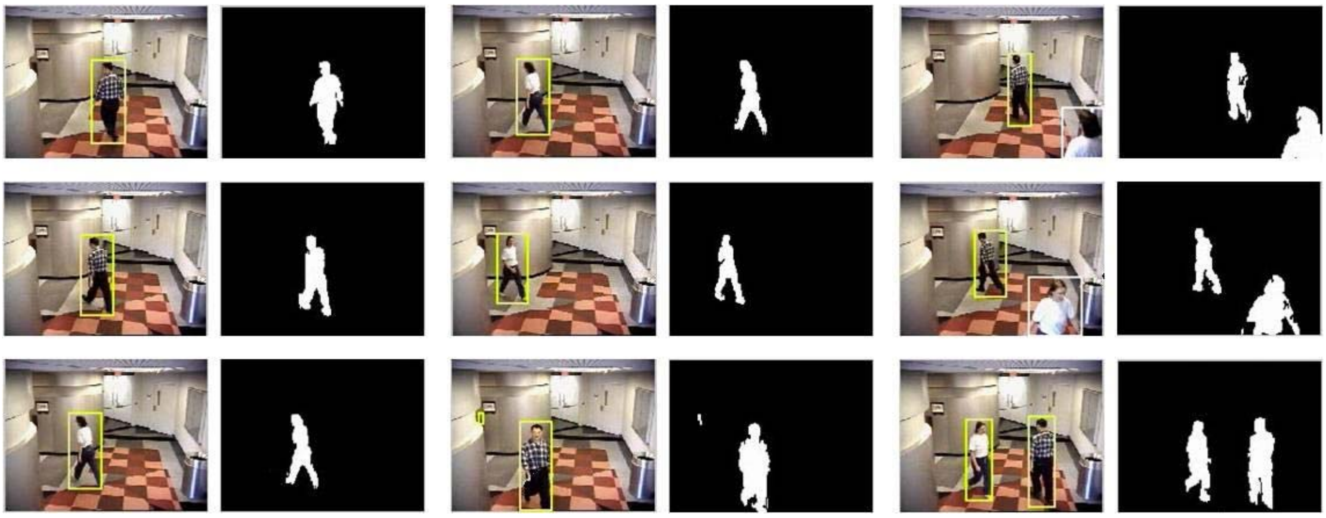
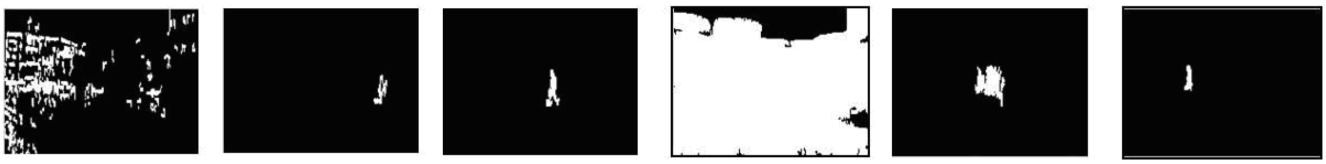
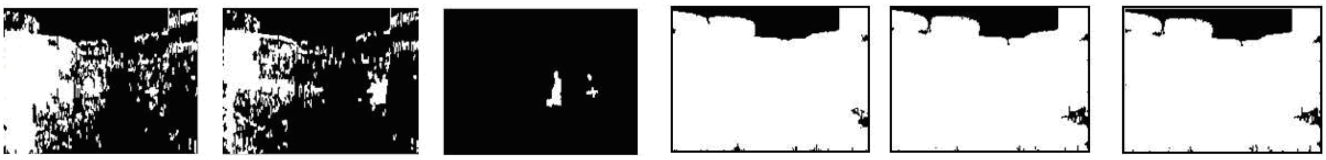


Fig. 7. Result of moving object detection in indoor scenes.



t+1 t+2 t+5 t+7 t+8 t+9

(a) Using adaptive parameter according to camera moving.



t+1 t+2 t+5 t+7 t+8 t+9

(b) Using constant parameter.

Fig. 8. Result of using variable parameter. 't+1 ~ t+2' is initial background modeling time and't+7' is time to start moving of camera

Figure 7 shows detection result in the indoor scene. This video focused on the problem of shadow effects which appear under moving humans. In our experiment, shadows almost do not occur on the floor. However, the problem of reflection on the wall is still remaining. Reflection problem is different from shadow effect because shadow is generated by intensity changes while reflection depends on color information. Color changes causes large variation to intensity, therefore it is difficult to solve by modeling method in gray-scale images.

Figure 8 shows adaptation result of abnormal condition. As can be seen in the results (t+2) and (t+7), the variable parameters find foreground and modeling background faster than fixed parameters in early stages and after the camera stops moving. Especially, after camera starts moving, for a little while many noises are detected, but our method removes the noises quickly, which occur in the background. Due to this function our method maintains continuous detection ability.

4. CONCLUSION

We have proposed an effective detection algorithm for surveillance system. For fast and accurate object detection, we used the Modified Temporal Difference with Stationary Pixel Remove and adaptive background modeling. Also for effective adaptation to abnormal condition, we used variable parameter. We can summarize our algorithm advantages as follows;

1) Simple and fast object detection
2) Robust to change environment condition
3) Express original object's shape with less loss
4) Fast adaptation in abnormal condition.

As future work, we will research how to remove 'Hole' which occurs in patches at slow moving objects. We also are going to control the reflection problem in indoor surveillance system.

ACKNOWLEDGEMENT

This research was supported by the Ministry of Commerce, Industry and Energy of Korea and Samsung Techwin. Co.

REFERENCES

Cheng-Chi Chang, Tsorng-Lin Chia, Chen-Kuei Yang (2005). Modified temporal difference method for change detection. Optical engineerin . Volume 44. page 28 – 31

Chris Stauffer, W.E.L Grimson (1999) Adaptive background mixture models for real-time tracking. Conference on Computer Vision and Pattern Recognition. Volume 2. Page - 252.

Chritopher P. Diehl, John B. Hampshire Ⅱ. (2002) Real-time Object Classification and Novelty Detection for Collaborative Video Surveillance. International Joint Conference on Neural Networks. Volume 3. Page 2620 – 2625.

Davide A. Migliore, Matteo Matteucci, Matteo Naccari (2006). A Revaluation of Frame Difference in Fast and Robust Motion Detection. The 4th ACM International workshop on Video surveillance and sensor networks. page 215 – 218.

Stefan Huwer, Heinrich Niemann (2000). Adaptive Change Detection for Real-Time Surveillance Application. International Workshop on Visual Surveillance(VS'2000). page 37 – 46

Ismail Haritaoglu, David Harwood, Larry S. Davis (2000) W4 : Real – Time Surveillance of People and Their Activities . Pattern analysis and machine intelligence. Volume 22, no 9. Page 809 - 830.

Masayiki Yokiki Yokoyama, Tomaso Poggio (2005) Contour - Based Moving Object Detection and Tracking. 2005. 2nd Joint IEEE International Workshop on Visual surveillance and Performance Evaluation of Tracking and Surveillance. Page 271 – 276

Mohamed F. Addelkader, Rama Chellappa, Qinfen Zheng (2006) Integrated Motion Detection and Tracking for Visual Surveillance. International Conference on Computer Vision System. Page 28 – 34.

Paul L.Rosin, Tim Ellis (1995) Image difference threshold strategies and shadow detection. British conference on Machine vision. Volume 1. Page 347 – 356.

R. T. Collins (2000). A system for video surveillance and monitoring: VSAM final report. Technical report. CMU-RI-TR-00-12

Ying-Li Tian, Max Lu, Arun Hampapur (2005). Robust and Efficient Foreground Analysis for Real-time Video Surveillance. CVPR'05. Volume 1. page 1182 – 1187.

Zoran Zivkovid (2004). Improved Adaptive Gaussian Mixture Model for Background Subtraction. International Conference of Pattern Recognition, Volume 2. page 28 – 31.