International Journal of Electronics Communication and Computer Engineering
Volume 4, Issue 2, ISSN (Online): 2249–071X, ISSN (Print): 2278–4209

Surveillance of Video to Detect Moving Objects Using Background Subtraction Algorithm

Ms. G. Selvi
M.E-Final Year(CSE),
Kalasalingam Institute of Technology
seana.vetri@gmail.com

Mr. V. Ramesh
M. Tech., Assistant Professor,
Kalasalingam Institute of Technology
ramesh_8607@yahoo.co.in

***Abstract* - Video Surveillance is needed for governments and government organizations to maintain social control, recognize and notice threats, and prevent criminal activity. In our approach, frames are taken from video and are converted to images. For each pixel in the image, the values of that pixel that are stored previously at the same location is taken and it is compared with the current pixel value to determine the background pixel. Then background pixel is removed and only foreground pixels that comprises moving object are detected. When the pixel is found as a background pixel, then its value is transmitted into neighboring pixel's background model. By this way, current frame is compared with available frames that are already stored in database. Our approach have efficiency in detection rate and computation speed. The algorithm called Vibe is used to find foreground pixels. Then we apply Temporal Difference technique to find foreground pixels. Performance is evaluated and compared for these two algorithms. It also requires a minimum of one comparison and per pixel one byte of memory is needed.**

***Keywords* - Background Subtraction, Image Segmentation, Pixel Classification, Surveillance.**

I. Introduction

Video Surveillance is major application in more organizations. Those organizations need to watch over video to detect illegal activity. So there is a need for continuous monitoring of scene with in video camera. Surveillance of video also needed in ATM booth, parking lot, Intelligent Highways, Bio/Pharmaceuticals, etc.This generates large volumes of data to record. To reduce data storage requirements (i.e) to reduce size of each video frame to record and to reduce total number of video frames to record, some idea must be established.

That idea is to use motion detection algorithms. The algorithms compares each pixel of a static background frame with corresponding pixel of the current frame of a video sequence. This technique builds a background model and compares this model with the current frame in order to detect parts in which place significant difference occurs. Here moving object is known as Foreground and part of scene is known as Background. If a static object from background starts moving, motion detection algorithm detects motion of object and a hole in scene in which place object occurs previously. This hole is referred to as Ghost and this is irrelevant. So it is discarded.

Since a static background model is suitable for video sequences which are shorter in an indoor environment, the model is invaluable for most practical situations. So there is a need for more sophisticated model. A first step in understanding the scene is detecting motion areas. Motion detected areas might be filtered and characterized for the detection of traffic surveillance, gait recognition, people counting, face detection, unsupervised bags, etc.

II. Review of Background Subtraction Algorithms

In background subtraction techniques, the current image is compared with already estimated image which have no objects. This is referred as the background model or background image [1].This comparison process is called Foreground detection. It divides the observed image into two sets of pixels that fully cover the entire image: 1) the foreground that contains the objects and 2) The background.

According to [2], a background subtraction technique must adapt to gradual or fast illumination changes. That are time of day, clouds, motion changes that is camera oscillations, moving background objects e.g tree leave or branches, and changes in the background scene e.g., parked cars. Background subtraction algorithms are required in some applications to be embedded in the camera. Robustness to noise and adapt to illumination changes are needed for the surveillance of outdoor scenes.

Review and evaluation of commonly implemented Background Subtraction algorithms is described in [3].Here evaluation is performed on a large video dataset that is made up of various real, synthetic video sequences. Five implemented motion detection methods, One Gaussian *(1-G)*, Gaussian Mixture Model *(GMM)*,Kernel Density Estimation *(KDE)*, Basic Motion Detection *(Basic)*, Minimum, Maximum and Maximum Inter-Frame Difference *(MinMax)* are described. Evaluation on Noisy Videos, Noise-Free with Perfectly Static Background Video and Multimodal Videos are compared.

Evaluation of BGS algorithms and measuring performance with post-processing techniques are described in [4].There are two techniques in Background modeling. Recursive techniques and Non-recursive techniques. It aims at detection of foreground. Post processing Techniques used are Noise removal, Blob processing, Optical flow test and Object level feedback. Main feature is computational efficiency, robustness to noise and simplicity.

In [5], Implementation of the Stauffer-Grimson algorithm is described. It uses a mixture of normal distributions to model a multimodal background image sequence. Approach used is to maintain a background image as a cumulative average of the video stream and to segment moving objects by thresholding a per-pixel distance between the current frame and the background image. It has advantage of Computational Speed and background is visible more frequently than any foregrounds.

Background subtraction techniques are surveyed and analysed with respect to three important attributes in [6].They are foreground detection, background maintenance, and postprocessing. In [7], Many existing algorithms for background removal are surveyed. Survey is also done on pre-processing algorithms, different background models, and way used to update background model and how they are initialized. And also performance measure of moving object detection algorithm is surveyed.

In [8], Background modeling is used to detect moving objects from static cameras. Classification of pixels is done in term of generation following the years of publication and the statistical tools used. First generation methods mainly focused are Single Gaussian, Mixture of Gaussians, Kernel Density Estimation and Subspace Learning using PCA. The improvements of these methods have been classified in term of strategies. The strategies are analyzed and their limitations are identified. In [9], Background subtraction is viewed as a signal estimation problem and proposed method is able to overcome various difficulties that are occurred in real applications.

In [10], an efficient foreground detection algorithm based on a new color space model and morphological filtering is proposed. Each pixel's color distortion and brightness distortion is used to detect the foreground pixels. The color distortion considers the vector's position in color space. Using this color features are assembled effectively. Also shadow elimination process is used to remove the moving shadow. By applying it to background subtraction, complete foreground object is detected. The post-processing and the filtering stage helps to achieve better results in terms of complex background.

All these unimodal techniques gives us satisfactory results in controlled environments and it is fast, easier to implement, and simpler. When dealing with videos that are captured in complex environments these sophisticated methods are necessary. In complex environments, moving background, camera ego motion, and high sensor noise are occurred [6].

Algorithm used in this paper works in a different way in handling new or fading objects in the background. In our method ghost that refers to region of the background that are discovered is added to the background model more quickly than an object that stops moving when a static object starts moving. Here samples are gathered from the past and used to update the sample values by omitting when they were added to the background models. This ensures a smooth exponential decaying lifespan for the sample values of the pixel models.

III. VIBE ALGORITHM

A Background Subtraction technique, called Visual Background extractor (ViBe) is used.

A. Pixel model and classification of pixels

Each background pixel is modeled with a collection of samples instead of using an explicit pixel model. so the current value of the pixel is compared to its nearest samples within the collection of samples. In comparing with existing algorithms this is an important difference. A new value is compared to background samples values and it should be close to some of the sample values instead of the majority of all values in collections.

V(x) is the value of pixel in a Euclidean color space at x in the image, and v_i is a background sample value with an index i. Each background pixel x is modeled by a collection of N background sample values

$$M(x) = \{v_1, v_2, \ldots. v_n\}$$

that are taken in previous frames. To classify a pixel value v(x) according to its corresponding background model M(x) ,comparison is done with the closest values within the set of samples. A sphere $S_R(v(x))$ of radius R centered on v(x) is defined. The pixel value v(x) is then classified as background if the cardinality(#) of the set intersection of this sphere and the collection of model samples M(x) is larger than or equal to a given threshold $\#_{min}$.

B. Background model initialisation from a single frame

It is the requirement for many applications to provide an uninterrupted foreground detection, when there is a presence of sudden light changes and this cannot be handled by the regular update mechanism of the algorithm properly. A possible solution to this problem is to provide a specific model update process that strengthen the pixel models to adapt to new lighting conditions. The best solution is to provide a technique/algorithm that will initialize the background model from a single frame. In this technique, response to sudden illumination changes is possible. Here the existing background model is discarded and a new background model is initialized as quickly as possible. For short video sequences in video-surveillance the second frame of a sequence has advantages and also it is advantageous for devices that embed a motion detection algorithm. Here aim is to provide a reliable foreground segmentation. The pixel models are filled with the values that are occurred in the neighborhood pixel of each pixel. This values taken in their neighborhood pixel in the first frame randomly.

A ghost may be introduced. Here it is refers to the presence of a moving object in the first frame. A ghost is a set of connected points or pixels that are detected as in motion but that are not corresponding to any real moving object. Sometimes the ghost is caused by the unfortunate initialization of pixel models with sample pixels that are coming from the moving object. The object moves in video makes the ghost object fade over time in subsequent frames. So here recovery of model will be very fast when a ghost appears in scene. This is done by update process. It also ensures a slow addition of real moving objects into the background model.

C. Updating the Background Model over Time

The background model is continuously updated with each new frame. The update process is adapted to lighting changes and handle new moving objects that appear in a scene.

International Journal of Electronics Communication and Computer Engineering
Volume 4, Issue 2, ISSN (Online): 2249–071X, ISSN (Print): 2278–4209

A sample that belongs to a foreground region in the more than 10s is never included by update policy.

Update process has three important components:

1) A memory less update policy, ensures a smooth decaying lifespan for the samples that stored in the background pixel models
2) A random time subsampling used to extend the time windows enveloped by the background pixel models
3) A mechanism guarantees spatial consistency by propagating background pixel samples spatially and allow background pixel models to adapt that are masked by the foreground.

A Memoryless Update Policy:

Many sample-based models are updated by using first-in first-out policies. Each pixel modal contain sample values that are taken previously. So for each pixel the oldest sample values are kept remains and it can't be discarded in any case. According to a uniform probability density function, the sample to be discarded is chosen randomly. This is an alternative way instead of removing the oldest sample from the pixel model. Then the new value replaces the selected sample. Many techniques have the idea of older values should be replaced first. So this idea gets complicated by using this technique. A conservative update policy is necessary for the stability of the process and conservative update policy guarantees that the models does not get difference over time.

Time subsampling:

Each background pixel model have fixed size time window. So to extend the size of the time window random time subsampling is used. It is not necessary to update each background pixel model for each new frame. We extend the expected lifespan of the background samples by making the background model update. Sometime the background model does not adapt to the periodic background motions. This is done by using fixed subsampling intervals. But the background model must adapt to the background motions. The solution is to use a random subsampling policy. When a pixel value has identified as a background pixel, the random process determines whether that pixel is used to update corresponding pixel model.

Background samples propagation:

The number of consecutive times that a pixel has been identified as foreground is counted. For this purpose detection support map technique is used here. If this number reaches a given threshold for a particular pixel location, the current pixel value at that location is inserted into the background model. The pixel values that are classified as foreground does not included in any background pixel model. The support map methods delay the inclusion of foreground pixels into background model. Sometimes foreground object may be included in background model. At that moment, time is needed to recover.

The background pixels that are close to each other share a similar temporal distribution. A new background sample of a pixel update the neighboring pixels model values. Background models will be updated with background samples from neighboring pixel locations from time to time that are hidden by the foreground. The Background model is adapted to changing illumination levels and can respond to changes in video when objects are added or removed in scene.

Since each pixel models have many samples, some irrelevant information may be inserted into the neighborhood model .But this does not affect the accuracy of the detection. The erroneous spreading of irrelevant information is blocked before it can spread further. This limitation prevents spreading of error information.

ViBe does not consider the video stream frame rate or color space, the scene content, the background itself, or background variation over time.

IV. Temporal Difference Algorithm

Temporal differencing method uses the pixel-wise difference between two or three consecutive frames in video imagery to extract moving regions. It is a highly adaptive approach to dynamic scene changes.

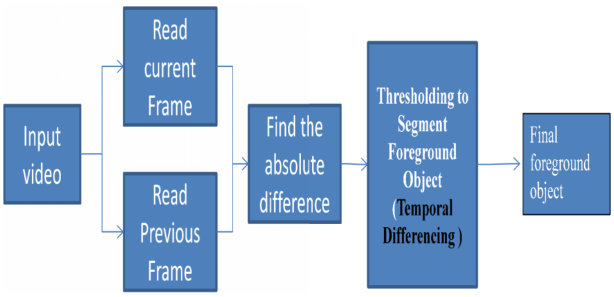


Fig.1. Background Subtraction Block Diagram

Fig.1 shows how foreground objects are segmented from video file using Temporal Difference Algorithm. The predetermined threshold value is used to segment foreground objects.

Let In(x) represent the gray-level intensity value at pixel position x and at time instance n of video image sequence I, which is in the range [0, 255]. T is the threshold initially set to a pre-determined value. Two-frame temporal differencing scheme suggests that a pixel is moving if it satisfies the following:

$$| I_n(x) - I_{n-1}(x) | > T$$

This method is computationally less complex .It is adaptive to dynamic changes in the video frames.

V. Implementation

The steps involved for the computation of ViBe are given.

• *Segmentation step:*

We compare a new pixel value to background samples to find two matches. Once two matches have been found, we step over to the next pixel and ignore the remaining background samples.

Operations involved during the segmentation step are:

1) Comparison of the current pixel value with the values of the background model. Mostly, the two first values of the background model of a pixel are close to the new pixel value.

2) We only need to compare the counter value after the comparison between the current pixel value and the second value of the background model.

• *Update step:*

1) 1 pixel substitution per 16 background pixels. Then we perform a similar operation,for a pixel in the neighborhood .first we locate which pixel in the neighborhood to select, then which value to substitute.
2) The cost of the update step is three additions on memory addresses per 16 background pixels.

• *Summary (average per pixel, assuming that most pixels belong to the background):*

1) 4 subtraction on bytes.
2) 3/16 addition on memory addresses.

Determination of Parameters

ViBe has the following parameters:

- Radius of the sphere (R) that is used to compare a new pixel value to pixel samples
- The time subsampling factor(ϕ)
- The number of samples (N) stored in each pixel model
- The number of close pixel samples ($\#_{min}$) needed to classify a new pixel value as background.

These parameters are fixed to N=20,R=20, ϕ=16 and $\#_{min}$=2.

Since ViBe has a low computational cost, it is well suited to an embedded implementation and the computational cost of ViBe can be further reduced by using low values for N and $\#_{min}$.In this way, ViBe is used to find foreground pixels of a video.

Temporal Difference Algorithm

The steps involved for the temporal difference technique are given.

- Each and every moving object tracking method requires an object detection mechanism either in every frame or when object first appears in video.
- We use the temporal information computed from a sequence of frames to reduce the number of false detections.
- This temporal information is usually in the form of frame differencing, which highlights regions that changes dynamically in consecutive frames.
- Segmentation of moving objects from fixed background objects is achieved.

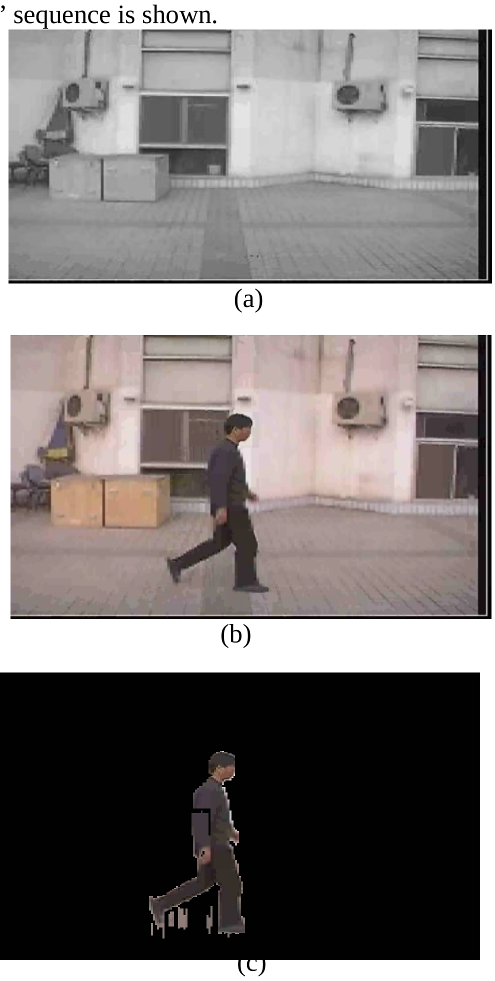
VI. Experimental Results

In this section, experimental results of the ViBE algorithm and of the temporal difference algorithm are presented. The system has been implemented in MATLAB. Tests were executed on a Intel core i3 with a CPU frequency of 3.10GHz and 4GB RAM.

Experiments were carried out on selected sequences from the PetsD1TeC1 dataset (available at http://www.research.ibm.com/peoplevision/PetsD1TeC1.avi)

We also tested videos taken from laboratory. They contain a single moving target that is a red remote controlled car. The videos are taken from a number of positions and are specifically designed so that background subtraction can work well on them. This sequence is available at http://www.cvc.uab.es/~bagdanov/master/videos/car-overhead-1.avi.

In Fig.2 Foreground segmentation map of background subtraction technique for one frame taken from the "house" sequence is shown.



(c)

Fig.2. Foreground segmentation. (a) Background image (b) Image with moving human (c) Moving object segmented from b

Both pets and house sequences are challenging. Because they contain background motion, moving trees, bushes,and illumination changes. The "pets" sequence is used to determine objective values for parameters of ViBe.

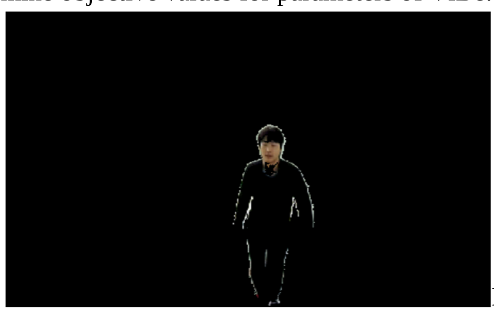


Fig.3. Foreground Segmentation

Fig.3 shows moving object is segmented from background image.This sequence available in the url http://ivylab.kaist.ac.kr/demo/vs/dataset/Stair.mpg.And also test is made on sequence available at http://www.cvc.uab.es/~bagdanov/master/videos/car-perspective-1.avi

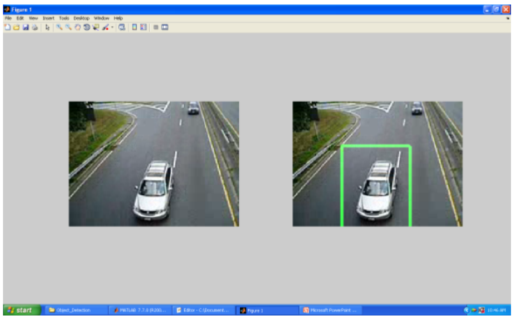


Fig.4. Object tracking

Fig.4- shows how moving objects are segmented from image using Temporal Difference Algorithm. Object detection can be achieved by building a background model for every pixel and then finding difference between frame and background model for each incoming frame. Any significant change in an image from the background model ensures that moving object is detected. The combined pixels making regions that are changing continuously are marked for further processing.

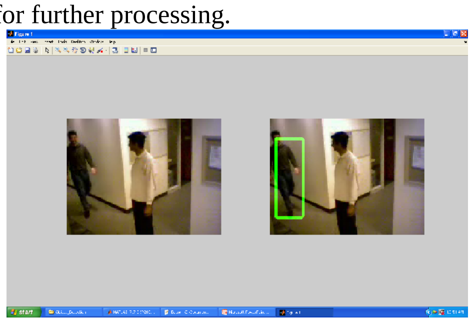


Fig.5. Object tracking

Fig. 5 shows how moving objects are segmented from image using Temporal Difference Algorithm. This sequence is available at
http://www.youtube.com/watch?v=Ib5lxGBhamI.
Performance is evaluated based on three parameters.

- Detection processing time.
- Tracking window size.
- False negative rate (FNR)

Table 1: Performance Evaluation

Video	Processing Time per frame	No of frames	False Negative Rate	Tracking Accuracy
traffc.avi	0.1198	120	0.1917	0.8083
Vipmen.avi	0.0818	283	0	100

VII. Conclusion

Pixel model and classification process is based on similarity between a candidate value and the corresponding background pixel model. Then it shows how ViBe can be initialized with a single frame. So the need to wait for several seconds to initialize the background model is lesser. Instead of keeping samples in the pixel models for a fixed amount of time, we ignore the insertion time pixel in the model and select a value to be replaced randomly. This results in smooth decaying lifespan for pixel samples and reducing required number of samples needing to be stored for each pixel. The spatial process is responsible for better resilience to camera motions. It also frees us from the need to post process segmentation maps in order to obtain spatially coherent results. The spatial propagation technique and update mechanism are combined with a conservative update scheme: no foreground pixel value should be included in any background model. This makes our method more effective to detect foreground objects.

References

1) O. Barnich and M. Van Droogenbroeck, “ViBe: A powerful random technique to estimate the background in video sequences,” in *Proc. Int.Conf. Acoust., Speech Signal Process.*, Apr. 2009, pp. 945–948.
2) M. Piccardi, “Background subtraction techniques: A review,” in *Proc.IEEE Int. Conf. Syst Man Cybern.*, The Hague, The Netherlands,Oct.2004, vol. 4, pp. 3099–3104.,
3) Y. Benezeth, P. Jodoin, B. Emile, H. Laurent, and C. Rosenberger,“Review and evaluation of commonly-implemented background subtraction algorithms,” in *Proc. IEEE Int. Conf. Pattern Recognit., Dec.*2008, pp. 1–4.
4) D. Parks and S. Fels, “Evaluation of background subtraction algorithms with post-processing,” in *Proc. IEEE Int. Conf. Adv. Video Signal Based Surveillance, Santa Fe, New Mexico, Sep. 2008, pp.*192–199.
5) P. Power and J. Schoonees, “Understanding background mixture models for foreground segmentation,” in *Proc. Image Vis. Comput.*,Auckland, New Zealand, Nov. 2002, pp. 267–271.
6) A. McIvor, “Background subtraction techniques,” in *Proc. Image Vis.Comput.*, Auckland, New Zealand, Nov. 2000.
7) S. Elhabian, K. El-Sayed, and S. Ahmed, “Moving object detection inspatial domain using background removal techniques—State-of-art,”*Recent Pat. Comput. Sci.*, vol. 1, pp. 32–54, Jan. 2008.
8) T. Bouwmans, F. El Baf, and B. Vachon, “Statistical background modeling for foreground detection: A survey,” in *Handbook of PatternRecognition and Computer Vision (Volume 4)*. Singapore: WorldScientific, Jan. 2010, ch. 3, pp. 181–199.
9) M. Dikmen and T. Huang, “Robust estimation of foreground in surveillance videos by sparse error estimation,” in *Proc. IEEE Int. Conf. PatternRecognit.*, Tampa, FL, Dec. 2008, pp. 1–4.
10) M. Sivabalakrishnan and D. Manjula, “An efficient foreground detection algorithm for visual surveillance system,” *Int. J. Comput. Sci. NetworkSec.*, vol. 9, pp. 221–227, May 2009.

Author’s Profile

G. Selvi
completed her school education in C. M. Hr. Sec. School, Tuticorin. She did her BE degree in Computer Science and Engineering in Mepco Schlenk Engineering college, Virudhunagar. She is presently doing her final year Master Of Engineering in Computer Science and Engineering in Kalasalingam Institute of Technology, Virudhunagar. Her current interests are Image Processing and Video Surveillance. She has presented her paper in three National Conferences and one International Conference.
E-mail : seana.vetri@gmail.com.