# A Novel Background Updating Algorithm Based On the Logical Relationship

NING LI DE XU BING LI
Institute of Computer Science and Engineering
Beijing Jiaotong University
Beijing, 100044
CHINA.

*Abstract:* Moving object detection and segmentation is a fundamental technology for video applications. Background subtraction is a basic component step for this task: given a frame in the video sequence, a moving object can be detected by spotting the parts of the frame that do not fit the background model. Therefore, a high performance background modeling and background updating algorithm is crucial for background subtraction. We work with a fixed camera in a stationary scene and a simple non-statistical background model is utilized, therefore we formulate the problem of background subtraction as the problem of background updating. This paper proposes a novel and efficient background updating algorithm based on the logic AND and XOR relationship of the change information in difference frames.

*Key-Words:* background subtraction (BS), background modeling, background updating (BU), dynamic thresholding, logic relationship.

## 1 Introduction

Automatic moving object detection and segmentation is the fundamental task for many video applications. There are of sorts approaches developed to deal with this task. The [1] gives an extensive investigate on moving object detection and segmentation algorithms. Roughly, these algorithms can be divided into 4 categories: Edge-feature based algorithm [2], Spatial-Temporal based algorithm [3], Motion feature based algorithm [4] and Change Detection based [5] [6] algorithm. The [1] claims that the Change Detection technology is more efficient in terms of the computation-cost requirement for real-time tracking.

The [7] classifies the change detection algorithm into temporal difference (TD) and background subtraction (BS). TD is very easy to implement, but it is not so effective to get the whole region of moving objects, especially the inner part of them. BS is commonly used mainly due to its simplicity and low computational intensity. BS can be defined as follows: given a frame in the video sequence, a moving object can be detected by spotting the parts of the frame that do not fit the background model. Therefore, a reasonable and effective background modeling and background updating algorithm is crucial for BS. The [8] gives a systematic discussion on background modeling methods. In our experiment, we use videos shot by fixed color cameras in stationary scene and constant illumination for simplicity. This assumption only requires a non-statistical background model. Therefore the BS is simply achieved by the subtraction of the current frame from a predefined background frame that should be updated periodically.

In this paper, we formulate the problem of BS as the problem of composing a new background frame using current background frame and patches extracted from video sequence. We firstly use Otsu's dynamic thresholding method to binarize the difference image BS and TD to locate the change information with respect to the background frame, and then propose a novel and efficient BU algorithm based on the logic AND and XOR operation on the change information in the difference frames.

The rest of the paper is organized as follows: Section 2 describes the proposed background updating algorithm. Section 3 presents the experiment results on two video sequences. Conclusive remarks are addressed at the end of this paper.

## 2 The proposed background updating algorithm

For a static scene, the simplest background model could be just an image of the scene without the intruding objects. To achieve the BS task, the background model should be updated periodically. Therefore, one basic and controversial question arises: what kind of objects should be deemed a part of a background frame. The [9] states "background will not be updated in regions of pixels belonging to stop objects in the scene." This statement can be implemented in such a way that when a still object moving away from its static position should be excluded from the background frame; on the other hand, a moving object becoming standstill should be included as a part of the background frame. We name the objects having this property as ghosts. And the BU task in our paper is to detect and rectify ghosts.

To achieve this goal, we develop a novel BU algorithm. The working flow of the algorithm is illustrated in Fig 1. We use 3 consecutive frames in a video record for each BU process. Inspired by [1], both TD mask and BS mask are combined as the primary step to locate the ghost region. Otsu's threshoding method is used to binarize the difference images (TD and BS), which are inputted the subsequent logical operation stage to determine if one background should be updated or not.

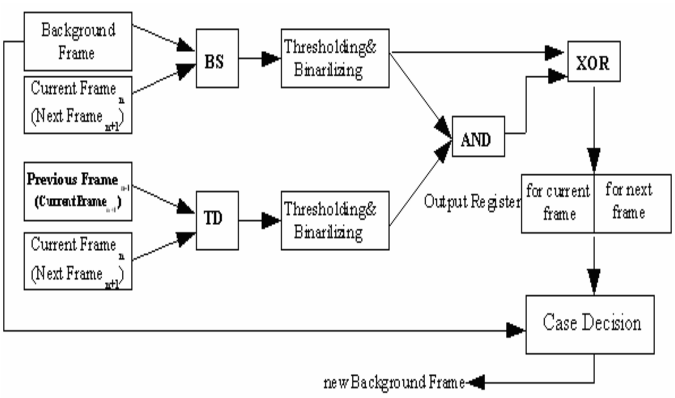


Fig.1 Flow Chart for Background Updating Strategy.

### 2.1 Change detection

In our experiment, we use frame subtraction to detect change. The color difference between two consecutive frames (TD) and between one frame and a background frame (BS) is denoted as "Change". We measure the change in HSV color space. The change can be induced either by ordinary moving of objects relative to background or ghosts. Therefore, the changing region unnecessarily means the region where an ordinary object motion occurs, it may be the ghosts region where the BU process proceeds.

Fig. 2 shows the ideal cases where changes occur for simplicity. In case 1, there is only ordinary moving object (full black mask) exist therefore no updating required; in case 2, the half black mask is a ghost which represents a still object added on the background; in case 3, the ghost is removed from background.

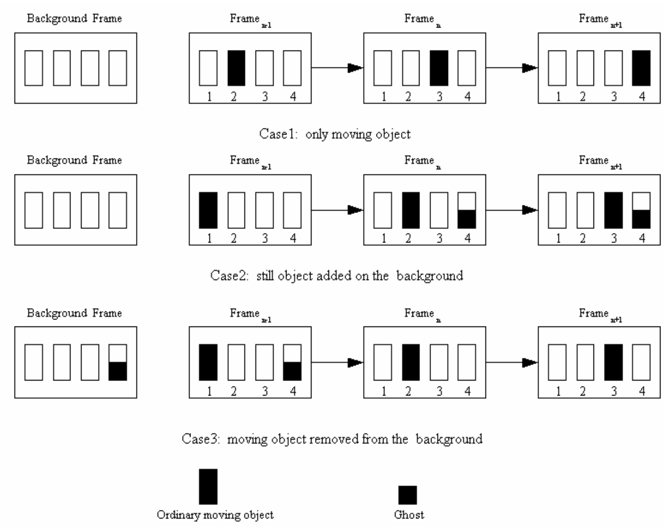


Fig.2 Ideal case of Change region for simplicity.

### 2.2 Dynamic thresholding

Otsu's threshoding method [10] is used to locate the threshold for binarizing image. It is based on the idea: find the threshold that minimizes the weighted within-class variance $\sigma_w^2$. Otsu's method works with a bimodal histogram. In motion detection application, pixels of the difference frame are classified into changed region and unchanged region, therefore the histogram of the difference frame is bimodal.

In our experiment, we operate this method on the HSV histogram so that it can deal with more complex scene than the operation on gray histogram. We denote the normalized histogram as P($l$), where $l$=1: L and L is the total number of HSV bins. A HSV level $t$ needs to be found to divide the histogram into two classes: $l$ = 1: $t$ and $l$ = ($t$ + 1): L so that $\sigma_w^2$ is minimum.

The class probabilities $q_1(t)$ and $q_2(t)$ are estimated as:

$$q_1(t) = \sum_{l=1}^{t} P(l) \quad (1)$$

$$q_2(t) = \sum_{l=t+1}^{L} P(l) \quad (2)$$

the class means, $\mu_1(t)$ and $\mu_2(t)$ are given by:

$$\mu_1(t) = \sum_{l=1}^{t} \frac{lP(l)}{q_1(t)} \quad (3)$$

$$\mu_2(t) = \sum_{l=t+1}^{L} \frac{lP(l)}{q_2(t)} \quad (4)$$

and the class variances are calculated as:

$$\sigma_1^2(t) = \sum_{l=1}^{t} [l - \mu_1(t)]^2 \frac{P(l)}{q_1(t)} \quad (5)$$

$$\sigma_2^2(t) = \sum_{l=t+1}^{L} [l - \mu_2(t)]^2 \frac{P(l)}{q_2(t)} \quad (6)$$

Finally, the $\sigma_w^2$ is calculated by:

$$\sigma_w^2(t) = q_1(t)\sigma_1^2(t) + q_2(t)\sigma_2^2(t) \quad (7)$$

Theoretically, we need calculate the $t$ from 0 to L and find the threshold $t$ resulting in the minimum $\sigma_w^2$. Fig. 3 shows the experimental result using Otsu's method to get the binary image.

(a) 1st frame (b)43rd frame (c) result image

Fig.3 Example for Otsu's thresholding method, (c) is the binary image of the difference image of (a) and (b).

## 2.3 Logical relationship operation

The 0 and 1 region in the binarized TD and BS mask mean the unchanged and changed regions respectively. We can use the logic relationship of these regions to determine the type of the change: motion induced or ghosts induced. If it is the later case, a BU stage will follow.

Each BU process consists of 3 consecutive frames (Framen-1, Framen, Framen+1) in a video record. According to Fig. 2 we use the logic operator AND and XOR for each case to determine if the background frame should be updated. The "1" region in BS image indicates ordinary moving object and ghost. The logic "1" region in MO image indicates the ordinary moving object with respect to the background. Therefore, after the XOR operation, the OUTPUT image extracts the ghost region, see Table 1.

**Case 1: only moving object**
$BS_n$ = ***bin***($Frame_n$ - Background Frame) = 0010
$TD_n$ = ***bin***($Frame_n$ - $Frame_{n-1}$) = 0110
$MO_n$ = $BS_n$ ***AND*** $TD_n$ = 0010
$OUTPUT_n$ = $MO_n$ ***XOR*** $BS_n$ = 0000;

$BS_{n+1}$ = ***bin***($Frame_{n+1}$ - Background Frame) = 0001
$TD_{n+1}$ = ***bin***($Frame_{n+1}$ - $Frame_n$ ) = 0011
$MO_{n+1}$ = $BS_{n+1}$ ***AND*** $TD_{n+1}$ = 0001
$OUTPUT_{n+1}$ = $MO_{n+1}$ ***XOR*** $BS_{n+1}$ = 0000;

**Case 2: still object added on the background**
$BS_n$ = ***bin***($Frame_n$ - Background Frame) = 0101
$TD_n$ = ***bin***($Frame_n$ - $Frame_{n-1}$) = 1101
$MO_n$ = $BS_n$ ***AND*** $TD_n$ = 0101
$OUTPUT_n$ = $MO_n$ ***XOR*** $BS_n$ = 0000;

$BS_{n+1}$ = ***bin***($Frame_{n+1}$ - Background Frame) = 0011
$TD_{n+1}$ = ***bin***($Frame_{n+1}$ - $Frame_n$ ) = 0110
$MO_{n+1}$ = $BS_{n+1}$ ***AND*** $TD_{n+1}$ = 0010
$OUTPUT_{n+1}$ = $MO_{n+1}$ ***XOR*** $BS_{n+1}$ = 0001;

Case 3: moving object removed
from the background

Table 1. Logic operation illustration for Fig.2

Case 2 and Case 3 are essentially the same each other, because we utilize "CHANGE" as the measure to determine when BU occurs. A still object moving away from its original position, or a moving object transforming to the standstill state in the scene all means "CHANGE".

## 2.4 Background updating

See the case decision stage of Figure 1, $OUTPUT_n$ and $OUTPUT_{n+1}$ for Framen and Framen+1 are used to make decision on BU. If the ratio of "1" region in $OUTPUT_n$ to that in $OUTPUT_{n+1}$ is smaller than a threshold, the current background frame should be covered with the region in $Frame_n$ where its corresponding coordinates in $OUTPUT_{n+1}$ is logic "1", see Table 2. The threshold value is set to 0.90 experientially.

```
If   OUTPUT_n / OUTPUT_n+1 < threshold
    Background Frame (i,j) = Current Frame n (i,j)
    where OUTPUTn+1 (i,j) EQ 1;
    % (i,j) is the coordinate of a pixel in a frame
Else
    background frame keeps;
End
```

Table 2. Background updating decision

## 3. Experiments

The test videos we select are the "Hall Monitor" in MPEG 4 test sequence and "Fire Extinguisher" shot in our laboratory. Both videos are collected under the condition of fixed color camera and illumination. In our experiment, we sample every 5 frames and use 3 consecutive sampled frames for each BU process.

See Fig.4, in the frame sequence (Hall Monitor) (b)-(d), only ordinary moving object exist, therefore (a) should not be updated (Case 1); in the sequence (e)-(g), the suitcase appears as the ghost and should be added on the background (Case 2), see (h).

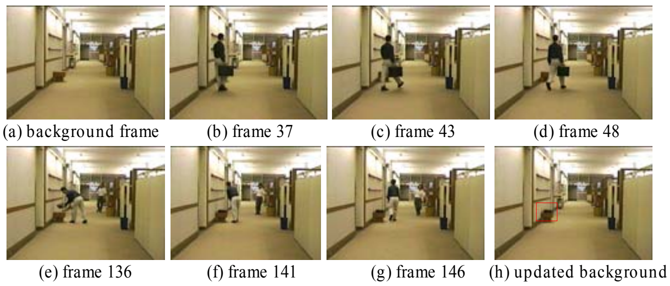


Fig.4 Example for algorithm processing, (h) is the updated background from frame (e)-(g).

Fig.5 shows the corresponding BS, TD, MO and OUTPUT image for the Case 2 in Fig4. We denote (e), (f) and (g) in Fig.4 as Framen-1, Framen and Framen+1 respectively. We get the binary image based on Otsu's thresholding method with HSV components quantized into 3, 3 and 12 levels respectively.

As illustrated in Fig.5, we can not get the whole region of a moving object in TD (see (b) and (g)) due to the overlap of the moving object in a short duration. The inner part of the moving object are classified as background area and therefore black. The drawback of TD results in a poor OUTPUT. In (d) and (i) much of light region is in fact the inner part of the two moving people after the XOR operation. We use seed growth method to calculate the area of each continuous light region in BS and OUTPUT. If the ratio of the continuous light region in OUTPUT to the corresponding region in BS is smaller than a threshold, the light region in OUTPUT is eliminated. We set the threshold 0.9 experientially.

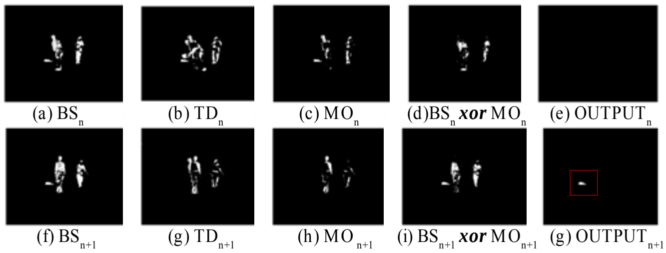


Fig.5 The first row is for $Frame_n$, (e) is the $OUTPUT_n$ after overlap region removal; the second row is for $Frame_{n+1}$, (g) is the refined $OUTPUT_{n+1}$.

To verify the universality of our approach, we use another video "Fire Extinguisher" to implement all BU cases. Fig.6 shows the sampled video sequence and the corresponding BU results: no BU required in terms of the video sequence (b)-(d); (h) is the updated background frame of (a) in terms of video sequence (e)-(g), where the fire extinguisher is excluded from the table; (l) is the updated background frame of (h) in terms of (i)-(k), where the fire extinguisher is added on the right side of the table.

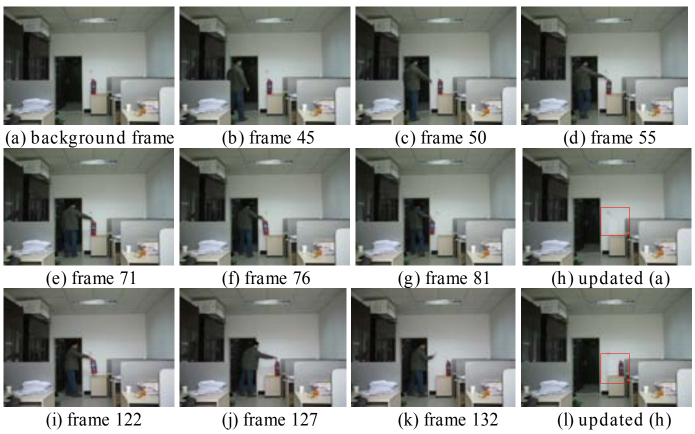


Fig.6 Example for algorithm processing, (h) is the background updating result from (e)-(g), (l) is the background updating result from (i)-(k).

## 4. Conclusion

Our approach makes use of the change content of both BS and TD. Their logic relationship is used to distinguish the ordinary moving object and ghost region with respect to a background frame. The method has lower computational intensity, and therefore applicable for BU. Another application would be irregular moving object detection. The future work would combine a statistical background model so that the algorithm is more robust for complex scenes.

*Reference:*

[1] Thou-Ho (Chao-Ho) Chen, Tsong-Yi Chen and Yung-Chuen Chiou, "An Efficient Real-Tim Video Object Segmentation Algorithm Based On Change Detection and Background Updating," IEEE Inter. Conf. on Image Processing, 2006, PP.1837 – 1840.

[2] C.Kim and J.N.Hwang, "Fast and Automatic Video Object Segmentation and Tacking for Content-Based Application," IEEE Trans. on Circuits and System for Video Technology, vol.12 (2), 2002, pp.122-129.

[3] Munchurl kim, Jae Gark Choi, Daehee Kim, Hyung Lee, Myoung Ho Lee, Chieteuk Ahn and Yo-Sung Ho, "A VOP Generation Tool: Automatic Segmentation of Moving Object in Image Sequences Based on Spatio-Temporal Information," IEEE Trans. on Circuits Syst. Video Technol, vol.9 (8), 2002, pp. 1216-126.

[4] YU Tian-li and ZHANG Yu-jin, "Video Retrieval Based on the Global Motion Information," ACTA ELECTRONICA SINICA, vol.29 (12A), 2001.

[5] R.T. Collins, A.J. Lipton, T. Kanade, H. Fujiyoshi D. Duggins, Y. Tsin,D. Tolliver, N. Enomoto and O. Hasegawa, "A System for Video Surveillance and Monitoring," VSAM Final Report, Carnegie Mellon University, Pittsburgh, 2000.

[6] C.Stauffer and W.E.L.Grimson, "Learning Patterns of Activity Using Real-Time Tracking," IEEE Trans. on Pattern Analysis and Machine Intelligence, 2000, pp. 747-757.

[7] Guo Jing, Chng Eng Siong and Deepu Rajan, "Fore ground Motion Detection by Difference Based Spatial Temporal Entropy Image," IEEE Region 10 Conference on TENCON, Volume A, vol. 1, 2004, pp.379 – 382.

[8] Massimo Piccardi, "Background subtraction techniques: a review," IEEE Inter. Conf. on Systems, Man and Cybernetics, vol.4, 2004, pp.3099 – 3104.

[9] Duarte Duque, Henrique Santos and Paulo Cortez, "Moving Object Detection Unaffected by Cast Shadows, Highlights and Ghosts," IEEE Inter. Conf. on Image Processing. ICIP, vol 3, 2005, pp. III – 413 -16.

[10] Guo Jing, Deepu Rajan and Chng Eng Siong, "Motion Detection With Adaptive Background and Dynamic Thresholds," Fifth Inter. Conf. on Information, Communications and Signal Processing, 2005, pp.41 – 45.