ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 9, September 2017

A Survey Methods in Adaptive Background Subtraction in Images

Neha Sakpal, Manoj Sabnis

PG Student, Dept. of IT, V.E.S Institute of Technology, Mumbai, India

Associate Professor, Dept. of IT, V.E.S Institute of Technology, Mumbai, India

ABSTRACT: Background subtraction is a widely used approach for detecting moving objects from static cameras. There are many challenges in the development of a good background subtraction algorithm and researchers has been devoted to developing new invention and improvement techniques to overcome all the limitations. Many aspects involved in order to produce best detection system to detect moving object for the outdoor environment. Most methods exist from low to complex computational algorithm and performance, and each with different strengths and weaknesses. Difference technique applied would be solved difference issues and challenges face on background subtraction.

KEYWORDS: Background Subtraction Process, Parametric and Non Parametric Approaches

I. INTRODUCTION

Background subtraction, also known as Foreground Detection, is a technique in the fields of image processing and computer vision wherein an image's foreground is extracted for further processing. In this perspective, motion detection is often the first step of a multi-stage computer vision system (car tracking, person recognition, wild-life monitoring, etc.). For this reason, it is usually required to be as fast and as simple as possible. Consequently, most Background Subtraction methods label in motion every pixel at time 't' whose color is significantly different from the ones in the background. This solution has proven successful whenever the camera is rigorously static with a fixed noise-free background. But detecting motion through background subtraction is not always as easy as it may first appear. Indeed, some videos with poor signal-to-noise ratio caused by a low quality camera, compression artifacts or a noisy environment, are likely to generate numerous false positives. False positives can also be induced by illumination changes (gradual or sudden), an animated background (waves on the water, trees shaken by the wind), or camera jitter to name a few. On the other hand, false negatives can also occur when a moving object is made of colors similar to the ones in the background (the so-called camouflage effect). In order to cope with those challenges, this paper provide a main methods of background subtraction. Those methods are more robust to noise and background instability than the basic background subtraction approaches. The following issues should be considered while performing background subtraction:

• Light changes: There is a possibility of frequent change in illumination in the outdoor scenes, resulting from clouds, rains, fogs, etc.

• Moving background: Natural phenomena such as storms and winds also cause changes in background due to the movement of tree branches, bushes, etc.

• Cast shadows: The background model should include the shadow cast by moving objects that apparently behaves itself moving, in order to have a more accurate detection of the moving object's shape.

• Bootstrapping: The background model should be properly set up even in the absence of a complete and static (free of moving objects) training set at the beginning of the sequence.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 9, September 2017

• Camouflage: Moving objects should be detected even if their chromatic features are similar to those of the background model.

The rest of the paper is organized as follows: Section 2 describes the main process of Background Subtraction. Section 3 presents the different methods of Background Subtraction Algorithm. Section 4 presents the conclusion of this paper.

II. Background Subtraction Algorithm

Most literatures on background subtraction algorithm consider four major steps namely pre-processing, background modelling, foreground detection and finally is data validation. following Figure. 1 depicted the overall view of a generic background subtraction.

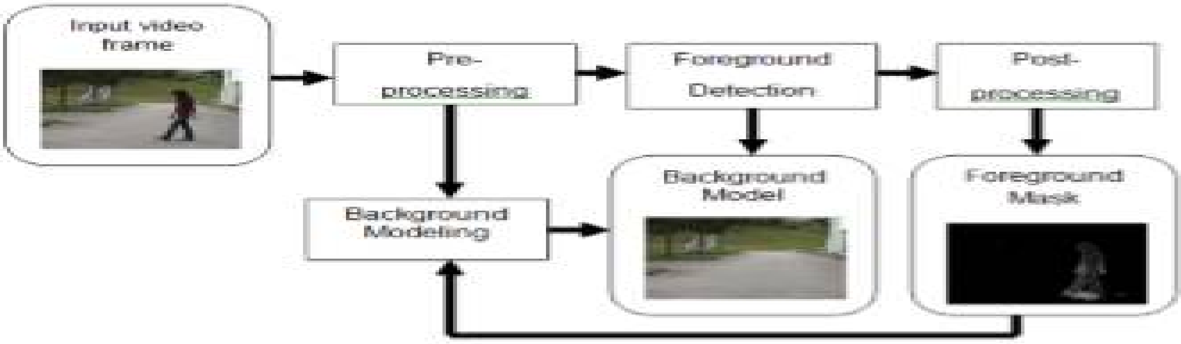

Figure. 1 Flow diagram of a generic background subtraction algorithm

A. Preprocessing

In most computer vision systems, simple temporal and/or spatial smoothing are used in the early stage of processing to reduce camera noise. Smoothing can also be used to remove transient environmental noise such as rain and snow captured in outdoor camera. For real-time systems, frame-size and frame-rate reduction are commonly used to reduce the data processing rate. If the camera is moving or multiple cameras are used at different locations, image registration between successive frames or among different cameras is needed before background modelling.

Another key issue in pre-processing is the data format used by the particular background subtraction algorithm. Most of the algorithms handle luminance intensity, which is one scalar value per each pixel. However, colour image, in either RGB or HSV colour space is becoming more popular in the background subtraction.

B. Background Modelling

Background modelling is at the heart of any background subtraction algorithm. Much research has been devoted to developing a background model that is robust against environmental changes in the background. We classify background modelling techniques into two broad categories – non-recursive and recursive. We focus only on highly-adaptive techniques, and exclude those that require significant resource for initialization. These include schemes described in , which store tens of seconds of video to construct initial background models that are characterized by eigen-images or temporal maximum, minimum, and maximum inter-frame differences of all identified background pixels.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 9, September 2017

C. Foreground Detection

Foreground detection compares the input video frame with the background model, and identifies candidate foreground pixels from the input frame. Except for the non-parametric model and the Mixture of Gaussian model, use a single image as their background models. The most commonly used approach for foreground detection is to check whether the input pixel is significantly different from the corresponding background estimate:

$$|\quad (x, y) - \quad (x, y)| > T \qquad \text{eq. (1)}$$

Another popular foreground detection scheme is to threshold based on the normalized statistics:

$$\frac{(,)- \quad (,)-}{} > \qquad \text{eq. (2)}$$

where μ and are the mean and the standard deviation of $(,) - \quad (,)$ for all spatial locations (x, y). Most schemes determine the foreground threshold T or Ts experimentally. Ideally, the threshold should be a function of the spatial location (x, y). For example, the threshold should be smaller for regions with low contrast.

Another approach to introduce spatial variability is to use two thresholds with hysteresis. The basic idea is to first identify "strong" foreground pixels whose absolute differences with the background estimates exceeded a large threshold. Then, foreground regions are grown from strong foreground pixels by including neighbouring pixels with absolute differences larger than a smaller threshold. The region growing can be performed by using a two-pass, connected-component grouping algorithm.

D. Data Validation

We define data validation as the process of improving the candidate foreground mask based on information obtained from outside the background model. All the background models have three main limitations: first, they ignore any correlation between neighbouring pixels; second, the rate of adaption may not match the moving speed of the foreground objects and third, non-stationary pixels from moving leaves or shadow cast by moving objects are easily mistaken as true foreground objects.

The first problem typically results in small false-positive or false-negative regions distributed randomly across the candidate mask. The most common approach is to combine morphological filtering and connected component grouping to eliminate these regions. Applying morphological filtering on foreground masks eliminates isolated foreground pixels and merges nearby disconnected foreground regions. Many applications assume that all moving objects of interest must be larger than a certain size. Connected-component grouping can then be used to identify all connected foreground regions, and eliminates those that are too small to correspond to real moving objects.

When the background model adapts at a slower rate than the foreground scene, large areas of false foreground, commonly known as "ghosts", often occur. If the background model adapts too fast, it will fail to identify the portion of a foreground object that has corrupted the background model. A simple approach to alleviate these problems is to use multiple background models running at different adaptation rates, and periodically cross-validate between different models to improve performance. Sophisticated vision techniques can also be used to validate foreground detection. Computing optical flow for candidate foreground regions can eliminate ghost objects as they have no motion. Colour segmentation can be used to grow foreground regions by assuming similar colour composition throughout the entire object. If multiple cameras are available to capture the same scene at different angles, disparity information between cameras can be used to estimate depth. Depth information is useful as foreground objects are closer to the camera than background.

The moving-leaves problem can be addressed by using sophisticated background modelling techniques like Mixture of Gaussian and applying morphological filtering for clean-up. On the other hand, suppressing moving shadow is much more problematic, especially for luminance-only video. A recent survey and comparison of many shadow suppression algorithms can be found.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 9, September 2017

III.BACKGROUND SUBTRACTION TECHNIQUES

A. Recursive Techniques

Recursive techniques do not maintain a buffer for background estimation. Instead, they recursively update a single background model based on each input frame. As a result, input frames from distant past could have an effect on the current background model. Compared with non-recursive techniques, recursive techniques require less storage, but any error in the background model can linger for a much longer period of time. Most schemes include exponential weighting to discount the past, and incorporate positive decision feedback to use only background pixels for updating. Some of the representative recursive techniques are described below:

- *Approximated median filter:*

Due to the success of non-recursive median filtering, McFarlane and Schofield propose a simple recursive filter to estimate the median. This technique has also been used in background modelling for urban traffic monitoring. In this scheme, the running estimate of the median is incremented by one if the input pixel is larger than the estimate, and decreased by one if smaller. This estimate eventually converges to a value for which half of the input pixels are larger than and half are smaller than this value, that is, the median.

- *Kalman filter:*

Kalman filter is a widely-used recursive technique for tracking linear dynamical systems under Gaussian noise. Many different versions have been proposed for background modelling, differing mainly in the state spaces used for tracking. The simplest version uses only the luminance intensity. Karmann and von Brandt use both the intensity and its temporal derivative, while Koller, Weber, and Malik use the intensity and its spatial derivatives. The internal state of the system is described by the background intensity Bt and its temporal derivative $'$, which are recursively updated as follows:

$$[\quad '\ ] = \ .[\ _{-1}\quad '_{-1}\ ] + \ .(\quad - \ .\ .[\ _{-1}\quad '_{-1}\ ]) \qquad \text{eq. (3)}$$

Matrix A describes the background dynamics and H is the measurement matrix. Their particular values used in are as follows:

$$A=[1\ 0.7\ 0\ 0.7\ ],\quad = [1\ 0\ ] \qquad \text{eq. (4)}$$

The Kalman gain matrix switches between a slow adaptation rate α1 and a fast adaptation rate α2 > α1 based on whether It−1 is a foreground pixel:

$$= [\ _1\quad _1\ ] \quad \text{if} \quad _{-1} \text{ is foreground, and } [\ _2\quad _2\ ] \text{ otherwise}$$

- *Mixture of Gaussians (MoG):*

Unlike Kalman filter which tracks the evolution of a single Gaussian, the MoG method tracks multiple Gaussian distributions simultaneously. MoG has enjoyed tremendous popularity since it was first proposed for background modelling. MoG maintains a density function for each pixel. Thus, it is capable of handling multimodal background distributions. On the other hand, since MoG is parametric, the model parameters can be adaptively updated without keeping a large buffer of video frames. The pixel distribution $(\quad = \quad)$ is modelled as a mixture of K Gaussians

$$(\quad = \quad) = \Sigma_{=1}\quad _{,}\ .\ (\ ;\ _{,}\ ,\ _{,}\ ) \qquad \text{eq. (5)}$$

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 9, September 2017

Where $(\; ; \; , \; , \;)$ is the Gaussian component with intensity mean $_{,}$ and standard deviation $_{,}$. $_{,}$ is the portion of the data accounted for by the i-th component. Typically, K ranges from three to five, depending on the available storage. For each input pixel , the first step is to identify the component $\hat{}$ whose mean is closest to . Component $\hat{}$ is declared as the matched component if $| \; - \; _{\hat{},\,-1}| \leq D \cdot \; _{\hat{},\,-1}$,where D defines a small positive deviation threshold. The parameters of the matched component are then updated as follows:

$$_{\hat{},} = (1 - \;)\; _{\hat{},\,-1} + \qquad \text{eq. (6)}$$
$$_{\hat{},} = (1 - \;)\; _{\hat{},\,-1} \qquad \text{eq. (7)}$$
$$^{2}_{\hat{},} = (1 - \;)\; ^{2}_{\hat{},\,-1} + \;(\; - \; _{\hat{},})^{2} \qquad \text{eq. (8)}$$

where α is a user-defined learning rate with 0≤α≤1. P is the learning rate for the parameters and can be approximated as follows:

$$\approx \frac{}{_{\hat{},}} \qquad \text{eq. (9)}$$

If no matched component can be found, the component with the least weight is replaced by a new component with mean , a large initial variance $_0$ and a small weight $_0$. The rest of the components maintain the same means and variances, but lower their weights to achieve exponential decay:

$$_{,} = (1 - \;)\; _{,\,-1} \qquad \text{eq. (10)}$$

Finally, all the weights are renormalized to sum up to one. To determine whether is a foreground pixel, we first rank all components by their values of $\frac{_{\hat{},}}{_{\hat{},}}$. Higher-rank components thus have low variances and high probabilities, which are typical characteristics of background. If $_1, \; _2, \ldots,$ is the component order after sorting, the first M components that satisfy the following criterion are declared to be the background components:

$$\sum_{=1} \; _{,} \geq \; , \qquad \text{eq. (11)}$$

where Γ is the weight threshold. It is declared as a foreground pixel if is within D times the standard deviation from the mean of any one of the background components. The computational complexity and storage requirement of MoG is linear in terms of the number of components K. Recent development in MoG technologies include a sensitivity analysis of parameters, improvements in complexity and adaptation and an extension to construct a panoramic background.

B. Non-recursive Techniques

A non-recursive technique uses a sliding-window approach for background estimation. It stores a buffer of the previous L video frames, and estimates the background image based on the temporal variation of each pixel within the buffer. Non-recursive techniques are highly adaptive as they do not depend on the history beyond those frames stored in the buffer. On the other hand, the storage requirement can be significant if a large buffer is needed to cope with slow-moving traffic. Given a fixed-size buffer, this problem can be partially alleviated by storing the video frames at a lower frame-rate r. Some of the commonly-used non-recursive techniques are described below:

- *Frame differencing:*

Arguably the simplest background modelling technique, frame differencing uses the video frame at time t−1 as the background model for the frame at time t. Since it uses only a single previous frame, frame differencing may not be able to identify the interior pixels of a large, uniformly-coloured moving object. This is commonly known as the aperture problem.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 9, September 2017

- *Median filter:*

Median filtering is one of the most commonly-used background modelling techniques. The background estimate is defined to be the median at each pixel location of all the frames in the buffer. The assumption is that the pixel stays in the background for more than half of the frames in the buffer. Median filtering has been extended to colour by replacing the median with the medoid. The complexity of computing the median is ( ) for each pixel.

- *Linear predictive filter:*

Compute the current background estimate by applying a linear predictive filter on the pixels in the buffer. The filter coefficients are estimated at each frame time based on the sample covariances, making this technique difficult to apply in real-time.

- *Non-parametric model:*

Unlike previous techniques that use a single background estimate at each pixel location, Elgammal use the entire history $_{-}\,,\ _{-}\,_{+1}, \dots,\ _{-1}$ to form a non-parametric estimate of the pixel density function $(\; = \;)$:

$$(\; = \;) = \frac{1}{}\sum_{=}^{-1} \; (\; - \;) \qquad \text{eq. (12)}$$

K (·) is the kernel estimator which was chosen to be Gaussian. The current pixel It is declared as foreground if it is unlikely to come from this distribution, i.e. ( ) is smaller than some predefined threshold. The advantage of using the full density function over a single estimate is the ability to handle multi-modal background distribution. Examples of multi-modal background include pixels from a swinging tree or near high-contrast edges where they flicker under small camera movement. The implementation uses the median of the absolute differences between successive frames as the width of the kernel. Thus, the complexity of building the model is the same as median filtering. On the other hand, the foreground detection is more complex as it needs to compute Equation for each pixel.

C. Multimodal Technique

Bayesian Modelling has three novel contributions. First, the method proposed provides a principled means of modelling the spatial dependencies of observed intensities. Second, unlike previous approaches, the foreground is explicitly modelled to augment the detection of objects without using tracking information. Third, instead of directly applying a threshold to membership probabilities, a MAP-MRF framework is used that competitively uses the foreground and background models for object detection. In this method high levels of detection accuracy are sustained even in the presence of nominal camera motion and dynamic textures. A few false positives and false negatives were detected by the proposed approach invariably at the edges of true objects, where factors such as pixel sampling affected the results. If a contiguous region of pixels was consistently detected corresponding to an object during its period within the field of view, a correct "object" detection was recorded. If two separate regions were assigned to an object, if an object was not detected, or if a region was spuriously detected, a misdetection was recorded.

D. Non- Parametric Techniques

Codebook model presents a real-time algorithm for foreground-background segmentation. Sample background values at each pixel are quantized into codebooks which represent a compressed form of background model for a long image sequence. This allows us to capture structural background variation due to periodic-like motion over a long period of time under limited memory. This method works well in compressed videos having irregular intensity distributions. The update mechanism does not allow the creation of new code words. This can be problematic if permanent structural changes occur in the background. In SOBS (Self-Organizing Background Subtraction) for each pixel, the corresponding weight vectors are initialized with the pixel value from the first image of the sequence. In order to represent each weight vector, the HSV colour space is chosen. After initialization, temporally subsequent samples are fed to the network. Each incoming pixel of the sequence frame is compared to the current pixel model C to determine if there exists a weight vector that best matches it. If a best matching weight vector is found, it means that belongs to the background and it is used as the pixel encoding approximation, and the best matching weight vector, together with its neighbourhood, is reinforced. The background model of this method adapts to gradual illumination changes. The selective update mechanism prevents the inclusion of foreground objects which has become static, into the background model.

E. Pixel- based Techniques

A Bayes decision rule for classification of background and foreground from selected feature vectors is formulated in Bayes decision theory. Under this rule, different types of background objects will be classified from foreground objects by choosing a proper feature vector. The stationary background object is described by the colour feature, and the moving background object is represented by the colour co-occurrence feature. Foreground objects are extracted by fusing the classification results from both stationary and moving pixels. This method works well under different conditions such as in videos containing wavering tree branches and curtains, flickering screens, lights, or water surfaces, moving escalators, opening/closing doors, removed/deposited objects, switching on/off lights, lighting condition changes from day to night or clouds and raining, and shadows of people on the ground surface. But it requires parameter tuning.

Texture based method presents a novel and efficient texture-based method for modelling the background and detecting moving objects from a video sequence. Each pixel is modelled as a group of adaptive local binary pattern histograms that are calculated over a circular region around the pixel. This method is tolerant to the multimodality of the background, and the introduction/removal of background objects. The method requires a non-moving camera, which restricts its usage in certain applications.

In SACON (SAmple CONsensus), a new efficient background modelling method, SAmple CONsensus is proposed and it is applied to background subtraction. SACON gathers background samples and computes sample consensus to estimate a statistical model at each pixel. SACON is easy to perform but highly effective in background modelling and subtraction. SACON is suitable for both indoor and outdoor applications. The training period for this algorithm must comprise at least 20 frames. To cope with lighting changes and objects appearing or fading in the background, two additional mechanisms (one at the pixel level, a second at the blob level) are added to the consensus algorithm to handle entire objects. This increases the complexity.

F. Region- based Techniques

Background Modelling via Classification is motivated by criteria leading to what a general and reasonable background model should be, and realized by a practical classification technique. The key idea of this approach is simple but effective: If a classifier can be used to determine which image blocks are part of the background, its outcomes can help to carryout appropriate block wise updates in learning such a model. A global consistency is utilized to eliminate noise in the updates. But this method requires manual parameter tuning. Based on the Spatial-Temporal correlations at pixels of the contour area of foreground blobs (PCFB), an efficient method for detecting ghost and left objects is proposed in this method. This method contains three main steps: firstly, the average background modelling together with background difference subtraction is adopted to draw foreground blobs. Secondly, a novel method based on temporal correlation of inter-frame at PCFB is introduced to detect candidate targets (ghosts and left objects).Thirdly, based on spatial correlation of intra-frame at PCFB, ghosts and left objects will be successfully discriminated respectively and background is updated accordingly. This algorithm could effectively detect ghosts and remove them immediately after an object leave the ghost area, and left objects could be differentiated from ghosts accurately. In this method there is a trade-off between computational complicity and real-time performance.

RECTGAUSS-Tex method presents a region-based method for background subtraction. It relies on colour histograms, texture information, and successive division of candidate rectangular image regions to model the background and detect motion. This method integrates texture and the Gaussian Mixture model. It is a multi scale rectangular region based motion detection and background subtraction algorithm. It filters noise during image

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 9, September 2017

differentiation. The choice of the coarsest rectangle size should be selected to be small enough to detect the object of interest. Thus, balancing the rectangle size for the detection of small objects might be incompatible for very small objects while filtering noise.

IV. CONCLUSION

Detecting motion through background subtraction is not always as easy as it may lead to many problems such as some videos with poor signal-to-noise ratio, noisy environment, illumination changes (gradual or sudden), an animated background (waves on the water, trees are shaken by the wind), or camera jitter to name a few. In this paper different Background Subtraction Techniques are introduced. These techniques are used to solve these problems which are faced in detecting static and moving background. These methods are more robust to noise and background instability than the basic background subtraction approaches.

REFERENCES

[1] S. Chen, J. Zhang, Y. Li and J. Zhang, "A Hierarchical Model Incorporating Segmented Regions and Pixel Descriptors for Video Background Subtraction," IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS, vol. 8, 2012.

[2] R. A. T. Jenifa, C. Akila and D. , "Rapid Background Subtraction from Video Sequences," International Conference on Computing, Electronics and Electrical Technologies [ICCEET] , p. 10, 2012.

[3] M. Piccardi, "Background subtraction techniques: a review*," IEEE International Conference on Systems, Man and Cybernetics , 2004.

[4] N. M. T. R. A. Shahrizat Shaik Mohamed, "Background Modelling and Background Subtraction Performance for Object Detection," 6th International Colloquium on Signal Processing & Its Applications (CSPA), 2010.