International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 -0056
Volume: 02 Issue: 05 | Aug-2015 www.irjet.net p-ISSN: 2395-0072

Background Subtraction for Effective Object Detection Using GMM&LIBS

R.Pravallika [1], Mr.G.L.N.Murthy [2]

[1] *P.G. Student, Department of Electronics & Communication Engg., Lakireddy Balireddy College of Engg., Andhra Pradesh, India*

[2] *Associate Professor, Department of Electronics & Communication Engg., Lakireddy Balireddy College of Engg., Andhra Pradesh, India*

***Abstract*- Object detection and tracking is a challenging problem in numerous applications like video surveillance, human computer interaction, video indexing and retrieval. In the current paper, an intensity range based object detection scheme is proposed. The proposed object detection scheme consists of two steps: the first one models the background from initial few frames and the second one extracts the objects based on local thresholding. The strength of this 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.**

In this paper, Gaussian Mixture Model and Local Illumination Based Background Subtraction model are to be analyzed and compared using kappa coefficient parameter values for effective object detection.

Key words: Background subtraction, Background Modeling, video segmentation.

I. INTRODUCTION

Background subtraction is generally considered a lower level image processing task. The segmentation result of the background subtraction stage is then fed into some higher level application, which aims to understand something about the scene. One such application is that of identifying and classifying moving objects based on their appearance. For reliable classification, the object must be described in a compact and robust fashion. The description may include its color distribution, size, shape, and textural information. In order for the description to be characteristic of the true object, a reliable segmentation must be provided. Otherwise, errors in the detection stage will give rise to misrepresentation, which may result in misclassification.

Object detection and tracking in video is a challenging problem and has been extensively investigated in the past two decades. It has applications in numerous fields, such as video compression, video surveillance, human-computer interaction, video indexing and retrieval etc. 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, background subtraction, temporal differencing, active contour models, and generalized Hough transforms. In surveillance system, video sequences are generally obtained through static cameras and fixed background. 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 [1]. In most of the suggested schemes, the object detected is accompanied with misclassified foreground objects due to illumination variation or motion in the background. Moreover, shadows are falsely detected as foreground objects during object extraction. Presently,The human face of a different skin colored people have different skin color an additional step is carried out to remove these misclassified objects and shadows for effective object detection. To alleviate this problem, a simple but efficient object detection technique is proposed, which is invariant to change in illumination and motion in the background. The proposed approach also neutralizes the presence of shadows in detected objects.

The suggested background model initially determines the nature of each pixel as stationary or non-stationary and considers only the stationary pixels for background model formation. In the background model, for each pixel location a range of values are defined.

Subsequently, in object extraction phase our scheme employs a local threshold, unlike the use of global threshold in conventional schemes.

Much work has been carried out related to cuurent 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 [2]. Here each pixel is modelled separately by a mixture of three to five Gaussians. TheW4 model presented by Haritaoglu *et al.* is a simple and effective method [3]. It uses three values to represent each pixel in the background image namely, the minimum intensity, the maximum intensity, and the maximum intensity difference between consecutive frames of the training sequence. McHugh *et al.* proposed an adaptive thresholding technique by means of two statistical models [4].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 [5]. 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. Instead of segmenting a frame pixel-by-pixel, Reddy *et al.* used an overlapping block-by-block approach for detection of foreground objects [6]. The scheme passes the texture information of each block through three cascading classifiers to classify them as background or foreground.

Generally, shadow removal algorithms are employed after object detection. Salvador *et al.* developed a three step hypothesis based procedure to segment the shadows [7]. It assumes that shadow reduces the intensities followed by a complex hypothesis using the geometrical properties of shadows. At the pixel level, Gaussian mixture model (GMM) is used, whereas at the global level statistical features of the shadow is utilized.

II. METHODOLOGY

Initially the video to be tracked or monitored is converted into frames. The background model is then assumed and each frame is updated by using the pixel by pixel method and apply the Gaussian mixture model and LIBS technology to each frame.

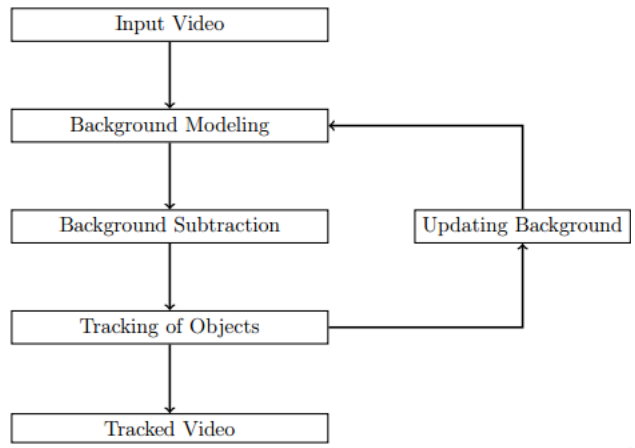


Fig: Moving object detection block diagram

A. GAUSSIAN MIXTURE MODEL

In practice, the illumination in the scene could change gradually (daytime or weather conditions in an outdoor scene) or suddenly (switching light in an indoor scene). A new object could be attended into the scene or a present object removed from it. In order to adapt to changes the training set can be updated by adding new samples and discarding the old ones. A reasonable time period T is chosen and at time t, $X_T = \{x^{(t)} \ldots\ldots\ldots\ldots, x^{(t-T)}\}$. For each new sample the training data set X_T is updated and re estimated it. However, among the samples from the recent history there could be some values that belong to the foreground objects and this estimate should be used as $\hat{p}(\vec{x}|X_T, BG+FG)$.In the current algorithm GMM with M components is used :

$$\hat{p}(\vec{x}|X_T, BG+FG) = \sum_{m=1}^{M} \hat{\pi}_m N(\vec{x};\hat{\vec{\mu}}_M, \hat{\vec{\sigma}}_m^2 I)$$

where $\hat{\vec{\mu}}_1 \ldots\ldots\ldots\ldots \hat{\vec{\mu}}_M$ are the estimates of the means and $\hat{\sigma}_1, \ldots\ldots\ldots\ldots, \hat{\sigma}_M$ are the estimates of the variances that described in the Gaussian components. The covariance matrices are assumed to be diagonal and the identity matrix I has peculiar dimensions. The mixing weights are denoted by $\hat{\pi}_m$ are non-negative and add up to one. Given a new data sample $\vec{x}^{(t)}$ at time t the recursive update equations are

$$\hat{\pi}_m \longleftarrow \hat{\pi}_m + \alpha(o_m^{(t)} - \hat{\pi}_m)$$

$$\hat{\vec{\mu}} \longleftarrow \hat{\vec{\mu}}_M + o_m^{(t)} (\alpha / \hat{\pi}_m) \vec{\delta}_m$$

$$\hat{\sigma}_m^2 \longleftarrow \hat{\sigma}_m^2 + o_m^{(t)} (\alpha/\hat{\pi}_m) (\vec{\delta}_m^{T} \delta_m - \vec{\hat{\sigma}}_m^2)$$

B. LIBS SCHEME

The LIBS scheme consists of two stages. The first stage deals with developing background model. This stage consists of two steps. First step is background model intilization. This step tries to classify each pixel as stationary or non-stationary in the frames required for background modeling. Next step of this stage is development of background model. Here a background model is developed considering stationary information of the pixel. In the second stage a local threshold based background subtraction method tries to find the objects by comparing any frame with the established background. Proposed scheme uses two parameters namely, window size W (an odd length window) and a constant C for its computation. The optimal values are selected experimentally. The stages and the parameter selection process of proposed scheme are described below in sequel.

The first frame or the combination of first few frames is considered as the background model. However, this model is susceptible to illumination variation, uneven lighting etc., and also to small changes in the background like waving of leaves. 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. In the proposed scheme few initial frames are considered for background modeling. Pixels in these frames are classified as stationary or non-stationary by analyzing their deviations from the mean.

1. *Development of Background Model*

The background is then modeled taking all the stationary pixels into account. The developed background model defines a range of values for each background pixel location around its true intensity. In development of background model algorithm, stationary pixels at any pixel location (i, j) in the frames form $f_{w/2}$ to $f_{n-[w/2]}$ are put into a vector
$\vec{R}$. Minimum and maximum value from it are determined and kept in two two-dimensional vector M(i, j) and N(i, j) respectively. The entire process is repeated for each pixel location in the frame. Finally, M(i, j) and N(i, j) will contain the minimum and maximum value of the stationary pixels from frames produced as output at respective pixel location (i, j). M(i, j) and N(i, j) represent the background model, defining a range of values for each background pixel location. From the above description it can be concluded that, for development of background model, frames having pixels as stationary or non-stationary are taken as input and at the end of the process min and max frames are produced in the form of background model as output.

2. *Extraction of Foreground Object*

After successfully developing the background model, a local thresholding based background subtraction is used to find the foreground objects. A constant C is considered that helps in computing the local lower threshold (T_L) and the local upper threshold (T_U) . These local thresholds help in successful detection of objects suppressing shadows if any.

Background subtraction algorithm takes the developed background model and a frame f as its input. It produces a segmented frame as its output consisting of foreground object if any with shadow suppressed. Algorithm is repeated for each location in the frame. At each pixel location threshold T(i, j) is calculated as T(i, j) = 1 C [M(i, j) + N(i, j)] where, C is a constant. Considering T(i, j) local thresholds are calculated as —

Local lower threshold: $T_L(i, j) = M(i, j) - T(i, j)$

Local upper threshold: $T_U(i, j) = N(i, j) + T(i, j)$

If f(i, j) value lies in between T_L and T_U, then it is a background pixel else a foreground pixel.

III. KAPPA COEFFICIENT

Cohen's kappa coefficient (κ) is a statistic which measures inter-rater agreement for qualitative (categorical) items. It is generally thought to be a more robust measure than simple percent agreement calculation, since κ takes into account the agreement occurring by chance. Some researchers have expressed concern over κ's tendency to take the observed category frequencies as given, which can have the effect of underestimating agreement for a category that is also commonly used; for this reason, κ is considered an overly conservative measure of agreement.

Group randomized trials are studies in which groups are assigned to treatments rather than individuals, but the units of observation are the members of those groups. The groups consist of members considered to have some social, geographic, or other connections rather than being composed of randomly selected individuals (Murray 1998). Analysis of data arising from group randomized trials imposes an inherent violation of the independence assumption required by traditional statistical techniques.

The groups are considered to be independent of each other, but the members within the groups may not be independent; observations on members within a group are likely to be correlated. The correlation between the members for each group need to be accounted for when analyzing group randomized trial data. Additionally, since the units of assignment are the clusters rather than the subjects within clusters, the degrees of freedom should be based on the number of clusters rather than the number of subjects.

Cohen's kappa measures the agreement between two raters who each classify *N* items into *C* mutually exclusive categories.

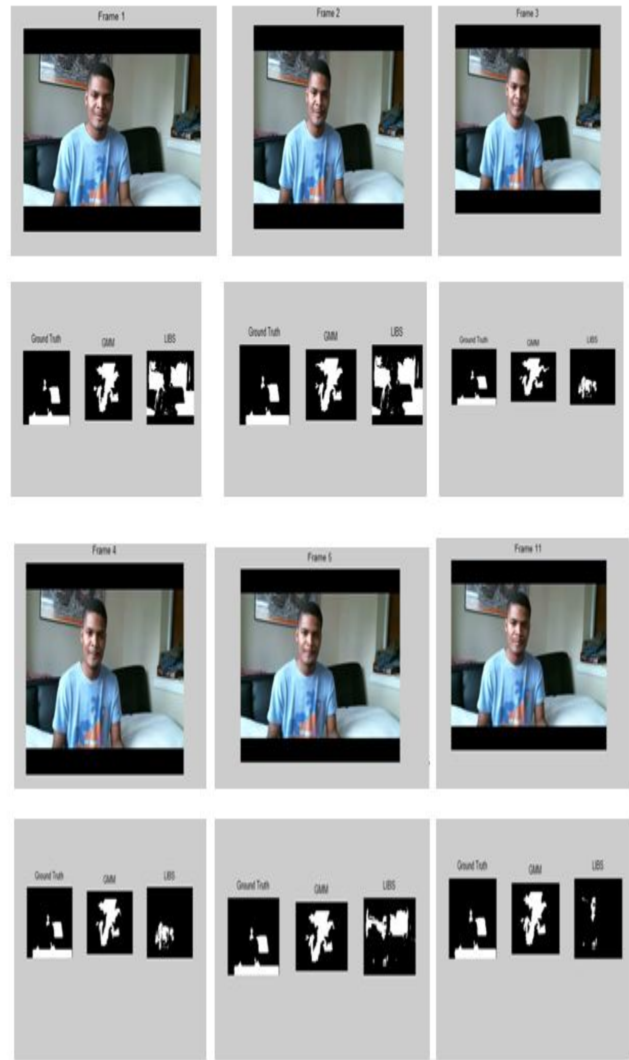
The output of this function is:

- Observed agreement percentage
- Random agreement percentage
- Agreement percentage due to true concordance
- Residual not random agreement percentage
- Cohen's kappa
- kappa error
- kappa confidence interval
- Maximum possible kappa
- k observed as proportion of maximum possible
- k benchmarks by Landis and Koch
- z test results

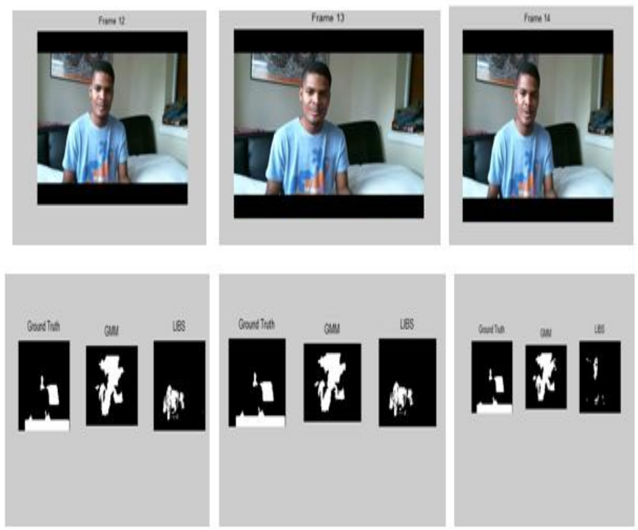
The equation for κ is:

$$K = Pr(a) - Pr(e) / 1 - Pr(e)$$

where $Pr(a)$ is the relative observed agreement among raters, and $Pr(e)$ is the hypothetical probability of chance agreement, using the observed data to calculate the probabilities of each observer randomly saying each category. If the raters are in complete agreement then κ = 1. If there is no agreement among the raters other than what would be expected by chance (as defined by $Pr(e)$), κ = 0.



IV. RESULTS AND DISCUSSION



Finally, entropy and PSNR values are calculated for the last frame which is shown in the figure below.

	MSE	PSNR	Entropy	Mean	Std
Ground Truth	0.0328	52.9735	65	0.1635	0.3260
GMM	0.0202	55.0834	72	0.2034	0.2350
LIBS	0.0026	67.0580	99	0.1537	0.0621

UNWEIGHTED COHEN'S KAPPA:

Observed agreement (po) = 0.0036

Random agreement (pe) = 0.0040

Agreement due to true concordance (po-pe) = -0.0004

Residual not random agreement (1-pe) = 0.9960

Cohen's kappa = -0.0004

kappa error = 0.0000

kappa C.I. (alpha = 0.0500) = -0.0004 -0.0003

Maximum possible kappa, given the observed marginal frequencies = 0.7637

k observed as proportion of maximum possible = -0.0005

Poor agreement

Variance = 0.0000

z (k/sqrt(var)) = -14.9913

p = 0.0000

V. CONCLUSION

Object tracking is an important computer vision application which consists of two closely related processes; object detection and tracking of the detected objects. Object detection in videos obtained from static camera and fixed background is achieved through background subtraction approach. In this approach a background model is developed considering the first frame or first few frames. Subsequently, a thresholding technique is utilized to extract foreground objects. Shadows are very often misclassified as foreground objects, which needs an additional step to remove before the detected objects can be tracked. Object tracks are computed by various approaches. Centroid in subsequent frames are searched in a fixed size window, which makes the algorithm more complex. Inorder to accommodate changes in the background scene, updating background model plays a vital role. Frequent and entire updating of the background model makes the method computationally inefficient.

In this paper, Proposed detection scheme starts with considering first few frames of the video. Pixels in these frames are classified as stationary or non-stationary according to their intensity along temporal axis. Considering the stationary pixel information, background model is developed. A local thresholding technique tries to extract foreground objects and suppresses shadows at low computational cost. Comparative analysis demonstrates the efficacy of the proposed detection scheme. a simple but robust scheme of background modelling and local threshold based object detection are proposed. We are consider a video corresponding GMM and LIBS figures are mentioned above section. Videos having variant illumination background, textured background, and low motion background are considered for simulation.GMM and LIBS schemes are compared with the kappa coefficient parameter both qualitatively and quantitatively.

International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 -0056
Volume: 02 Issue: 05 | Aug-2015 www.irjet.net p-ISSN: 2395-0072

REFERENCES

1. A. Yilmaz, O. Javed, and M. Shah, "Object tracking: A survey," ACM Comput. Surv., vol. 38, Dec. 2006.
2. C.Wren, A. Azarbayejani, T. Darrell, and A. Pentland, "Pfinder: Realtime tracking of the human body," IEEE Trans. Patt. Anal. Mach. Intell.,vol. 19, no. 7, pp. 780–785, Jul. 1997.
3. I. Haritaoglu, D. Harwood, and L. Davis, "W4: Real-time surveillance of people and their activities," IEEE Trans. Patt. Anal. Mach. Intell., vol. 22, no. 8, pp. 809–830, Aug. 2000.
4. J.McHugh, J. Konrad,V. Saligrama, and P. Jodoin, "Foreground-adaptive background subtraction," IEEE Signal Process.Letters, vol. 16, no. 5, pp. 390–393, May 2009.
5. O. Barnich and M. Van Droogenbroeck, "ViBe: A universal background subtraction algorithm for video sequences," IEEE Trans. Image Process., vol. 20, no. 6, pp. 1709–1724, Jun. 2011.
6. V. Reddy, C. Sanderson, and B. Lovell, "Improved foreground detection via block-based classifier cascade with probabilistic decision integration," IEEE Trans. Circuits Syst. Video Technol., vol. 23, no. 1, pp.83–93, Jan. 2013.
7. E. Salvador, A. Cavallaro, and T. Ebrahimi, "Cast shadow segmentation using invariant color features," Comput. Vis. Image Understand., vol. 95, no. 2, pp. 238–259, 2004.

K.Tejasri has received her B.Tech from JNTU Kakinada with Electronics & Communications Engineering Specialization and at present she is pursuing M.Tech from JNTU Kakinada University, AP,India in the area of Systems and Signal Processing .She has actively involving in college level activities and technical symposiums.

Mr. G.L.N.MURTHY has received his Master's Degree from JNTU Ananthapur in the Discipline of Digital Systems & compouter .He is pursuing his Ph.D in ECE.He has been actively involved in teaching and research for the past 13 Years and he is working as Assoc Professor, in the Dept of ECE,Lakireddy Balireddy College Of Engineering, Mylavaram. AP, India