International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-8 Issue-4, April 2019

Model Based Approach for Effective Segmentation of Images Based On Background Subtraction

Pavan Kumar Tadiparthi, Srinivas Yerramalle

Abstract: Image segmentation is considered as a vital part of image analysis which better understanding of these images is possible. Among the algorithms available to analyze images under motion, background subtraction is considered to be the most imperative. In this article an attempt is made to propose a methodology of image segmentation based on background subtraction by a proposing and developing a model based on truncated Gaussian distribution. The experimentation is carried on CDnet 2014 data set and results are analyzed using the metrics.

Index terms: Image segmentation; Background subtraction; Truncated Gaussian distribution; performance metric; Bench mark images.

I. INTRODUCTION

With the latest updates in the area of technology, latest innovations have been populated for better understanding in particular regarding the visualization of images. As the technology is advancing, sophisticated mechanisms/tools are to be developed for better understanding of the images and also in practical applications related to security. In order to identify the final details regarding the acquired images, process under visualization cameras many methodologies have been evolved based on region based techniques[1], Edge based techniques[2], contour based techniques[3], graph cuts[4], artificial neural networks[5], optimization techniques [6], among this models majority of the models aim at identifying the background and subtracting the background information from the foreground to have a more detail information regarding the images. In these approaches, the objects of interest are identified based on the features foresaid. However, in practicality, the images are well interpreted by understanding the pattern of pixels and with this assumption approaches are to be developed basing on analyzing the objects and therefore better understanding about the objects can be possible. Very little work in this direction as been carried out using Gaussian mixture models [7]. However, in reality most of the images are depicted with in the finite regions and therefore, we consideration of the normal range proposed by Gaussian mixture model (GMM) that is –α toα is hardly necessary. Therefore it is necessary to truncated this limits to a finite boundaries [8]. The boundaries are estimated as A and B, where A is the minimum pixel intensity value and B denotes the maximal pixel intensity value of the image considered and rest of the pixels are assumed to within this limits.

Manuscript published on 30 April 2019.
***** Correspondence Author (s)

Pavan Kumar Tadiparthi, Associate Professor, Department of Information Technology, MVGR College of Engineering, India

Srinivas Yerramalle, Professor, Department of Information Technology, GIT, GITAM University, India.

Therefore, in this article a truncated Gaussian mixture model is proposed for effective image segmentation. The rest of the article is articulated as follows. Section II of the article introduces the Truncated GMM together with the parameters to be updated based on E.M algorithm. The need for truncation is also proposed. Section III highlights the data set considered. In section IV the methodology is proposed .section V of the article presents the experimentation process and results derived there are. The outputs derived are subjected to performance metrics and the evaluation processes together with mathematical formulas are highlighted in corresponding the section VI, and the article is summarized in the concluding section VII.

II. TRUNCATED GAUSSIAN MIXTURE MODEL (TGMM)

Of late much prominence is given for analyzing the static objects from the video sequences, through Finite Normal Mixture Models (FNMM). In FNMM, each object is exemplified by a Gaussian Distribution (G.D)and the whole object is well thought-out to be a blend of these G.D, i.e., they consider the probability density function of the object acquired from a video frame follow a FNM Distribution. The range of objects collected within these static frames is considered to be finite. For this reason approximating the unlimited range (- ∞ to +∞) to an object using Finite Normal Mixture Models is apparently not sensible. As a result, it is desirable to visualize that each object portrays a Truncated Gaussian Distribution and the objects within the entire frame follow a **TGMM**. In assortment models, the imperative issue is the number of components 'K' (district groups). Typically the numbers of distinct groups are implicitly known as prior in EM algorithm, and these groups need to be optimized for which E M methodology is considered. The difficulty with E.M algorithm for object detection is w.r.t. the identification of primary values of the estimates within the objects. The regularly practiced method is based on Bayesian priori and by the usage of the concepts of random sampling. This method succumbs to excellent outcomes when the chosen number is huge, also the time consumed will be heavy. When the selected number is undersized, it is extremely probable that some minute groups within the objects may not be sampled, which influences the segmentation accuracy.

II. PROBABILITY DENSITY FUNCTION OF TGMM

In low-level object analysis, the complete object dataset is understood to be following a **TGMM**.

To model, it is habitual to presuppose that each object within the segmented frame follows a Normal (Gaussian) Distribution. The Probability Density function (P.D.F) is given by,

$$f(Z)=\frac{1}{\sqrt{2\pi}\sigma}\, e^{-\frac{1}{2}\left(\frac{z-\mu}{\sigma}\right)^{2}} \;;\ -\infty < E<+\infty - \infty< \mu <\infty;\ 0< \sigma \qquad \text{-- (1)}$$

The value of 'E'(the no of times particular word is occurring) below some value E_L and above some value E_M cannot be possible to exit. Then the resulting distribution of the word from the object corpus is a Truncated Normal Distribution, with the probability distribution function

$$g(z)=\begin{cases} & -\infty\le Z<\infty \\ \dfrac{f(z)}{\int_{Z_L}^{Z_M} f(z)dz} & Z_L \le Z \le Z_M \\ & Z_M \le Z<\infty \end{cases} \qquad \text{-- (2)}$$

This implies $g(z)=\frac{f(z)}{B-A},\ z_L<z<z_M$

Where f(E) is as defined in equation (1)

$$A=\int_{-\infty}^{Z_L}\frac{e^{\left(\frac{-1}{2}\left(\frac{z-\mu}{\sigma}\right)^{2}\right)}}{\sqrt{2\pi\sigma}}dz \ \ and\ \ B=\int_{-\infty}^{Z_M}\frac{e^{\left(\frac{-1}{2}\left(\frac{z-\mu}{\sigma}\right)^{2}\right)}}{\sqrt{2\pi\sigma}}dz \qquad \text{-- (3)}$$

The lower and upper truncation points are E_L and E_M respectively. The degrees of truncation are (A) and (1-B). Where A and B are called truncating limits. In this method the values of A and B are chosen such that A is the minimum probability obtained from Term Frequency Independent Data Frequency (TFIDF) and B is the maximum TFIDF for evaluating in the objects under investigation. The term frequencies are evaluated and from which the minimal and maximum values of the frequency of terms indicated by A and B Values.

If E_L is replaced by -∞, or E_M by ∞, the distribution is singly truncated from above, or below, respectively.

These are classified according to the degrees of truncation. It can be seen that when the truncations are large, the distribution bears little resemblance to a normal distribution. The case E_L =μ and E_M = ∞ produces a half-normal distribution.

The mean no of times particular word is occurring of the i^th^ region is

$$E(z)=\mu_i+\frac{\sigma_i^{2}\left[f(Z_L)-f(Z_M)\right]}{A-B} \qquad \text{-- (4)}$$

The Variance of the word intensities in the i^th^ object is

$$V(Z)=\left[1+\frac{\left[\left(\frac{Z_L-\mu_i}{\sigma_i}\right)Z_L-\left(\frac{Z_L-\mu_i}{\sigma_i}\right)Z_M\right]}{B-A}\right]\sigma_i^{2} \qquad \text{-- (5)}$$

The cumulative distributive function is G(E)

$$G(z)=\frac{F(Z)-F(z_L)}{B-A}\quad for\ z_L\le z\le z_M \qquad \text{-- (6)}$$

Where F(E) is the cumulative density function associated with the normal variant Z .

The coefficient of variation of the word in the ith object is

$$C=\frac{\sqrt{Var(Z)}}{E(Z)} \qquad \text{-- (7)}$$

The moments of a doubly truncated normal variant are

$$m_o=\frac{1}{2}\left[\exp\left(\frac{z_M-\mu}{\sigma\sqrt{2}}\right)-\exp\left(\frac{1-\mu}{\sigma\sqrt{2}}\right)\right]$$

$$m_1=\frac{\sigma}{\sqrt{2\pi}}\left[\exp\left(-\frac{z_L-\mu}{\sigma\sqrt{2}}\right)^{2}-\exp\left(-\frac{z_M-\mu}{\sigma\sqrt{2}}\right)^{2}\right]+\mu m_o \qquad \text{-- (8)}$$

Since the entire object is a collection of regions which are characterized by doubly truncated normal variants. We assume that the word from the object corpus follows a K-component finite mixture of Truncated Gaussian Distribution and its probability density is of the form, where K is the number of words in a object, α_i >0 are weights such that $\sum_{i=1}^{K}\alpha_i=1$,and

$$g_i(Z\mid\mu_i,\sigma_i^{2})=\frac{e^{\left(\frac{-1}{2}\left(\frac{z-\mu_i}{\sigma_i}\right)^{2}\right)}}{\sqrt{2\pi\sigma}(A-B)}\quad -z_L<z<z_M,\ 0<\sigma_i,\ z_L<\mu_i<z_M \qquad \text{--(9)}$$

μ_i,σ_i^2 and $g_i(z/\theta_i)$ are the mean, variance and probability density function of the pixels in the i^th^ object respectively

α_i is the probability of occurrence of the i^th^ component of the Truncated Gaussian Mixture Model i.e., the probability of the i^th^ object. Generally it can be taken as the ratio of the size of the i^th^ object region to the size of the entire object data, such that $\sum_{i=1}^{K}\alpha_i=1$.

The likelihood function of sample observations $z_1, z_2, \ldots z_N$, drawn from an object with P.D.F $h(z,\theta)$ = $\sum\alpha_i g_i(z_s,\theta)$ is given by

$$L(\theta)=\pi_{s=1}^{N}\left(\sum_{l=1}^{K}\alpha_i g_i(z_s,\theta)\right) \qquad \text{-(10)}$$

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication (BEIESP)

$= \prod_{s=1}^{N}(\sum_{l=1}^{K}\alpha_i \frac{e^{\left(\frac{-1}{2}\left(\frac{(z-\mu_i)}{\sigma_i}\right)^2\right)}}{\sqrt{2\pi\sigma}(A-B)}$, where A & B are as given in equation (3)

This implies

$$\text{Log L}(\theta) = \log(\sum_{k=1}^{K}\alpha_i g_i(z_s,\theta))$$

$$= \sum_{i=1}^{N} Log\,(\sum_{i=1}^{K}\alpha_i g_i(z_s,\theta)) \quad -(11)$$

The first step of the EM algorithm requires the estimation of some reasonable initial estimates for both parameters $\theta^{(0)}$ and component weights $\alpha^{(0)}$ from the observed sample. The idea of the EM algorithm is then to iteratively calculate maximum likelihood estimate of the unknown parameters θ.

III. DATASET

In order to experiment the proposed methodology, in this article, we have considered a bench mark data set of video images from CDnet 2014.The data set contains 6 different video categories with a total of 31 videos comprising of 80,000 frames. These video are categorized like Baseline, Camera jitter, Dynamic background, shadow, thermal, intermittent object motion.

IV. METHODOLOGY

To identify the background more effectively the following steps are to be considered

1. *Post processing*: In order to extract the background image, it is necessary to initially process the input objects such that the deviations among the object is low. These pixels with low disparity are considered to be background pixels else considered as a foreground. However, they are similar deviations may be accorded to both foreground and background inforamtions.so we need to have a clear distinguish illumination should to be verified and the proper illumination helps to interpret the background images more apparently.

2. *Adaptive background subtraction:*

Here, we have best possible threshold value can be estimated using the adaptive background thresholding technique. Here, we have estimate the difference in values between two consecutive frames 't'and 't-1'.one of the frame is considered as current frame and other one is assumed to be the reference frame.

The frame is represented as

F(x, y) =C(x, y)-R(x, y)

F(x, y) =1 if f(x, y) >=T -(12)

(Where T is the threshold value, (using the methodology proposed by N.Otsu [21]) and zero otherwise. Here C(x, y) denotes the current frame and R(x, y) represents the considered reference background image, F(x, y) denotes the deviation between the present frame and reference frame.

The pixels with high threshold values are given as input to the Truncated Gaussian mixture Model. The probability density functions (pdf) against each of the intensity values are given as input to the model and the respective values are estimated. These values which are below the threshold value considered as background information else they are considered as foreground information.

V. EXPERIMENTATION

In order to experimentation the proposed model based on the truncated Gaussian mixture model presented in section II of the article. Each of the considered images and the static frames are obtained. The background objects of the video sequences are identified by the adaptive background subtraction method presented in section IV. of the article. Each of the videos is subsequently processed according to the process presented in section IV. In general, background pixels are considered to have a minimum threshold value. Pixels having high threshold values are provided as input to the truncated Gaussian mixing model (TGMM) presented in section II of the article and based on the composite probability of the probabilities, each pixel is classified as a background pixel or a pixel close-up

The complete experimentation carried out in a mat laboratory environment in the data set described in section III of the article and the results obtained using the performance evaluation methods presented in section VI. The results were also compared with the model based on the Gaussian Mixture Model (GMM).

VI. PERFORMANCE EVALUATION AND EXPERIMENTAL RESULTS

To evaluate the performance of the proposed modeling method, we have considered different quality measurement metrics are Accuracy, Recovery, Accuracy, F Score, MSE, RMSE, FNR, FPR, PSNR, PWC. Recovery expressed in terms of the exact number of foreground pixels that are classified as foreground pixels. The precision is expressed in terms of the number of exact foreground pixels against the assigned foreground pixels; The performance of the developed model validated by the value of the calculated precision, if it is high, it represents a high performance. On the other hand, if the proposed method assigns most of the pixels to the background, the output precision value may be high but proportionally the value of the recovery drops.

Performance metrics are expressed as

Precision=TP/ (TP+FP) - (13)

Recall=TP/ (TP+FN) - (14)

Accuracy=TP+TN/(TP+TN+FP+FN) -(15)

F-score = $\frac{(2*Precision*recall)}{(Precision +recall)}$ - (16)

MSE=FP+FN/M*N - (17)

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication (BEIESP)

$RMSE=MSE$ - (18)

$FNR=FN/(TP+FN)$ - (19)

$FPR=FP/(FP+TN)$ - (20)

$PSNR=10\log_{10}(R^2/MSE)$ - (21)

$$PWC=\frac{100*(FN+FP)}{(FN+TN+FP+TP)}$$ - (22)

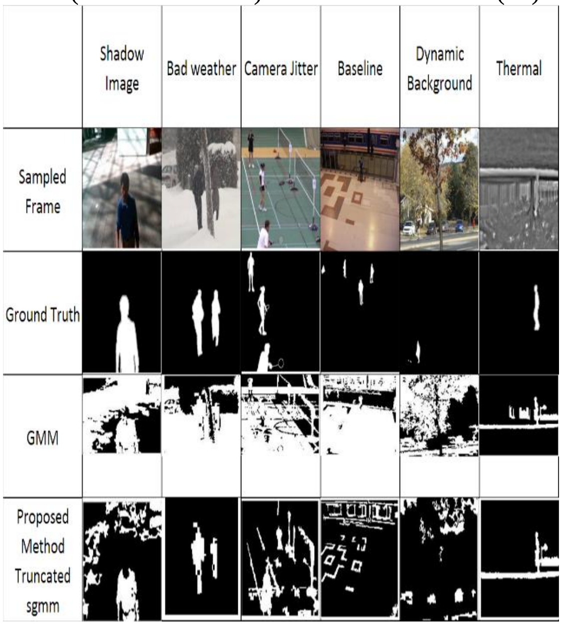

Fig1: Foreground detection of thermal, camerajitter,Dymanicbackground,Shadow,badweather,baseline from the CDNet2014 Dataset.

TP: number of foreground pixels classified as foreground, FN: number of foreground pixels classified as Background. FP: number of pixels of background pixels classified as foreground.TN: number of background pixels classified as background.

The experimentation is carried out with the proposed model considering the set of reference data CDnet2014 presented in section III. The results derived from the considered data are presented in the following figure .1.

We evaluate the proposed method that is analyzed in section II. The scenarios used to evaluate the thermal model, shadow, fluctuation of the camera, dynamic background. In each scenario there are many videos. We select a typical framework for each video. Figures 1. (a) to (f) are selected from six categories in the CDNet 2014 data set. Figure 1 (1) shows the original picture of the video and Figure 1. (2) are the results of the data of the fundamental truth. Fig.1 (3) - (4) are the object detection results of the background modeling methods of the state of the art. Tables I to IV present 10 metrics of quality performance of the proposed modeling method in the 2014 CDNET data set.

The effectiveness of the proposed method can be confirmed by withdrawal, accuracy, F-score, accuracy and other metrics. For each evaluation metric, the results of the model proposed in different scenes through Tables 1 to 6

TABLE.1

Evaluation Metrics of different methods on SHADOW video from CD net DATASET		
Metrics\ Methods	GMM	TRUNCATED GMM
PRECISION	0.0139	0.0178
RECALL	0.135	0.0538
ACCURACY	0.9474	0.9832
F-SCORE	0.0193	0.0495
MSE	0.0155	0.03
RMSE	0.1302	0.1516
FPR	0.0737	0.0379
FNR	0.832	0.932
PSNR	67.5661	76.917
PWC	5.2785	3.5925

TABLE.2

Evaluation Metrics of different methods on BADWEATHER video from CD net DATASET		
Metrics\ Methods	GMM	TRUNCATED GMM
PRECISION	0.0438	0.0806
RECALL	0.0583	0.0452
ACCURACY	0.9425	0.9843
F-SCORE	0.0597	0.0464
MSE	0.0213	0.0304
RMSE	0.1248	0.1535
FPR	0.0578	0.0459
FNR	0.705	0.8329
PSNR	66.0952	70.6053
PWC	5.7813	4.6856

TABLE.3

Evaluation Metrics of different methods on CAMERA JITTER video from CD net DATASET		
Metrics\ Methods	GMM	TRUNCATED GMM
PRECISION	0.032	0.0516
RECALL	0.032	0.0218
ACCURACY	0.9456	0.9863
F-SCORE	0.026	0.0532
MSE	0.0131	0.0789
RMSE	0.1156	0.2886
FPR	0.0536	0.1603
FNR	0.056	0.9618
PSNR	65.683	73.6308
PWC	5.5255	3.0238

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication (BEIESP)

TABLE.4

Evaluation Metrics of different methods on BASELINE video from CD net DATASET		
Metrics\ Methods	GMM	TRUNCATED GMM
PRECISION	0.0071	0.0136
RECALL	0.0241	0.0767
ACCURACY	0.9423	0.9653
F-SCORE	0.0161	0.0139
MSE	0.0026	0.0153
RMSE	0.053	0.1372
FPR	0.0256	0.0145
FNR	0.7247	0.9925
PSNR	67.0696	72.5137
PWC	3.5326	2.252

TABLE.5

Evaluation Metrics of different methods on DYNAMIC BACK GROUND video from CD net DATASET		
Metrics\ Methods	GMM	TRUNCATED SGMM
PRECISION	0.0021	0.0152
RECALL	0.0237	0.0523
ACCURACY	0.9676	0.9832
F-SCORE	0.0132	0.0178
MSE	0.0135	0.0183
RMSE	0.1116	0.1273
FPR	0.0432	0.0255
FNR	0.7356	0.8523
PSNR	69.2044	67.3431
PWC	5.2423	3.4567

TABLE.6

Evaluation Metrics of different methods on THERMAL video from CD net DATASET		
Metrics\ Methods	GMM	TRUNCATED SGMM
PRECISION	0.0131	0.0343
RECALL	0.0324	0.0835
ACCURACY	0.9765	0.9956
F-SCORE	0.0132	0.0154
MSE	0.0086	0.0065
RMSE	0.0865	0.076
FPR	0.0116	0.0102
FNR	0.5578	0.9832
PSNR	70.3876	71.1748
PWC	3.1932	2.104

VII CONCLUSION

In this article, a model proposed for the effective background subtraction based on the Truncated Gaussian mixture model is presented. The proposed model is compared with that of the existing model based on GMM using the metrics like precision, recall, FPR, FNR, F-Score, and Accuracy etc. The results derived are valued against assessment metrics. The results are tabulated and presented in table- I to table -IV.and graphs 2-61.From the graphs it can be clearly observed that the proposed models perform well with respect to all the considered metrics, when compared to that of models based on the GMM, it clearly showcases better performance accuracy.

REFERENCES

1. Shi J and Malik.J, “Normalized cuts and image segmentation,” IEEE TransMachine Intell, 2000, vol.22, pp.888–905.
2. Otsu N, Discriminant and Least Square Threshold Selection. In: Proc 4IJCPR, 1978, pp.592–596.
3. Zarrar Javaid, Charles P. Unsworth, Mark Boocock, Peter McNair, "Fast identification & modelling of osteoarthritic holes in the human knee with contour interpolated radial basis functions", Engineering in Medicine and Biology Society (EMBC) 2013 35th Annual International Conference of the IEEE, pp. 4438-4441, 2013.
4. D. Freedman, M.W. Turek, "Illumination-invariant tracking via graph cuts", Computer Vision and Pattern Recognition 2005. CVPR 2005. IEEE Computer Society Conference on, vol. 2, pp. 10-17 vol. 2, 2005.
5. S. U. Indira, A. C. Ramesh, "Image Segmentation Using Artificial Neural Network and Genetic Algorithm: A Comparative Analysis", Process Automation Control and Computing (PACC) 2011 International Conference on, pp. 1-6, 2011.
6. P. Suganthan, “Particle Swarm Optimizer with Neighborhood Optimizer,” Proceedings of the Congress on Evolutionary Computation, pp. 1958-1962, 1999.
7. C. Stauffer and W. Grimson. Adaptive background mixture models for real-time tracking. in IEEE Proc. CVPR, 1999.
8. Hausman, J. A., and D. A. Wise (1976): “The evaluation of results from truncated samples: The New Jersey negative income tax experiment,” Annals of Economic and Social Measurement, 5, 421–445.
9. Nagesh Vadaparthi, Srinivas Yerramalle, and Suresh Varma.P: An Efficient Approach for Medical Image Segmentation Based on Truncated Skew Gaussian Mixture Model Using K-Means Algorithm”, International Journal of Computer Science and Telecommunications, (ISSN 2047-3338), 2(6):79-86, August 2011.
10. Conrad Sanderson, Kuldip K. Paliwal, “Fast features for face Recognition under illumination direction changes”, in Pattern Recognition Letters Vol 24, No.14, pp.2409-2419, 2003
11. J. E Canny, “A computational approach to edge detection,” IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 8, pp. 679-698, 1986
12. G. Heitz and D. Koller. Learning spatial context: Using stuff to find things. In ECCV, 2008
13. Y. L. Chang and X. Li, “Adaptive image region-growing,” IEEE Trans. on Image Processing, vol. 3, no. 6, pp. 868-872, 1994
14. Haralick, Shapiro: Survey: Image segmentation Techniques. In: Proc. of Int. Conf. CVGIP 1985, vol. 29, pp. 100–132 (1985)
15. Pal, S.K., Pal, N.R.: A Review on Image Segmentation Techniques. Pattern Recognition 26(9), 1277–1294 (1993)
16. Haritha D., Srinivasa Rao K. and Satyanarayana Ch. “Face recognition algorithm based on doubly truncated Gaussian mixture model using hierarchical clustering algorithm coefficients”, International journal of Computer science issues. 9(2), 388-395 (2012)
17. C. Gu, J. J. Lim, P. Arbelaez, and J. Malik. Recognition using regions. In CVPR, 2009.
18. Wang Kejun, Ding Yuhang, Zhuang Dayan and Wang Dazhen, “Threshold segmentation for hand vein image,” Control Theory and Application, Beijing: Publishing House of Electronics Industry, 2005, pp.19–22.
19. Kittler J. and Illingworth J., “On threshold selection using clustering.
20. S.Bileschi and L. Wolf. A unified system for object detection, texture recognition, and context analysis based on the standard model feature set. In BMVC, 2005.
21. L. Brown, A. Senior, Y. Tian, J. Vonnel, A. Hampapur, C. Shu, H. Merkl, and M. Lu. Performance evaluation of surveillance systems under varying conditions. in IEEE Proc. PETS Workshop, 2005.
22. N.otsu,”A threshold selection on method from gray level histograms”.Automatica ,1975, vol.11, pg. 23-27.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication (BEIESP)