Scientific Research

On Segmentation of Moving Objects by Integrating PCA Method with the Adaptive Background Model

Noureldaim Emadeldeen, Mohammed Jedra, Noureldeen Zahid

Laboratory of Conception and Systems, Faculty of Sciences, Mohamed V University, Rabat, Morocco.
Email: emadnour2009@hotmail.com, jedra@fsr.ac.ma, zahid@fsr.ac.ma

ABSTRACT

Tracking and segmentation of moving objects are suffering from many problems including those caused by elimination changes, noise and shadows. A modified algorithm for the adaptive background model is proposed by linking Gaussian mixture model with the method of principal component analysis PCA. This approach utilizes the advantage of the PCA method in providing the projections that capture the most relevant pixels for segmentation within the background models. We report the update on both the parameters of the modified method and that of the Gaussian mixture model. The obtained results show the relatively outperform of the integrated method.

Keywords: Pixels; Gaussian Mixture Model; Principle Component Analysis; Background Model; Noise Process; Segmentation

1. Introduction

Object segmentation is an important technique used in many systems, especially in the field of computer vision. Background subtraction which involves calculating a reference image, labels the pixels corresponding to foreground objects. N. Friedman and S. Russell were proposed a Gaussian mixture model (GMM) for the background subtraction [1]. C. Stauffer and W. Grimson developed an algorithm for foreground segmentation based on the Gaussian mixture model [2,3]. W. Grimson and et al used tracking information in multicamera calibration, for object detection and classification [4]. T. Bouwmans, F. ElBaf and B. Vachon surveyed the background modeling using mixture of Gaussians for object detection [5]. F. Zbu and K. Fujimura demonstrated a face tracking method using GMM and EM algorithm [6]. P. Kaew-TraKulPong and R. Bowden proposed a method to detect moving shadows using Gaussian mixture model. The shadow detection also reduces the effect of small repetitive motions in the background scene [7]. On the other hand, the binding ellipsis, due to S. Cheung and C. Kamath is inappropriate for complex-shaped objects such as human beings. They used frame-differencing as it produces minimal false foreground trails behind objects [8].

The principal component analysis is as a linear transformation filter. It can be used to simplify a dataset and creates new classifiers. H. Moon and P. Jonathan introduced a statistical pattern recognition PCA algorithm which has become the famous benchmark PCA algorithm [9]. H. Kim, D. Kim, and S. Yang proposed PCA mixture models applicable to many areas, where applied as the alphabet recognition [10]. B. Antoni, M. Vijay and V. Nuno introduced a generalization of the Stauffer-Grimson background subtraction algorithm. However, the background subtraction algorithm was evaluated through extensive experiments, involving scenes with dynamic backgrounds, and high detection. Moreover, in their generalization the low false positive rates were obtained regardless of object scale suggesting that it may be suitable for surveillance on dynamic scenes [11]. I. Gómez, D. Olivieri, X. Vila, and S. Orozco proposed a method for foreground, background segmentation and created a feature vector for discriminating and tracking several people in the scene [12]. Y. Sheikh and M. Shah Introduced single distribution and kernel density [13]. Y. Yong and W. Ya-fei used Gaussian mixture model with spatial local correlation and Markov Chain [14]. R. Hassanpour, A. Shahbahrami and S. Wong proposed Gaussian mixture model and considered the higher level a priori knowledge of spatial and temporal relationship between the pixels in a video sequence [15]. G. Rajkumar, K. Srinivasarho and P. Sribivasa proposed a new image segmentation algorithm used the finite mixture of doubly truncated bivariate Gaussian distribution by integrating with the hierarchical clustering [16]. In fact, we cannot detect the boundaries of object with a good precision and it is difficult to separate a moving object and its shadow. In our

work, we attempt to integrate the PCA method with Gaussian mixture. Accordingly, we link the Gaussian mixture with PCA to obtain the feature of the shape in the scenes and classification of each object.

The remainder of the paper is organized as follows: in the Section 2, we present adaptive background subtraction algorithm based on Gaussian mixture model followed by PCA algorithm placed in the Section 3. The linkage between PCA and the adaptive model is explained in the Section 4. The applications of the algorithm with their experimental results are presented with comments in the Section 5.

2. Gaussian Mixture Models

In the adaptive mixture of Gaussian in our work, the pixel processes is considered at time series. We consider each pixel as an independent statistical process, incorporates all pixel observations. The mixture model does not distinguish components that correspond to background from those associated with foreground objects and record the observed intensity at each pixel over the previous N frames.

In the same manner as C. Stauffer and W. Grimson [3], we used the same number k of components ($K = 3, 5$). We consider the values of a particular pixel, with initial $\varnothing_0(x_0, y_0)$ over time t and we can define the history of this particular pixel as
$\{X_1, \cdots, X_t\} = \{I(\varnothing_{j-1}, j) : 1 \leq j \leq t\}$, where I is the image sequence. The probability of pixels observation data X is defined as follows:

$$p(X_t) = \sum_{k=1}^{K} W_{k,t} * \eta\left(X_t, \mu_{k,t}, \sum_{k,t}\right) \quad (1)$$

where $W_{k,t}$ is the weight of each component k^{th} at time t, $\mu_{k,t}$ is the mean and $\sum_{k,t}$ is the covariance mixture of each component k^{th} at time t respectively, η is a Gaussian probability density defined as follows:

$$\eta\left(X_t, \mu_{k,t}, \sum_{k,t}\right)$$
$$= \frac{1}{(2\pi)^{\frac{n}{2}} \left|\sum_k\right|^{\frac{1}{2}}} \exp\left\{\left(X_t - \mu_{k,t}\right)^T \sum_{k,t}^{-1} \left(X_t - \mu_{k,t}\right)\right\} \quad (2)$$

For simplicity we use $\sum_{k,t} = \sigma_k^2 I$, where I identity matrix. Next, we update the mixture of each pixel consecutively by reading in new pixel values and update each matching mixture component. We initialize the parameters W, μ and σ. They are updated for each frame in the mixture. We calculate W for each kth Gaussian component in the mixture. When the current pixel value matches none of the distributions, the least probable distribution is updated with the current pixel values. Then the parameters of the mixture are updated as follows:

$$W_{k,t} = (1-\propto)*W_{k,t-1} + \propto\left(M_{k,t}\right) \quad (3)$$

where the constant $\propto$ is a learning rate to update the components weights, $M_{k,t}$ is a dummy variable with value 1 for matching models and 0 otherwise. We introduce the following criterion for screening the preliminary Gaussian matching distribution with Mahalanobis distance less than a threshold T

$$\min\left(\left\|X_t - \mu_{k,t}\right\|_{\sum_k^{-1}}^2\right) < T \quad (4)$$

to collect the targeted pixels, with the update recursive equations

$$\mu_{k,t} = (1-p)*\mu_{k,t-1} + p*X_{k,t} \quad (5)$$

and

$$\sigma_{k,t}^2 = (1-p)*\sigma_{k,t-1}^2 + p*\sum_{k=1}^{K}\left(X_{k,t} - \mu_{k,t}\right)^T\left(X_{k,t} - \mu_{k,t}\right) \quad (6)$$

where $p = \alpha * \eta\left(X_i | \mu_t, \sigma_t\right)$ indicates how to accelerate the update. Thus, we can find the foreground resulted from the background and then find the best segment of the image. The non active background Gaussian is treated as foreground. In the ordering, we selected the ratios W_k / σ_k Large values of ratios are associated with distributions which have a high weight, low variance. The first B distributions chosen under the expression:

$$B = \arg\min_b\left(\sum_{k=1}^{b}\left(W_{k,t}\right)^3 > T\right) \quad (7)$$

where the threshold T is a prior probability estimated from the background process. Movements caused by the background (such as tree branches shaking, surface fluctuations, etc.) were considered to segment for the foreground objects.

3. Principle Component Analysis (PCA)

We used observation data $X = \{X_i\}$ where $i = 1, \cdots, N$. First we consider each of the N frames in one color (R) such that we can define the frame vectors for fixed i with respect to the jth pixels index:

$$R_i = \left[r_{i,j}\right] i = 1, \cdots, N \quad (8)$$

The arithmetic average of pixels in each component for the N_i frame vector can be defined as follows:

$$\bar{r}_i = \frac{1}{N_i}\sum_{j=1}^{N_i} r_{i,j} \quad (9)$$

where i^{th} is number of pixels in the frame i.

Next, we compute the difference $\varnothing_i = R_i - \bar{r}_i$ for the ith image $\varnothing_i$ and construct the difference matrix A as:

$$A = [\varnothing_1, \cdots, \varnothing_N]$$

The covariance matrix Ψ of each frame vector is defined by:

$$\Psi = AA^T = \frac{1}{N}\sum_{i=1}^{N}\varnothing_i\varnothing_i^T \tag{10}$$

We compute the eigenvectors and corresponding eigenvalues by:

$$\Psi V = \lambda V \tag{11}$$

Using Singular *V*alue Decomposition (SVD) method, where V is the set of eigenvectors associated with the corresponding eigenvalues λ. To apply the method of PCA we need to define the projection S_i by applying the translation matrix V^T to the vectors $(X_i - \bar{r}_i)$ as follows:

$$S_i = V^T (X_i - \bar{r}_i) \tag{12}$$

Next, we construct each frame to obtain the features of the shape. Then we can define the estimated data $\hat{x}_i$ by:

$$\hat{x}_i = V \cdot S_i + \bar{r} \tag{13}$$

One property of PCA is that the projection onto the principal subspace minimizes squared errors. To find the feature principal subspecies of PCA, the sequential reconstruction errors are estimated from:

$$\varepsilon = \sum_{i=1}^{N}\|X_i - \hat{x}_i\|^2 \tag{14}$$

the major advantage of the PCA method comes from its generalization ability. It reduces the feature space dimension by considering the variance of the input data. This method determines which projections are referable for representing the structure of input data. Those projections are selected in such a way that the maximum amount of information (*i.e.* maximum values of variances, where Eigenvalues represent the largest variances) is obtained in the smallest dimension of feature space.

4. Extended Gaussian Mixture

Since the Gaussian mixture model does not vary greatly with illumination changes and noise, we propose in this study an algorithm based on linking PCA method with GMM.

We extend Gaussian mixture of observing pixel data X_i from (1, 2), using the method of PCA. The estimated $\hat{x}$ is obtained by projecting the data on the axis, then with the new data $S = \{S_t\}$, $t = 1, N$, we propose to compute $p(X_t)$ based on S such that:

$$p(S_t) = \sum_{k=1}^{K}W_{k,t} * \eta\left(S_t, \hat{\mu}_{k,t}, \hat{\sum}_{k,t}\right) \tag{15}$$

$$\eta\left(S_t, \hat{\mu}_{k,t}, \hat{\sum}_{k,t}\right) = \frac{1}{(2\pi)^{\frac{n}{2}}|\sum|^{\frac{1}{2}}}\exp\left\{(S_{k,t})^T \sum_{k,t}^{-1}(S_{k,t})\right\} \tag{16}$$

and again for simplicity we assume $\hat{\sum}_{k,t} = \hat{\sigma}_{k,t}^2 I$, where I is identity matrix. In linking PCA with GMM it is evident that the integrated $p(S_t)$ based on $S = \{S_t\}$ by its physical structure, contains the maximum amount of information or has the minimum squared errors as compared to $p(X_t)$ based on $X = \{X_t\}$ of the GMM in handling the limitation and shadow that contribute to noise in the adaptive Gaussian model.

To update each mixture of each pixel entails reading the new pixel values and select the matching components, then compute the mean obtained with $\hat{x}$ and finally injecting that mean into the covariance of PCA method form (5, 6). Symbolically the update is based on:

$$\hat{\mu}_{k,t} = (1-p) * \hat{\mu}_{k,t-1} + p * \hat{x}_{k,t} \tag{17}$$

and

$$\hat{\sigma}_{k,t}^2 = (1-p) * \hat{\sigma}_{k,t-1}^2 + p * \sum_{k=1}^{K}(S_{k,t} - \hat{\mu}_{k,t})^T (S_{k,t} - \hat{\mu}_{k,t}) \tag{18}$$

where $p = \alpha * \eta(S_t|\hat{\mu}_t, \hat{\sigma}_t)$ indicates how to accelerate the update. In the integrated method we insert the estimates of the mean $\hat{\mu}_{i,t}$ and covariance $\hat{\sum}_{i,t}$ of each component in PCA into GMM respectively. We simply select the components with largest weights and re-estimate with (7) as the follows:

$$\hat{W}_{k,t} = (1-\alpha) * \hat{W}_{k,t-1} + \alpha\left(\hat{M}_{k,t}\right) \tag{19}$$

And their Gaussian matching parameters must satisfy:

$$\min\left(\|S_{k,t} - \hat{\mu}_{k,t}\|_{\Sigma_{k,t}^{-1}}^2\right) < T \tag{20}$$

with the estimates the ratios $\hat{W}_k/\hat{\sigma}_k$ Large values of ratios are associated with distributions which have a high weight, low variance and first $\hat{B}$ Gaussian distributions:

$$\hat{B} = \arg\min_b \left(\sum_{k=1}^{b}(\hat{W}_{k,t})^3 > T\right) \tag{21}$$

Then for estimating the distribution of data by integrating PCA and mixture Gaussian, we need to perform both the appropriate partitioning of the component and the estimation of model parameters. We can perform this task successfully due to the important property of the PCA method.

5. Experiment Results and Analysis

In our application, we performed a quantitative comparison between two algorithms for segmentation and objects

detection: The first algorithm based on Gaussian mixture model (GMM) and the second competing algorithm based on linking PCA with the adaptive background model.

In our experiment, each pixel is checked against the current frame of the background model by comparing it with every component in the Gaussian model under each of the two algorithms.

For each of the GMM method and PCA-GMM method, this process continues until a matching is found with k components ($k = 3, 5$). With changes in time, we used the data sequentially in the range 121 to 1000 frames and parameters $(\alpha = 0.0067)$ and $(p = 0.005)$.

The experiment resulted in two frame's sequences in the different backgrounds (tree, lightness).The frame's dimensions were 167×120. From **Figure 1** and **Figure 2**, it is obviously to observe the impact of the noise and lightness associated with GMM method.

Next we select each component of mixture model, minimum reconstruction error over the data need to be estimated. After the selection of one color in data, the estimation of the mean in the feature space, and the differences between mean and each color were computed.

After we applied the PCA to solve the problem of Gaussian mixture and obtained improved result in the two cases, see **Figure 1** and **Figure 3**, the performance enhanced remarkably with the PCA-GMM proposed detection method.

Clearly, the detection accuracy in terms of recall and precision of PCA + GMM method is consistently higher than the mixture of Gaussians approach, the GMM method, as illustrated by **Figure 2** and **Figure 4** that show the perframe detection rates.

Consequently, the PCA + GMM method reflects more accurate detection of object results than the GMM method at each level of the proposed approach.

6. Conclusion

As the appearance of scene varies with illumination changes and noise, it is difficult to separate and track objects in a dynamic scene environment. In this paper, we reported an application study based on a proposed
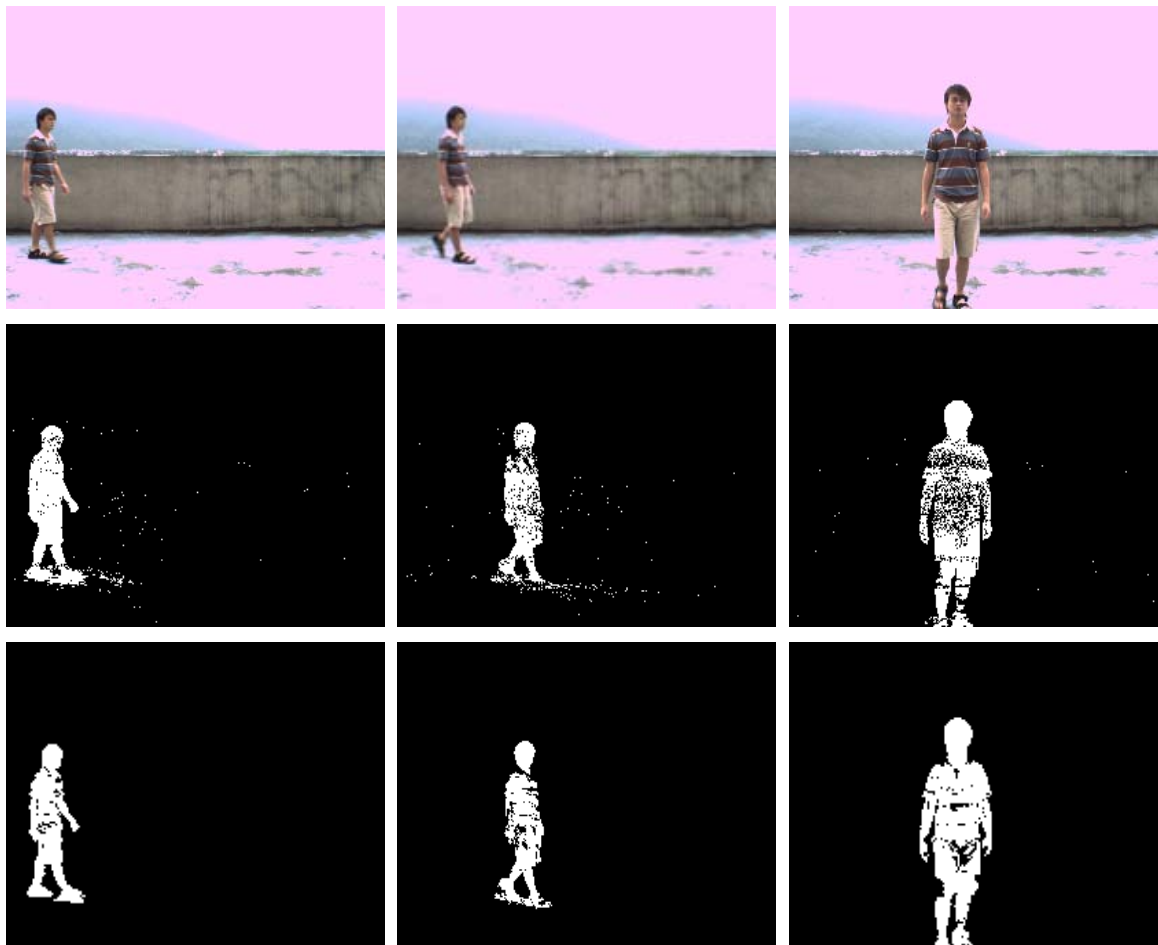

Figure 1. Sample frames, the top row are original images, the second row are the results obtained by Gaussians mixture method, and the third row results obtained by the PCA + GMM.
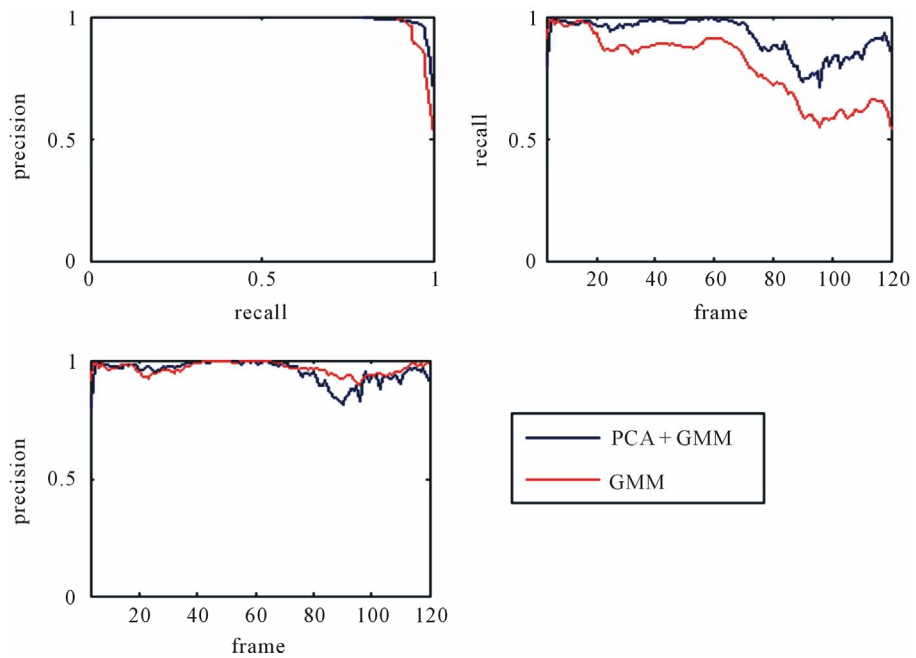
Figure 2. Precision-recall plots, for scenes in the Figure 1.
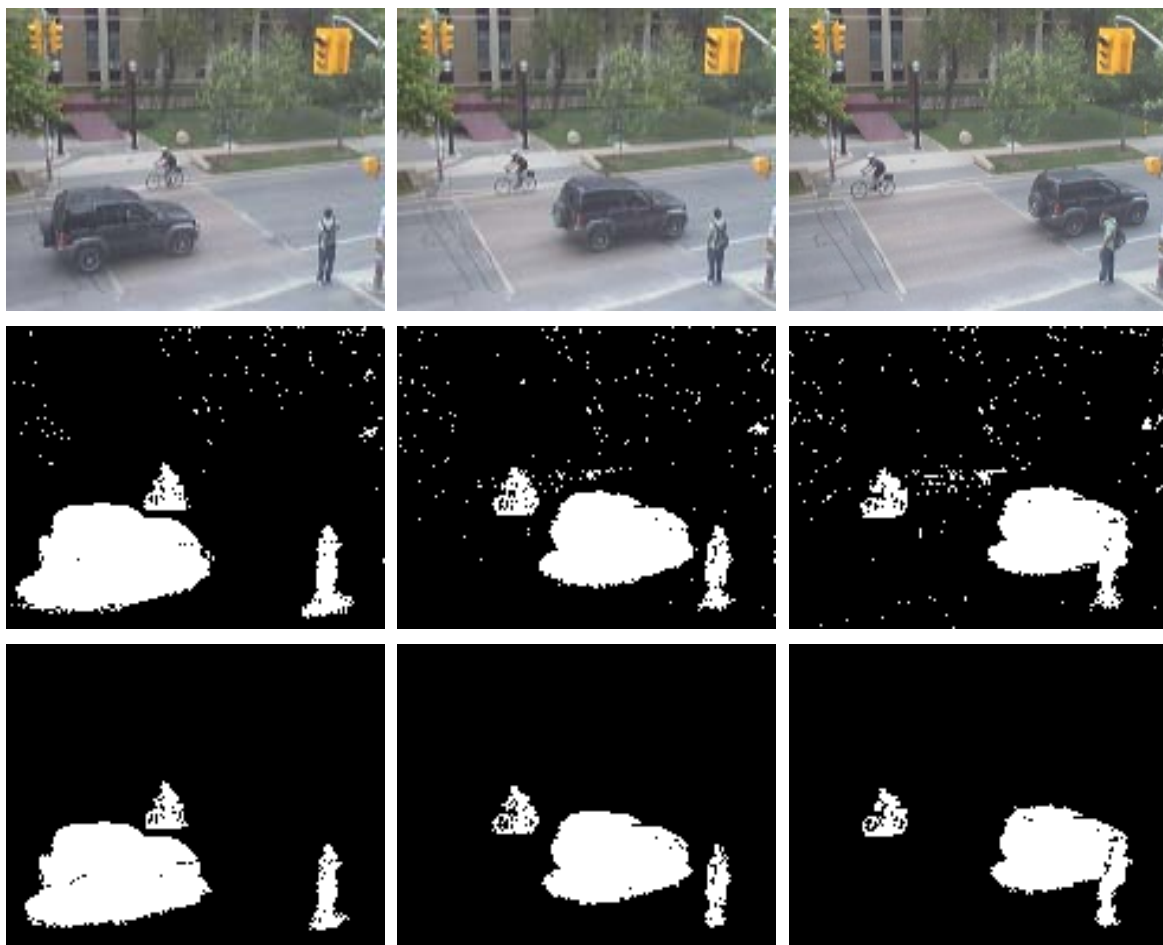

Figure. 3. Sample frames, in the top row are the original images, in the second row are the results obtained by using Mixture of Gaussians method and in the third row are the results obtained by the PCA + GMM.
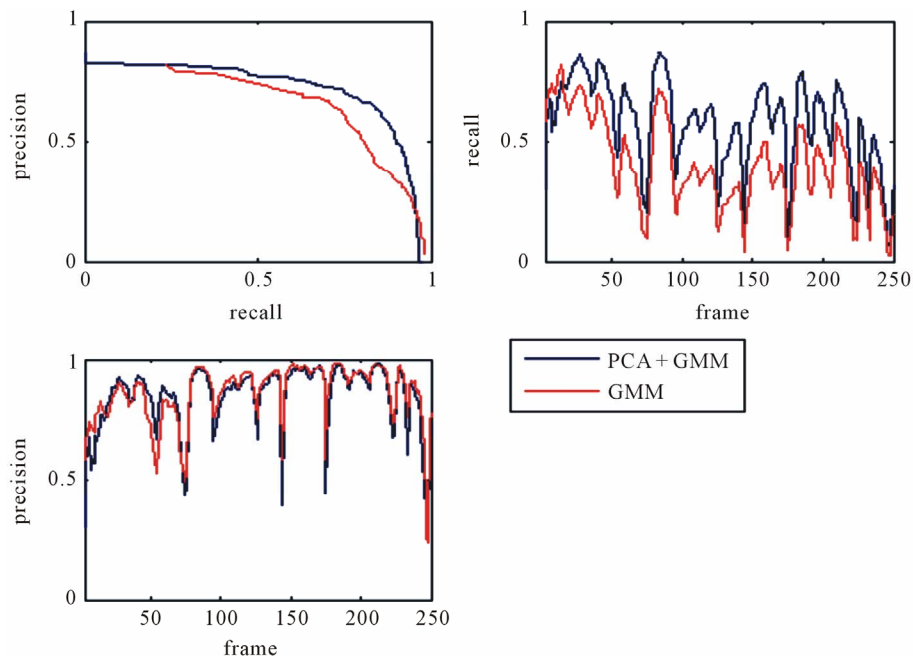
Figure 4. Precision-recall plots, for scenes in the Figure 3.

algorithm that links PCA and the Gaussian Mixture Model (GMM), to address this difficulty. The proposed algorithm has manipulated the variation to ease the segmentation of the foreground from the background. The study confirms that the proposed algorithm produces relatively good results compared to GMM method.

REFERENCES

[1] N. Friedman and S. Russell, "Image Segmentation in Video Sequences," *13th Conference on Uncertainty in Artificial Intelligence*, 1997, pp. 175-181.

[2] C. Stauffer and W. Grimson, "Adaptive Background Mixture Models for Real-Time Tracking," *Proceedings of IEEE Computer Vision and Pattern Recognition*, Vol. 2, 1999, pp. 246-252.

[3] C. Stauffer and W. Grimson, "Learning Patterns of Activity Using Real-Time Tracking," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, Vol. 22, No. 8, 2000, pp. 747-757. doi:10.1109/34.868677

[4] W. Grimson, C. Stauffer, R. Romano and L. Lee, "Using Adaptive Tracking to Classify and Monitor Activities in a Site," *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 23-25 June 1998, pp. 22-29.

[5] T. Bouwmans, F. El Baf and B. Vachon, "Background Modeling using Mixture of Gaussians for Foreground Detection," *Recent Patents on Computer Science*, 2008, pp. 219-237.

[6] Y. Zhu and K. Fujimura, "Driver Face Tracking Using Gaussian Mixture Model (GMM)," *Proceedings of IEEE on Intelligent Vehicles Symposium*, 9-11 June 2003, pp. 587-592.

[7] P. KaewTraKulPong and R. Bowden, "An Improved Adaptive Background Mixture Model for Real-Time Tracking with Shadow Detection," *2nd European Workshop on Advanced Video Based Surveillance Systems* (*AVB*S01), September 2001.

[8] S. Cheung and C. Kamath, "Robust Background Subtraction with Foreground Validation for Urban Traffic Video," *EURASIP Journal on Applied Signal Processing*, Vol. 14, 2005, pp. 2330-2340. doi:10.1155/ASP.2005.2330

[9] H. Moon and P. Jonathan, "Computational and Performance Aspects of PCA-Based Face-Recognition Algorithms," *Perception*, Vol. 30, No. 3, 2001. pp. 303-321. doi:10.1068/p2896

[10] H. Kim, D. Kim and S. Yang, "An Efficient Model Order Selection for PCA Mixture Model," *Pattern Recognition*, 2003, pp. 1385-1393.

[11] B. Antoni, M. Vijay and V. Nuno, "Generalized Stauffer-Grimson Background Subtraction for Dynamic Scenes," *Journal of Machine Vision and Applications*, Vol. 22, 2011, pp. 751-766.

[12] I. Gómez, D. Olivieri, X. Vila and S. Orozco, "Simple Human Gesture Detection and Recognition Using a Feature Vector and a Real-Time Histogram Based Algorithm" *Journal of Signal and Information Processing*, Vol. 2, No. 4, 2011, pp. 279-286. doi:10.4236/jsip.2011.24040

[13] Y. Sheikh and M. Shah, "Bayesian Modeling of Dynamic Scenes for Object Detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 27, No. 11, 2005, pp. 1778-1792. doi:10.1109/TPAMI.2005.213

[14] Y. Yong and W. Ya-Fei, "Moving Object Detection Based on Spatial Local Correlation," *China Academic Journal Electronic Publishing House*, *Opto-Electronic Engineering*, Vol. 36, No. 2, 2009.

[15] R. Hassanpour, A. Shahbahrami and S. Wong, "Adaptive Gaussian Mixture Model for Skin Color Segmentation," *Proceeding of World Academic of Science Engineering and Technology*, Vol. 31, 2008, pp. 1307-6884.

[16] G. Rajkumar, K. Srinivasarho and P. Sribivasa, "Image Segmentation Method Based on Finite Doubly Truncated Bivariate Gaussian Mixture Model with Hierarchical Clustering," *International Journal of Computer Science*, Vol. 8, No. 2, 2011.