Towards the Development of Effective Video Segmentation Based on Skew Gaussian Mixture Model

Pavan Kumar Tadiparthi, Srinivas Yarramalle

Abstract: video analysis has gained an exponential demand with its usage in security cameras and in most of the real time applications for monitoring the law order. In order to have a precise analysis background subtraction and foreground detection processed are generally considered in the most of the approaches. However to have a more precise output from the dynamic motion images, this article proposes a methodology based on skew Gaussian mixture model. The results are analyzed against the existing methods using quality assessment measures.

Index terms: Performance analysis; image segmentation; skew Gaussian; Background subtraction; quality metrics.

I. INTRODUCTION

In most of the image analysis techniques highlighted in case of video sequences, background subtraction is considered to be a crucial element. This Background subtraction helps to identify the reference background and there by identify the movement objects using the disparity among the background an input image. Many methodologies uses the concept of both model and non-model based approaches of which majority of the articles are mostly focusing on extracting the regions, contours and also some recent articles have been proposed by the author in this direction using Generalized Rayleigh Distribution [1]. The above article highlights about a methodology for effective identification of foreground images using the GRD methodology proposed. In some practical situations, the background methodologies used for subtraction from the foreground considers in particular cases of a temporarily halted objects and try to process the foreground detection. The results in these cases may not yield better results.

The most insightful methodology to consider the background is to take into account the final details of the objects during time intervals ‘t’ and ‘t-1’. This consideration helps to generate better understanding about the image to be segmented also.

In some practical situations are may came across the images that are skewed and having highest kurtosis and these sort of images have a limitation for segmenting based on Generalized Rayleigh distribution (GRD).

Therefore, to counter attack such variations in kurtosis, the present methodology on skew Gaussian is considered.

The main limitation with respect to the Generalized Rayleigh Distribution (GRD) is that its considerations of low range sale factor ‘α’ which is mostly equal to 0.5 and any image having scale above this can’t be well interpreted. In contrary, the Skew Gaussian can consider the ranges up to plus or minus 3 and hence better perception can be obtained using this method. Therefore the present article aims at proposing the image segmentation methodology based on skew Gaussian.

The rest of the article is highlighted as follows, In section II, a brief overview of Skew Gaussian is presented, and section III highlights the considered data set. In section IV of the article presence the methodology, section V proposes the experimentation and Results derived together with performance is presented is corresponding section VI. The final section VII summarization the article with conclusion.

II. FINITE SKEW NORMAL MIXTURE DISTRIBUTION

Every image is a collection of several image regions. In each image region, the image data is quantized by pixel, which is a random variable because of the fact it is influenced by random factors like Vision, brightness, contrast etc. To model the pixel intensities in a image region, it is necessary to assume that the pixels in each image region follow a skew normal distribution. The probability density function of the pixel intensity is given by [Azzalini A. (1985), Tsung I Lin et al (2007)]

\[ f(z) = \Phi(z), \Phi(\alpha z) \quad -\infty < z < \infty \]

where, \( \Phi(\alpha z) = \int_{-\infty}^{\alpha z} \Phi(t) \ dt \)

and, \( \Phi(z) = \frac{e^{-\frac{1}{2}z^2}}{\sqrt{2\pi}} \)

Let, \( y = \mu + \sigma z \)

\[ z = \frac{y - \mu}{\sigma} \]

Substituting equations (2), (3), and (4) in equation (1),

\[ f(y) = \frac{2}{\sqrt{\pi}} e^{-\frac{1}{2}z^2} \left[ \int_{-\infty}^{\frac{z}{\sqrt{2}\sigma}} e^{-\frac{1}{2}t^2} \ dt \right] \]

Revised Manuscript Received on February 08, 2020.

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

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

Published By: Blue Eyes Intelligence Engineering & Sciences Publication

DOI: 10.35940/ijeat.C5712.029220
Towards the Development of Effective Video Segmentation Based on Skew Gaussian Mixture Model

The mean pixel intensity of an image region is

$$E(z) = \mu + \sqrt{\frac{2}{\pi}} \cdot \delta(y)$$

Where,

$$\delta(y) = \frac{1}{\sqrt{1 + \lambda^2}}$$

The variance of the skew normal distribution is

$$\text{var}(z) = \left\{1 - \frac{2}{\pi}, \delta^2(\lambda)\right\} \cdot \sigma^2$$

The moments of the skew normal distribution are given by,

$$\mu = m_1 = -a_1 \left(\frac{m_3}{b_1}\right)^{\frac{1}{3}}$$

$$\sigma^2 = m_2 = -a_1^2 \left(\frac{m_3}{b_2}\right)^{\frac{2}{3}}$$

$$\delta(\lambda) = \left\{a_1^2 + m_2 \left(\frac{b_1}{m_3}\right)^{\frac{1}{3}}\right\}^{-\frac{1}{2}}$$

where,

$$a_1 = \sqrt{\frac{2}{\pi}}; \quad b_1 = \left(\frac{4}{\pi - 1}\right) a_1;$$

$$m_1 = n^{-1} \sum_{i=1}^{n} y_i;$$

$$m_2 = (n - 1)^{-1} \sum_{i=1}^{n} (y_i - \bar{y})^2;$$

$$m_3 = (n - 1)^{-1} \sum_{i=1}^{n} (y_i - \bar{y})^3;$$

Since, the entire image is a collection of regions which are characterized by skew normal variants, we assume that the pixel intensities in the image region follows a k-component finite skew normal distribution and its probability density function is of the form

$$h(\lambda) = \Sigma_{i=1}^{k} \alpha_i \cdot g(y_i | \mu_i, \sigma_i^2, \lambda)$$

where, $\lambda$ is the skewness parameter and $k$ is the number of regions, $\alpha_i > 0$ are weights such that $\Sigma_{i=1}^{k} \alpha_i = 1$ and

$$g(y_i | \mu_i, \sigma_i^2, \lambda) = \frac{1}{\sqrt{2\pi}} \cdot \frac{1}{\sigma} \cdot \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{(y_i - \mu_i)^2}{2\sigma^2}} \cdot \phi \left(\lambda \left(\frac{y_i - \mu_i}{\sigma}\right)\right)$$

where, $\phi \left(\lambda \left(\frac{y_i - \mu_i}{\sigma}\right)\right) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{t^2}{2}} dt$.

The first step of 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 EM algorithm is to evaluate the maximum likelihood estimates of the unknown parameters $\theta$ by iterative process.

III. DATA SET

In order to validate the efficacy of the developed algorithm, the methodology is proposed by taking the benchmark data set of CDNet 2014.

It consists of six categories of video frames ranging up to 80000. It also containing the frames pertaining to shadows, illumination effects.

IV. METHODOLOGY

In this methodology each of the image is considered and the pixels are estimated into back ground and foreground images based on the threshold selected.
The threshold is considered as a difference between the current frame and the previous frame. The pixels with high threshold values are given as input to the skew Gaussian model is presented in section II of the article. The probability density function (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 are considered as background information else they are considered as foreground information.

V. EXPERIMENTATION

The experimentation carried out in MATLAB environment using the bench mark data set 2014 and extracting the image background features in line with the proposed model based on skew Gaussian distribution.

The initial estimation of the parameters are estimated by using K-means algorithm and E.M algorithm is considered for further updating of these parameters. The segmentation is carried by maximizing the probability function of the skew Gaussian model is presented in the section II of the article. The results were also compared with the model based on Gaussian Mixture Model.

VI. PERFORMANCE EVALUATION AND EXPERIMENTAL RESULTS

In order to validate the proposed model, we have considered the assessment metrics precision, Recall, Accuracy, F-Score, MSE, RMSE, FNR, FPR, PSNR, PWC The efficacy of the model can be justified by the value of computed precision, if it is high, it signifies high performance. On the other side, if method allocates the most of the pixels to background, the output precision value may be high, but proportionally the value of recall falls.

The formulas for computations of the above metrics are given by

\[
\text{Precision}=\frac{TP}{TP+FP}
\]

\[
\text{Recall}=\frac{TP}{TP+FN}
\]

\[
\text{Accuracy}=\frac{TP+TN}{TP+TN+FP+FN}
\]

\[
\text{F-score}=\frac{2\times \text{Precision}\times \text{Recall}}{\text{Precision}+\text{Recall}}
\]

\[
\text{MSE}=\frac{FP+FN}{M+N}
\]

\[
\text{RMSE}=\text{MSE}
\]

\[
\text{FNR}=\frac{FN}{TP+FN}
\]

\[
\text{FPR}=\frac{FP}{FP+TN}
\]

\[
\text{PSNR}=10\log_{10}\left(\frac{R^2}{\text{MSE}}\right)
\]

\[
\text{PWC}=100\times\frac{(FN+FP)}{(FN+TN+FP+TP)}
\]

The results derived using the proposed methodology is presented in the following tables 1 to 4 and the graphs based on assessment metrics in Fig. 2 to 4.

<table>
<thead>
<tr>
<th>Evaluation Metrics of different methods on THERMAL video from CDnet DATASET</th>
<th>GMM</th>
<th>SGMM</th>
</tr>
</thead>
<tbody>
<tr>
<td>PRECISION</td>
<td>0.0238</td>
<td>0.0568</td>
</tr>
<tr>
<td>RECALL</td>
<td>0.175</td>
<td>0.0237</td>
</tr>
<tr>
<td>ACCURACY</td>
<td>0.9561</td>
<td>0.9912</td>
</tr>
<tr>
<td>F-SCORE</td>
<td>0.0323</td>
<td>0.0615</td>
</tr>
<tr>
<td>MSE</td>
<td>0.0152</td>
<td>0.0032</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.1212</td>
<td>0.0324</td>
</tr>
<tr>
<td>FNR</td>
<td>0.0205</td>
<td>0.0040</td>
</tr>
<tr>
<td>FPR</td>
<td>0.7032</td>
<td>0.8562</td>
</tr>
<tr>
<td>PSNR</td>
<td>64.7802</td>
<td>77.826</td>
</tr>
<tr>
<td>PWC</td>
<td>4.3823</td>
<td>2.8162</td>
</tr>
</tbody>
</table>

The threshold is considered as a difference between the current frame and the previous frame. The pixels with high threshold values are given as input to the skew Gaussian model is presented in section II of the article. The probability density function (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 are considered as background information else they are considered as foreground information.
Towards the Development of Effective Video Segmentation Based on Skew Gaussian Mixture Model

Fig. 2 Precision of proposed method on Thermal dataset

Fig. 3 Recall of proposed method on Thermal dataset

Fig. 4 Accuracy of proposed method on Thermal dataset

Fig. 5 F-Score of proposed method on Thermal dataset

Fig. 6 MSE of proposed method on Thermal dataset

Fig. 7 RMSE of proposed method on Thermal dataset
**Fig. 8** FPR of proposed method on Thermal dataset

**Fig. 9** FNR of proposed method on Thermal dataset

**Fig. 10** PSNR of proposed method on Thermal dataset

**Fig. 11** PWC of proposed method on Thermal dataset

**Table 2**

<table>
<thead>
<tr>
<th>Metrics\Methods</th>
<th>GMM</th>
<th>SGMM</th>
</tr>
</thead>
<tbody>
<tr>
<td>PRECISION</td>
<td>0.0346</td>
<td>0.0708</td>
</tr>
<tr>
<td>RECALL</td>
<td>0.083</td>
<td>0.0572</td>
</tr>
<tr>
<td>ACCURACY</td>
<td>0.9652</td>
<td>0.9953</td>
</tr>
<tr>
<td>F-SCORE</td>
<td>0.0255</td>
<td>0.0532</td>
</tr>
<tr>
<td>MSE</td>
<td>0.012</td>
<td>0.0232</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.0067</td>
<td>0.1435</td>
</tr>
<tr>
<td>FPR</td>
<td>0.0208</td>
<td>0.0038</td>
</tr>
<tr>
<td>FNR</td>
<td>0.805</td>
<td>0.9218</td>
</tr>
<tr>
<td>PSNR</td>
<td>80.321</td>
<td>88.1032</td>
</tr>
<tr>
<td>PWC</td>
<td>3.0483</td>
<td>1.6321</td>
</tr>
</tbody>
</table>

**Fig. 12** Precision of proposed method on Camera Jitter dataset
Towards the Development of Effective Video Segmentation Based on Skew Gaussian Mixture Model

Fig. 13 Recall of proposed method on Camera Jitter dataset

Fig. 14 Accuracy of proposed method on Camera Jitter dataset

Fig. 15 F-Score of proposed method on Camera Jitter dataset

Fig. 16 MSE of proposed method on Camera Jitter dataset

Fig. 17 RMSE of proposed method on Camera Jitter dataset

Fig. 18 FPR of proposed method on Camera Jitter dataset
**TABLE 3**

Evaluation Metrics of different methods on dynamic back ground video from CD net DATASET

<table>
<thead>
<tr>
<th>Metrics\Methods</th>
<th>GMM</th>
<th>SGMM</th>
</tr>
</thead>
<tbody>
<tr>
<td>PRECISION</td>
<td>0.012</td>
<td>0.0209</td>
</tr>
<tr>
<td>RECALL</td>
<td>0.023</td>
<td>0.0128</td>
</tr>
<tr>
<td>ACCURACY</td>
<td>0.9632</td>
<td>0.9932</td>
</tr>
<tr>
<td>F-SCORE</td>
<td>0.017</td>
<td>0.0323</td>
</tr>
<tr>
<td>MSE</td>
<td>0.003</td>
<td>0.0035</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.0161</td>
<td>0.0342</td>
</tr>
<tr>
<td>FPR</td>
<td>0.023</td>
<td>0.0042</td>
</tr>
<tr>
<td>FNR</td>
<td>0.032</td>
<td>0.8218</td>
</tr>
<tr>
<td>PSNR</td>
<td>83.523</td>
<td>84.1021</td>
</tr>
<tr>
<td>PWC</td>
<td>3.4325</td>
<td>2.6281</td>
</tr>
</tbody>
</table>

*Fig.19 FNR of proposed method on Camera Jitter dataset*

*Fig.20 PSNR of proposed method on Camera Jitter dataset*

*Fig.21 PWC of proposed method on Camera Jitter dataset*

*Fig.22 Precision of proposed method on dynamic back ground dataset*

*Fig.23 Recall of proposed method on dynamic back ground dataset*
Towards the Development of Effective Video Segmentation Based on Skew Gaussian Mixture Model

Fig. 24 Accuracy of proposed method on dynamic back ground dataset

Fig. 25 F-Score of proposed method on dynamic back ground dataset

Fig. 26 MSE of proposed method on dynamic back ground dataset

Fig. 27 RMSE of proposed method on dynamic back ground dataset

Fig. 28 FPR of proposed method on dynamic back ground dataset

Fig. 29 FNR of proposed method on dynamic back ground dataset
Fig.30 PSNR of proposed method on dynamic background dataset

Fig.31 PWC of proposed method on dynamic background dataset

Fig.32 Precision of proposed method on Shadow dataset

Fig.33 Recall of proposed method on Shadow dataset

Fig.34 Accuracy of proposed method on Shadow dataset

| TABLE 4 |
|---|---|
| Evaluation Metrics of different methods on SHADOW video from CD net DATASET |
| Metrics\Methods | GMM | SGMM |
| PRECISION | 0.0232 | 0.0532 |
| RECALL | 0.0132 | 0.0587 |
| ACCURACY | 0.9448 | 0.9932 |
| F-SCORE | 0.0132 | 0.0182 |
| MSE | 0.0367 | 0.0072 |
| RMSE | 0.0040 | 0.0902 |
| FPR | 0.0132 | 0.0182 |
| FNR | 0.9325 | 0.9642 |
| PSNR | 73.2872 | 85.262 |
| PWC | 3.0023 | 2.6055 |
Towards the Development of Effective Video Segmentation Based on Skew Gaussian Mixture Model

Fig.35 F-Score of proposed method on Shadow dataset

Fig.36 MSE of proposed method on Shadow dataset

Fig.37 RMSE of proposed method on Shadow dataset

Fig.38 FPR of proposed method on Shadow dataset

Fig.39 FNR of proposed method on Shadow dataset

Fig.40 PSNR of proposed method on Shadow dataset
Fig.41 PWC of proposed method on Shadow dataset

VII. CONCLUSION

In this article, a model proposed for the effective segmentation of images based on the background information. The initial values of parameters are processed to obtain the convergent values using E.M algorithm. The experimentation results derived and also it is compared with data of the existing models based on Gaussian Distribution using the performance metrics like FPR, FNR, F-score, precision, Recall, etc. Specified in table – I to table – 4 and figures 2-41. From the above information presented in the tables and graphs it can be clearly identify that the proposed methodology perform well with respect to all the parameters and the results. The proposed methodology can be very well suited for most of the applications pertaining to segmentation.

REFERENCES


AUTHORS PROFILE

PavanKumar Tadiparthi, is working as a Associate Professor in the department of information Technology, MVGR College of Engineering, Vizianagram, Andhra Pradesh, india.He has about 13 years of teaching experience and his area of interest is image processing. He is a life member of ISTE, MIE

Srinivas Yarramalle, is a Professor in the department of Information Technology, GITAM India and has about 23 years of teaching experience GITAM University, and Visakhapatnam. His areas of interests include Speech Processing, Data mining, and Software reusability apart from Image processing. He is having more than 200 publications at national and International level. He is the author for 4 books and a life member of ISTE, CSI, IE, ISTAM, IISA and ISPS.