ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 5, May 2015

Background Subtraction for Effective Object Detection and its Parametric Evaluation

Siddhartha A. Meshram, Mr. Vishal Raskar

Department of Electronics & Tele-communication Engineering

*M.E Scholar,M. E Teacher , Imperial College Of Engineering and Research, Pune(M.H)India

Abstract: Background subtraction is a powerful mechanism for detecting change in a sequence of images that finds many applications. The background subtraction methods apply probabilistic models to background intensities evolving in time nonparametric and mixture-of Gaussians. The main difficulty in designing a robust background subtraction algorithm is the selection of a detection threshold. In this we adapt threshold to varying video statistics by means of two statistical models. In addition to a nonparametric background model we introduce a foreground model based on small spatial neighborhood to improve discrimination sensitivity we also apply a Markov model to change labels to improve spatial coherence of the detections, the proposed methodology is applicable to other background models as well. The strength of the scheme lies in its simplicity and the fact that it defines an intensity range for each pixel location in the background to accommodate illumination variation as well as motion in the background. The efficacy of the scheme is shown through comparative analysis with competitive methods. Both visual as well as quantitative measures show an improved performance and the scheme has a strong potential for applications in real time surveillance.

Keyword: Background subtraction, tracking, object detection, surveillance.

I] INTRODUCTION

Object detection and tracking in video is a challenging problem and it has been extensively investigated in the past two decades and it has applications in numerous fields, such as video compression, video surveillance, human-computer interaction, Object detection and object tracking are two closely related processes. The former involves locating object in the frames of a video sequence while the latter represents the process of monitoring the object's spatial and temporal changes in each frame. Object detection can be performed through various approaches, such as region-based image segmentation, temporal differencing, active contour models and generalized Hough transforms. In surveillance system video sequences are generally obtained through static cameras and fixed background.

The background subtraction is an important first step for many vision problems and it separates objects from background clutter usually by comparing motion patterns and facilitates subsequent higher-level operations such as tracking object identification cause the environment can change substantially both in the short term and throughout the lifetime of the vision system background subtraction algorithms are expected to be robust. This is not always easy to guarantee and many methods have been proposed in the literature. A popular approach called background subtraction is used in this scenario, where moving objects in a scene can be obtained by comparing each frame of the video with a background. In most of the suggested schemes the object detected is accompanied with misclassified foreground objects due to illumination variation or motion in the background.

In particular the background subtraction method is the most commonly used motion detection method and updating the background model are the keys to the ideal background model. The desired effect of moving object extraction with a self-adaptive background model is established in which self-adaptive updates according with the changes in the scene. The moving objects detection system needs to deal with the following important situations sudden and gradual changes in the light conditions shadow region small movements of non-static objects such as waving tree and ocean waves; ghost. Then most important tasks in all such applications is change detection i.e. automatic segmentation of a video sequence into static and changed (e.g. moving) areas. The numerous algorithms developed to date, the simplest ones are based on thresholding intensity differences. Since the detection results are sensitive to threshold selection (false positives versus misses) various threshold adaptation methods have been proposed.

The model parameters have been estimated from previous frames, and the detection process involves thresholding the resulting PDF to thresholding probabilities instead of intensities and the approach is more robust more constitutes a powerful tool for change detection. In parallel change detection methods that were developed based on the maximum a posteriori probability (MAP) criterion. While some methods were formulated in discrete domain and used MRFs as prior models other methods used variation formulations in continuous domain only embodying the spirit of the MAP criterion. The change detection performs well, it is computationally complex; approximate, but faster, solution methods were developed. Then revisit background subtraction from the hypothesis testing point of view and make two contributions. Assuming spatial ergodicity we augment the background model with an explicit foreground model and estimate its parameters from a small spatial neighborhood.

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 5, May 2015

To capture background dynamics, these approaches lack the two most compelling (and dynamic) aspects of the SG method: (1) the ability to account for transitory events, due to motion of foreground objects; and (2) simple model management. The absence of a hidden discrete state variable, the dynamic texture will slowly interpolate through all these states both the transition from occluded to turbulent, and turbulent to normal waves, will generate outliers which are incorrectly marked as foreground. Object detection in surveillance system background modeling plays a vital role and Wren *et al.* have proposed to model the background independently at each pixel location which is based on computation of Gaussian probability density function (pdf) on the previous pixel values.

The Stauffer and Grimson developed a complex procedure to accommodate permanent changes in the background scene. In this each pixel is modeled separately by a mixture of three to five Gaussians. It uses three values to represent each pixel in the background image namely the minimum intensity, the maximum intensity and difference between consecutive frames of the training sequence. Jacques *et al.* brought a small improvement to the W4 model together with the incorporation of a technique for shadow detection and removal [5]. McHugh *et al.* proposed an adaptive thresholding technique by means of two statistical models. One of them is nonparametric background model and the other one is foreground model based on spatial information. In ViBe each pixel in the background can take values from its preceding frames in same location or its neighbor, then it compares this set to the current pixel value in order to determine whether that pixel belongs to the background and adapts the model by choosing randomly which value to substitute from the background model. Then scheme adopts a clustering-based feature, called fuzzy color histogram (FCH), which has an ability of greatly attenuating color variations generated by background motions while highlighting moving objects.

Then scheme passes the texture information of each block through three cascading classifiers to classify them as background or foreground. The results are then integrated with a probabilistic voting scheme at pixel level for the final segmentation. Generally, shadow removal algorithms are employed after object detection. Salvador *et al.* developed a three step hypothesis based procedure to segment the shadows. It assumes that shadow reduces the intensities followed by a complex hypothesis using the geometrical properties of shadows. It confirms the validity of the previous assumption. Choi *et al.* in their work of distinguished shadows from moving objects by cascading three estimators, chromaticity, brightness, and local intensity ratio. A novel method for shadow removal using Markov random fields (MRF) where shadow model is constructed in a hierarchical manner. It is observed that most of the simple schemes are ineffective on videos with illumination variations, motion in background, and dynamically textured indoor and outdoor environment etc. On the other hand, such videos are well handled by complex schemes with higher computational cost.

Furthermore, to remove misclassified foreground objects and shadows, additional computation is also performed. Keeping this in view, we suggest here a simple scheme called Local Illumination based Background Subtraction (LIBS) that models the background by defining an intensity range for each pixel location in the scene. Subsequently a local thresholding approach for object extraction is used. Simulation has been carried out on standard videos and comparative analysis has been performed with competitive schemes. Visual-based object tracking is critical to automatically monitor object activities in video sequences. To view and detect object actions from a monocular scene, object occlusions usually incur detection errors due to objects in crowded environments. In this work, an object tracking scheme is proposed to overcome the occlusion effects and then to improve the accuracy of counting characters in a visual surveillance system. Up to the present, conventional work has applied computer-vision skills to detect motions and understand activities of characters in a stationary camera. In Marcenaro *et al*.'s study for relieving the effect of dynamic occlusion a multi-mode method to improve accuracy and efficiency for tracking multiple targets in a crowed scene.

A multi-class statistical model for the tracked objects but the background model is a single Gaussian per pixel. The attempt to mediate the effect of shadows appears to be somewhat successful but it is not clear behavior of their system would exhibit for pixels which did not contain these three distributions. For example pixels may present a single background color or multiple background colors resulting from repetitive motions, shadows, or reflectance. The subtraction regions of interest are identified from the foreground image based on the completeness of region contours. Additionally, detecting character regions from the interested ones is then considered under the physical constraints of human body including the dynamic range of pixel gray-level values and the shapes of regions. As objects conforming to the constraints ellipses having the minimized areas that can cover the regions are constructed to derive their corresponding position parameters, ellipse radiuses, and centroids. In the modified overlap tracker, four tracking states comprising new target, leaving target, merged target and split target are used to understand characters in the current frame. New target means an object entering a video scenes.

II. BACKGROUND SUBTRACTION

Background subtraction involves two distinct processes that work in a closed loop background modeling and foreground detection. In this background modeling a model of the background in the field of view of a camera is created and periodically updated for example to account for illumination changes. In foreground detection a decision is made as to whether a new intensity fits the background model the resulting change label field is fed back into background modeling so that no foreground intensities contaminate the background model. Let I be a grayscale image sampled on 2-D lattice $n \epsilon A \in R^2$ extension to color images is straightforward. We denote a sequence

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 5, May 2015

of such images with I(n). We consider a nonparametric background model for its simplicity and performance.

$$P_B\,(I^K(\,\mathrm{n})) = \frac{1}{N}\sum_{i-1}^{N} K(I^k\,(n) - I^{k-1}(n)) \qquad (1)$$

At each background location n of frame k, the model uses intensity from recent frames to estimate background PDF, where K is a zero-mean Gaussian with variance that, for simplicity, we consider constant throughout the sequence. If a similar foreground model is available then change labels can be estimated by evaluating intensity in a new frame against these two models.

$$\frac{P_B I(n)}{P_F I(n)} = \eta\,\frac{\pi_F}{\pi_B} \qquad (2)$$

This entails testing labels (background) and (foreground) at each pixel of the current image by means of a binary hypothesis test (from now on the superscript is omitted to simplify notation) where on left are the probabilities I(n) of observing given it is the projection of either the background scene or a foreground object. On the right-hand side, are the prior probabilities of observing background π_B or foreground π_F, and a cost term? The ratio biases the decision based on the priori probabilities, the cost term accounts for unequal penalties assigned to the four decision/truth scenarios.

$$P_F I(n) = \frac{1}{N_f(n)}\sum_{m\epsilon N_f} K(I(n) - I(m)) \qquad (3)$$

Without an explicit foreground model, P_k is usually considered uniform. Assuming fixed prior probabilities and collecting all constants in this leads to a fixed-threshold background test. A pixel is labeled as moving if its probability is sufficiently small and this simple test is prone to randomly-scattered false positives, even for low system. The additional post-processing has been proposed to correct such errors it is ad hoc and thus difficult to generalize. However this model is susceptible to illumination variation dynamic objects in the background and also to small changes in the background like waving of leaves etc. A number of solutions to such problems are reported where the background model is frequently updated at higher computational cost and thereby making them unsuitable for real time deployment.

III] EXTRACTION OF FOUR GROUNDS

In most background subtraction algorithms, P_F it is assumed uniform thus preventing any decision bias by moving objects. An exception is the work of El Gammal *et al.* [3], who proposed foreground modeling for human body and of Sheikh and Shah, who proposed a general foreground model using past frames. While the first model is object-specific, the second one necessitates slow object motion as otherwise background samples contaminate P_F. Although this could be mitigated by object tracking such an approach would be illogical (track an object in order to detect it). Instead, we propose a foreground model based on small spatial neighborhood, i.e. in the same frame. Recently, we have demonstrated that periodicity in time also holds spatially; local-in-time and local-in-space models produce equivalent background characteristics.

$$\mathrm{P_{fi}\;I(n)} = \frac{1}{\mathrm{Nf_i(n)}}\sum\nolimits_{\mathrm{m\epsilon Nf_{i(n)}}} \mathrm{K(I(n) - I(m))} \qquad (4)$$

After successfully developing the background model a local thresholding based background subtraction is used to find the foreground objects. This can be accomplished by modeling labels as a Markov random field of which is a particular realization. MRF models have been successfully used in motion detection reducing scattered false detections and smoothing region boundaries. We propose a Markov model within the binary hypothesis test while maintaining non parametric. Our approach extends early methods using single-Gaussian and uniform and shares Markovianity with more recent formulations. Also, despite the use of accelerated simulated annealing in, the computational complexity is high. Although one can seek local minima by means of one-at-a-time search such as the iterated conditional modes algorithm in this case a binary solution is identical to our binary hypothesis test. Also, our approach uses spatial periodicity whereas the one is based on temporal periodicity which necessitates slow motion or tracking of foreground objects.

$$\frac{\mathrm{P_B I(n)}}{\mathrm{P_F I(n)}} = \eta\,\frac{\mathrm{P(E{=}e^F)}}{\mathrm{P\,(E{=}e^B)}} \qquad (5)$$

Since E there is a priori probabilities on the right-hand side are Gibbs distributions characterized by the natural temperature γ cliques c and potential function V defined on. Choosing the cliques and potential function is crucial to the Gibbs model's effectiveness. We use 2 element cliques from the second-order, spatial Markov neighborhood (eight immediate neighbors), commonly used in image processing. An extension to higher neighborhood orders and more complex cliques is straightforward and increases spatial coherence of labels. The potential it should produce no penalty within a patch of identical label resulting in a high probability but it should penalize label fields with severe fragmentation. The potential is commonly used in such scenarios for its simplicity:

$$\mathrm{V(n,m)} = \begin{cases} 0, & \text{if } \mathrm{e(n)=e(m)} \\ 1, & \text{if } \mathrm{e(n)\neq e(m)} \end{cases} \qquad (6)$$

Clearly, the inclusion of the foreground model improves the performance over the fixed-threshold approach and over the MRF approach although only slightly. For example for a given miss rate (FNR) the foreground-model approach results in a lower false alarm rate (FPR) than the fixed-threshold procedure, while for a given false alarm rate it produces a slightly lower miss rate.

$$\frac{\mathrm{P_B I(n)}}{\mathrm{P_F\;I(n)}} = \Theta\;\exp\frac{1}{\gamma}\;\mathrm{Q_F(n)} - \mathrm{Q_B(n)} = \upsilon(\mathrm{n}) \qquad (7)$$

Where $\mathrm{Q_F(n)}$ four ground and $\mathrm{Q_B(n)}$ background neighbours of n, an adjustment of in the fixed-threshold approach cannot accomplish the same error rates as the foreground-model approach. The inclusion of foreground model results in a clear decrease of misses and a slight increase of false alarms compared to the fixed-threshold method. However, most of the false alarms and misses are corrected by the addition of Markov model which significantly improves the detection

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 5, May 2015

performance. Again, although gains due to the foreground model are modest, it is clear that the miss rate within the "truck" object is reduced when compared with Markov-only model shows results for video captured by network camera. Then combined foreground-MRF model significantly outperforms fixed thresholding as well as the joint model and produces accurate results shows iterative evolution of one object, shows the background probability as brightness level (top)). The complementarity of background and foreground probabilities (where is low, is high) leading to reinforced threshold adaptation, shows the corresponding evolution of the label field; since most gains occur in the first 2–3 iterations the process may be quickly terminated.

A similar convergence can be observed for the Markov model, the window size (W) used during classification of a pixel as stationary or non-stationary. A graphical variation among these three parameters is for the "Lobby" video sequence, may be observed that for and the achieved maximum of 99.47%. It may be observed that LIBS accurately detects objects in almost all cases with least misclassified objects. Then Moreover the shadows in "Intelligent Room" sequence are also removed by the proposed algorithm. Furthermore, object detection performance of LIBS scheme is superior to GMM and EGMM schemes; however it has similar performance with Reddy *et al.*'s scheme. But, LIBS scheme is computationally efficient compared to Reddy *et al.*'s scheme as the latter uses three cascading classifiers followed by a probabilistic voting scheme.

A Markov model within the binary hypothesis test while maintaining non parametric. Our approach extends early methods using single-Gaussian and uniform and shares Markovianity with more recent formulations. In is a two-Gaussian mixture and is uniform, we use more accurate nonparametric models. Also, despite the use of accelerated simulated annealing in, the computational complexity is high. Although one can seek local minima by means of one-at-a-time search, such as the iterated conditional modes algorithm, in this case a binary MAP-MRF solution is identical to our binary hypothesis test. Also, our approach uses spatial ergodicity whereas the one is based on temporal ergodicity which necessitates slow motion or tracking of foreground object.

IV] PERFORMANCE ANALYSIS

Clearly, the inclusion of the foreground model improves the performance over the fixed-threshold although only slightly in each case. For example for a given miss rate (FNR) the foreground-model approach results in a lower false alarm rate (FPR) than the fixed-threshold procedure while for a given false alarm rate it produces a slightly lower miss rate. Thus an adjustment of in the fixed-threshold approach cannot accomplish the same error rates as the foreground-model approach. For a given the inclusion of foreground model results in a clear decrease of misses and a slight increase of false alarms compared to the fixed-threshold method. However most of the false alarms and misses are corrected by the addition of Markov model which significantly improves the detection performance. Again although gains due to the foreground model are modest, it is clear that the miss rate within the "truck" object is reduced when compared with Markov-only bottom row shows results for video captured by network camera.

The combined foreground-MRF model significantly outperforms fixed thresholding as well as the joint model P_B/P_F and produces accurate results. The iterative evolution of one object shows the background probability as brightness level P_B(top) followed by the evolution P_{Fi}.

The complementarities of background and foreground probabilities leading to reinforced threshold adaptation. The corresponding evolution of the label field; since most gains occur in the first 2–3 iterations the process may be quickly terminated.

V] CONCLUSION

In this we have tested the proposed models on other surveillance videos and confirmed the gains reported here. A Background subtraction based on these models currently serves as our main change detection algorithm in research on visual behavior analysis and classification. However the inclusion of a foreground model tends to grow the detected regions rather than shrink them. It is thus critical that the initial label field have as few false positives as possible. This can be accomplished via false discovery rate control that uses thresholding of significance scores instead of probabilities.

REFERENCES

[1] T.Aach and A. Kaup "Bayesian algorithms for adaptive change detection in image sequences using Markov random fields," Signal Process, Image Commun., vol. 7, pp. 147–160, 1995.

[2] C. Stauffer and E. Grimson, "Learning patterns of activity using realtime tracking," IEEE Trans. Pattern Anal. Machine Intell., vol. 22, no. 8, pp. 747–757, 2000.

[3] A. Elgammal, R. Duraiswami, D. Harwood and L. Davis, "Background and foreground modeling using nonparametric kernel density for visual surveillance," Proc. IEEE, vol. 90, pp. 1151–1163, 2002.

[4] W. Huang, I. Gu and Q. Tian, "Statistical modeling of complex backgrounds for foreground object detection," IEEE Trans. Image Process., vol. 13, no. 11, pp. 1459–1472, Nov. 2004.

[5] N. Paragios and V. Ramesh, "A MRF-based approach to real-time subway monitoring," in Proc. IEEE Conf. Computer Vision Pattern Recognition, 2001, pp. 58–65.

[6] J.Migdal and E. L. Grimson, "Background subtraction using Markov thresholds," in Proc. IEEE Workshop Motion and Video Computing, WACV/MOTIONS'05, 2005, vol. 2, pp. 58–65.

[7] Y. Sheikh and M. Shah, "Bayesian modeling of dynamic scenes for object detection," IEEE Trans. Pattern Anal. Machine Intell., vol. 27, no. 11, pp. 1778–1792, 2005.

[8] P.-M. Jodoin, M. Mignotte and J. Konrad, "Statistical background subtraction using spatial cues," IEEE Trans. Circuits Syst. Video Technol., vol. 17, pp. 1758–1763, Dec. 2007.

[9] J. McHugh, "Probabilistic Methods for Adaptive Background Subtraction," M.S. thesis, Boston Univ., Boston, MA, Jan. 2008.

[10] J. McHugh, J. Konrad, V. Saligrama, P.M. Jodoin and D. Castanon, "Motion detection with false discovery rate control," in Proc. IEEE Int. Conf. Image Processing. Oct. 2008.