

# Change Detection based Real Time Video Object Segmentation

Mrs.Megha V.Gupta (P.G. Student), Dr.S.D.Sawarkar (Principal),

Department of Computer Engineering, Datta Meghe College of Engineering  
Airoli, Navi Mumbai-400708

## Abstract

*Segmentation of video foreground objects from background has many important applications, such as human computer interaction, video compression, multimedia content editing and manipulation. The key idea in our paper is to obtain the moving object region which can be set as the possibility foreground, and the other region set as background. An efficient video object segmentation algorithm is proposed based on change detection and background updating that can quickly extract the moving object from video sequence. The change detection is used to analyse temporal information between successive frames to obtain the change region. Then, the combination of frame difference mask and background subtraction mask is adopted to acquire the initial object mask and further solve the uncovered background problem and still object problem. Moreover, the boundary refinement is introduced to overcome the shadow influence and residual background problem. The advantage of change detection based approaches is the low computational load and system complexity enabling real-time applications.*

**Key Words:** frame difference, change detection, background, object, threshold, mask

## 1. Introduction.

Segmentation of digital video plays an important role for content based multimedia applications. The new video coding standard MPEG-4 for instance relies on the decomposition of image sequence frames into semantically meaningful video objects (VO) to provide content based functionalities. These content based functionalities allow separate coding and encoding of objects and permit the manipulation of the original scene by simple operations on the bit stream [1].

Moving object segmentation techniques can be separated into motion-based methods and spatio-temporal methods [2]. Motion based methods can be classified as 2D [3-7] and 3D techniques. 2D motion based methods have usually a relatively low computational complexity and facilitate simple implementation, but generally lack robustness. 3D methods have a relatively high computational load but are generally more robust. Change detection has been proposed to obtain temporal information for the extraction of Video Object Planes (VOPs) from image sequences [8-10]. Change detection based methods proposed so far have employed frame difference information of two successive frames (the current and the previous frame) only. An algorithm

based on change detection using a special relaxation algorithm which increases noise strength has been proposed in [8]. In this case a memory is used for change detection masks in order to improve temporal coherency of the resulting object masks. Neri et al. in [9] proposed an algorithm based on change detection which separates potential foreground regions employing a higher order statistics (HOS) significance test to inter-frame differences. A pixel based background registration technique, which uses a change detection mask, is proposed to obtain reliable background information in [10]. This technique compares each incoming frame with the background image to decide whether a pixel corresponds to the background or a foreground object, however time consuming morphological filters are used to construct the final complete object mask. Because only frame by frame differences are monitored in the methods proposed so far, a moving foreground is required for successful segmentation, failing at slow movements and temporary poses of objects.

The main idea is to identify the set of pixels that have undergone under some significant change between the last image of the sequence and the previous images. These groups of pixels create what is often known as the change mask. Detecting and

representing this change mask provides valuable information for the applications described before.

A change of each pixel is detected if a difference value exceeds pre-set thresholds. However, this method will yield effective results only if the signal to noise ratio is very high. In addition, the uncovered background, still object situation, light changing, and shadow effects will also degrade the detecting accuracy.

## 2. Change Detection Method.

A conventional approach for the detection of targets or changes in images is to pairwise subtract successive images in the sequence.[11] Change detection methods segments each frame into two regions namely changed and unchanged regions in case of a static camera or global and local motion regions in case of a moving camera. Spatio-temporal change detection deals with the former case where unchanged region corresponds to background and changed regions to the foreground object(s).[12] It aims at real-time processing and different motion in the foreground object does not need to be distinguished for the targeted application (that is scene is only separated into 2 classes of objects: foreground and background). There are various methods of video object segmentation, but the faster video object segmentation techniques are based on change detection(with or without preprocessing to cater for global motion) approach followed by further post processing.[13]

- a) **Pixel-based method:** Change detection can be performed by comparing successive frames. The simplest way to compute the dissimilarity between two frames is to compare corresponding pixels from two successive images. [14]
- b) **Block-based method:** Block sampling of the video frames can be performed in order to increase the quality of change detection but also to decrease the computation time. Use of blocks allows a processing which is intermediate, between local level like pixel-based methods and global level as histogram-based methods. Main advantage of block-based methods is their relative insensitivity to noise and camera or object motion. [14]

## 3. Algorithm Description.

Baseline mode is designed for stable situations. That is, the camera is still, and there is no light changing and no shadows. It is based on change detection and background registration technique.

Unlike other change detection algorithms, the change detection mask here is not only generated from the frame difference of current frame and previous frame but also from the frame difference between current frame and background frame, which can be produced by background registration technique. Since the background is stationary, it is well-behaved and more reliable than previous frame. Besides, still objects and uncovered background problems can be easily solved under this scheme. The block diagram of baseline mode is shown in Figure 1. There are five parts in baseline mode: Frame Difference, Background Registration, Background Difference, Object Detection, and Post processing.

### Step I - Frame/Background Difference

The differencing includes frame differencing and background differencing. In the frame difference, the frame difference between current frame and previous frame is calculated and thresholded. It can be presented as

$$D(x, y, t) = |I(x, y, t) - I(x, y, t-1)| \quad (1)$$

$$FDM(x, y, t) = 1 \text{ if } FD \geq Th$$

$$0 \text{ if } FD < Th \quad (2)$$

where  $I$  is frame data,  $FD$  is Frame Difference, and  $FDM$  is Frame Difference Mask. Pixels belonging to  $FDM$  are moving pixels. Note that there is a parameter  $Th$  needed to be set in advance.

### Step II - Background Registration

The goal of background registration step is to construct a reliable background information from the video sequence. According to  $FDM$ , pixels not moving for a long time are considered as reliable background pixels. The procedure of Background Registration can be shown as

$$SI(x, y, t) = SI(x, y, t-1) + 1, \text{ if } FDM = 0$$

$$0, \text{ if } FDM = 1 \quad (3)$$

$$BG(x, y, t) = I(x, y, t), \text{ if } SI(x, y, t) = Fth$$

$$BG(x, y, t-1), \text{ else} \quad (4)$$

$$BI(x, y, t) = 1, \text{ if } SI(x, y, t) = Fth$$

$$BI(x, y, t-1), \text{ else} \quad (5)$$

where  $SI$  is Stationary Index,  $BI$  is Background Indicator, and  $BG$  is the background information. The initial values of  $BI$ ,  $BG$ , and  $BI$  are all set to "0."

Stationary Index records the possibility if a pixel is in background region. If is high, the possibility is high; otherwise, it is low. If a pixel is "not moving" for many consecutive frames, the possibility should be high. When the possibility is high enough, the current pixel information of the

position is registered into the background buffer BG which is shown as (4). Besides, *Background Indicator* is used to indicate whether the background information of current position exists or not, which is shown as (5). Note that (3)–(5) also imply that a background updating ability is also included in *Background Registration*, that is, if background changes, new background information will be updated in to the background buffer. The parameter *Fth* needed to set in advance, which indicates the number of consecutive frames pixel is not moving.

### STEP III- Background difference mask

After Background Difference, another change detection mask named Background Difference Mask (BDM) is generated the operations of Background Difference, can be shown as.

$$BD(x, y, t) = |I(x, y, t) - BG(x, y, t - 1)| \quad (6)$$

$$BDM(x, y, t) = 1, \text{ if } BD \geq Th$$

$$0, \text{ if } BD < Th \quad (7)$$

where *BD* is background difference, *BG* is background frame, and *BDM* is Background Difference Mask, respectively.

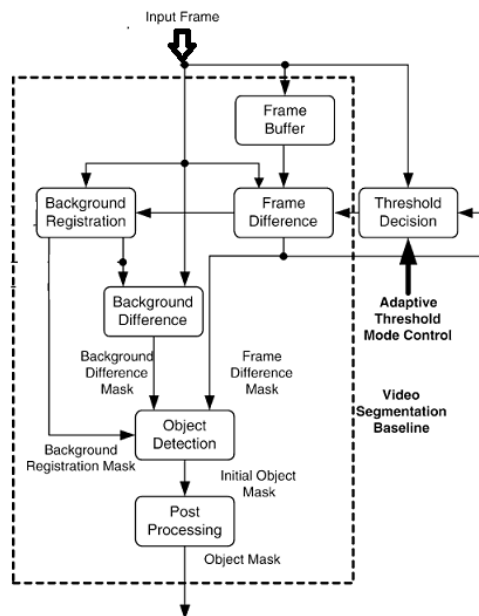


Figure. 1. Block diagram of video segmentation Algorithm Using change detection

For cases 3 to 6 in the decision table, the criteria are background difference because the background information exists. If both the frame difference and the background difference are significant, the pixel is part of a moving object. On

the other hand, if both the frame difference and the background difference are insignificant, the pixel should not be included in the object mask.

Table 1. Situations of object detection

Situation	FDM	BDM	BI	IOM
Stationary	0	-	0	0
Moving	1	-	0	1
Background	0	0	1	0
Moving Object	1	1	1	1
Still Object	0	1	1	1
Uncovered Background	1	0	1	0

Therefore, for the third and fourth cases in Table I, our result is the same as the result of using only the frame difference for change detection.

Cases 5 and 6 are situations that frame difference based change detection cannot handle properly, but the background difference works. One of the problems that confuse the conventional change detector (only frame difference is used) is that the object may stop moving temporarily or move very slowly. In these cases, the motion information disappears if we check the frame difference only. However, if we have background difference information, we can see very clearly that these pixels belong to the object region and should be included in the object mask. For case 6, since both the uncovered background region and the moving object region have significant luminance change, distinguishing the uncovered background from the object is not very easy if only the frame difference is available. In this algorithm, the uncovered background region is handled correctly because we recognize that this region matches the background information even though frame difference suggests significant motion.

### Step IV – Object Detection

Both of FDM and BDM are input into Object Detection to produce Initial Object Mask. The procedure of Object Detection can be presented as the following equation.

$$IOM(x, y, t) = BDM(x, y, t), \text{ if } BI(x, y, t) = 1$$

$$FDM(x, y, t), \text{ else} \quad (8)$$

This process can deal with the six situations shown in Table I, where “-” means “not available.” Note that the last two situations are easily misclassified by other change detection based segmentation algorithms, where *BDM* information is not available. In other algorithms, the still objects are often taken as background objects because they are not included in *FDM*, and the uncovered background is often taken

as foreground object because it is included in *FDM*. Both of these two situations need complex post-processing algorithms to compensate the misclassification, which are not needed in the proposed algorithm.

### Step V- Post processing

The Initial Object Mask (IOM) generated by Object Detection has some noise regions because of irregular object motion and camera noise. Also, the boundary may not be very smooth. Therefore, there are two parts in Post processing: noise region elimination and boundary smoothing.

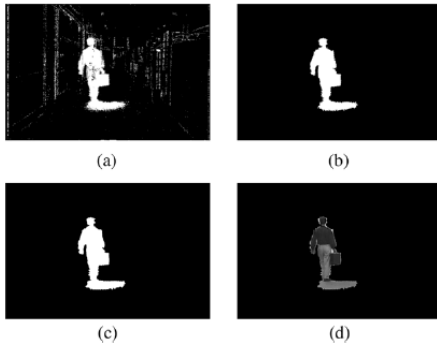


Figure. 2. Illustration of post-processing. (a) Initial object mask; (b) after noise elimination; (c) after morphological closing operation; (d) generated VOP.

The connected component algorithm [8] can mark each connected region with a special label. Then we can filter these regions by their area. If the area of a region is small, it may be a noise region and can be eliminated. Background regions, which are indicated by “0” in *IOM*, are first filtered, that is, background regions with small area are eliminated. This process eliminates holes in the change detection mask, which often occur especially when the texture of foreground objects is insignificant. Then foreground regions, which are indicated by “1” in *IOM*, are then filtered. This process removes noise regions. Next, the morphological close–open operations are applied to smooth the boundary of object mask. In addition, Stationary Index is further revised with *IOM* by

$$SI(x, y, t) = 0, \text{ if } IOM(x, y, t) = 1 \quad (9)$$

This process can avoid still objects to be registered into the background buffer. Figure 2 shows the effect of *Postprocessing*. Figure 2(a) is, where the white parts are those indicated by “1” in, and the black parts are those indicated by “0” in. After noise region elimination, the mask can be improved as

shown in Figure 2(b). After boundary smoothing, the improved mask is shown in Figure 2(c). Finally, the generated VOP is shown in Figure 2(d).

### 4. Adaptive Threshold mode.

The threshold is a very critical parameter for change detection based algorithms. If the optimal threshold cannot be decided automatically, these kinds of video segmentation algorithms are hardly used in real applications. Therefore, the automatic threshold decision is very important in our video segmentation system. After evaluating many thresholding methods for change detection, Rosin *et. al.* [17, 19] recommends three thresholding methods: Euler-number [20], Poisson-noise modelling [17], and Kapur method [18]. The Euler-number thresholding is based on the assumption that the number of regions of change in a difference image will tend to be stable over a wide range of threshold values. The Poisson-noise model thresholding is based on the assumption that observations (number of pixels over a specific threshold) in an image usually follow a Poisson distribution. The Kapur thresholding is entropy based. These three thresholding methods perform well for change detection. However, the loads of computation of the Euler number thresholding and Poisson noise modeling thresholding are high and not suitable in real-time. In addition, the Poisson noise modeling thresholding is sensitive to its parameter, the window size. The Euler method tends to under-threshold some images. The Kapur thresholding is sensitive to the noise level and under-threshold a difference image.

The proposed non-parametric algorithm [16] computes a threshold of each block of an image adaptively based on the scatter of regions of change (ROC) and averages all thresholds for image blocks to obtain the global threshold. First, the output  $D_n$  of change detection at time instant  $n$  is divided into  $K$  equal-sized blocks. Then a ROC scatter estimation algorithm is applied, where each image block  $W_k$ ,  $k = \{1, 2, \dots, K\}$ , is marked either as containing ROC, denoted  $W_k^r$ , or not containing ROC, denoted  $W_k^b$ . The threshold  $T_k^b$  of a  $W_k^b$  is computed by a noise statistical-testing algorithm. The threshold  $T_k^r$  of a  $W_k^r$  is computed by a noise-robust thresholding method. That is, the threshold  $T_k^f$  of a  $W_k$  in  $D_n$  is

$$T_k = \begin{cases} T_k^r & \text{of } W_k^r \\ T_k^b & \text{of } W_k^b \end{cases} \quad (10)$$

defined as

Finally, the global threshold  $T_n$  of a difference image  $D_n$  is

$$T_n = \frac{1}{K} \sum_{k=1}^K T_k \quad (11)$$

The ROC in  $D_n$  are, in general, scattered over the  $K$  image blocks. Let  $i$  be a pixel in  $D_n$  that varies between 0 and 255.  $i$  is high in ROC which is caused by strong changes such as motion or significant illumination changes and is low in non ROC which is caused by slight changes such as noise or slight illumination changes. We use the first moment,  $mk$ , of the histogram of each image block  $W_k$  as a measure for determining if an image block contains ROC. If  $mk$  of  $W_k$  is greater than a threshold  $T_m$ , the image block is regarded as a block containing ROC, and marked as  $W_k^r$ , otherwise, it is marked as  $W_k^b$ , i.e.,

$$W_k = \begin{cases} W_k^b & : m_k \leq T_m \\ W_k^r & : m_k > T_m \end{cases} \quad (12)$$

To find  $T_m$ , we first compute the  $mk$  of each block and then descending sorts the  $mk$  values. A straight line between the first bin and the last filled bin is then drawn.  $T_m$  is selected to maximize the perpendicular distance between the line and the sorted first moment curve. Figure 3 shows an example of adaptive thresholding, where we used relative  $mk$

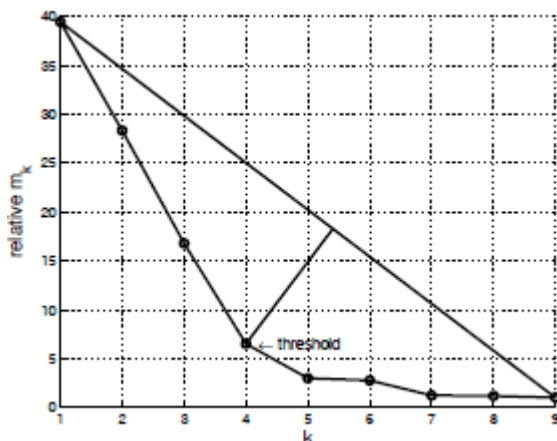


Figure. 3. An example of adaptive  $mk$  thresholding ( $K = 9$ ).

### 5. Experimental Results:

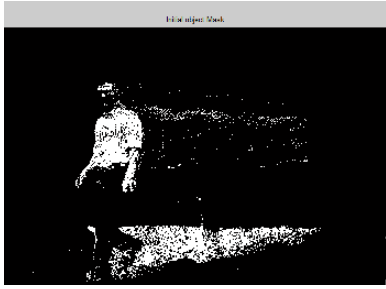
a) Using Adaptive threshold technique:



b) Using Pixel –based Change detection:



c) Initial Object Mask



## 6. Discussions.

The algorithm cannot deal with strong light sources. Texture and luminance of background should be different with the foreground moving object. The method is designed for moving objects segmentation, and background should never move. The project work is still in progress to get the desired results.

## 7. Conclusions.

This paper proposes, background registration and change detection based video segmentation algorithm with a real time adaptive threshold techniques to decide the parameters automatically, This algorithm can generate segmentation results with low computation complexity and high efficiency compare to other change detection based video segmentation algorithm. Finally, this algorithm is designed for moving objects segmentation; therefore, for stable and accurate results, the foreground object should not be still for a long time.

## 8. References.

1. Sikora T., The MPEG-4 Video Standard Verification Model (IEEE Trans. on Circuits and Systems for Video Technology), 7(1), 1997, 19-31.
2. Zang D.S. and Lu G., Segmentation of Moving Object in Image Sequence: A Review (Circuits, Systems and Signal Processing (Special Issue on Multimedia Services)), 20(2), 2001, 143-183.
3. Kim M., Choi J.G., Kim D., Lee H., Lee M.H., Ahn C., Ho Y-S., A VOP Generation Tool : Automatic Segmentation of Moving Objects in Image Sequences Based on Spatio-Temporal Information (IEEE Trans. on Circuits and Systems for Video Technology), 9(8), 1999, 1216-1226.
4. Park H.W., Schoepflin T., Kim Y., Active Contour Model with Gradient Directional Information: Directional Snake (IEEE Trans. on Circuits and Systems for Video Technology), 11(2), 2001, 252-256.
5. Gatica-Perez D., Gu C., Sun M-T., Semantic Video Extraction Using Four-Band Watershed and Partition Lattice Operators (IEEE Trans. on Circuits and Systems for Video Technology), 11(5), 2001 603-618
6. Kim C. and Hwang J-N., Fast and automatic video object segmentation and tracking for content-based applications (IEEE Trans. on Circuits and Systems for Video Technology), 12(2), 2002, 122-129.
7. Kim B-C. and Park R-H., A fast automatic VOP generation using boundary block segmentation (Real-Time Imaging), 10(2), 2004, 117-125.
8. Mech R. and Wollborn M., A Noise Robust Method for 2D Shape Estimation of Moving Objects in Video Sequences Considering a Moving Camera (Signal Processing), 66(2), 1998, 203-217.
9. Neri A., Colonnese S., Russo G., Talone P., Automatic moving object and background separation (Signal Processing), 66(2), 1998, 219-232.
10. Chien S-Y., Huang Y-W., Hsieh B-Y., Ma S-Y., Chen L-G., Fast video segmentation algorithm with shadow cancellation, global motion compensation, and adaptive threshold techniques (IEEE Trans. on Multimedia), 6(5), 2004, 732-748.
11. Urhan O., Ertürk S., Video segmentation with block based change detection using numerous preceding image frames, New Trends in Computer Networks, Advances in Computer Science and Engineering:
12. "The Essential guide to video processing" A book by Alan Conrad Bovik.
13. "Advances in image and video segmentation" A book by Yu Jin Zhang.
14. "Video Segmentation Based on Image Change Detection for Surveillance Systems". Tung-Chien Chen.
15. A Review of Real-time Segmentation of Uncompressed Video Sequences for Content-based Search and Retrieval Sébastien Lefevre 1, 2, Jérôme Holler 1, Nicole Vincent.
16. ChangSu and AishyAmer "A Real Time Adaptive Thresholding for video change detection" .IEEE International Conference on Image Processing (ICIP), pp 157-160, Oct-2006.
17. P.L.Rosin,"Thresholding for change detection," Computer Vision and Image Understanding, vol. 86, pp. 79-95, 2002.
18. J. Kapur, P. Sahoo, and A. Wong, "A new method for gray-level picture thresholding using the entropy of histogram," Computer Vision, Graphics Image Process, vol. 29, no. 3, pp. 273-285, 1985.
19. P. L. Rosin and E. Ioannidis, "Evaluation of global image thresholding for change detection," Pattern Recognition Letters, vol. 24, pp. 2345-2356, 2003.
20. P. L. Rosin and T. Ellis, "Image difference threshold strategies and shadow detection," In Proc. British Machine Vision Conference, pp. 347-356, 1995.