Available online at www.sciencedirect.com

SciVerse ScienceDirect

Procedia Computer Science 10 (2012) 153 – 160

Procedia Computer Science

The 3rd International Conference on Ambient Systems, Networks and Technologies (ANT)

Real-time Video Segmentation using Student's t Mixture Model

Dibyendu Mukherjee, Q. M. Jonathan Wu

Department of Electrical and Computer Engineering,
University of Windsor,
Canada

Abstract

Mixture models for video segmentation have mainly revolved around Gaussian distributions for a long time due to their simplicity and applicability. In this work, we propose a novel real-time video segmentation algorithm based on Student's t mixture model. Though, Student's t-distribution has been used for image segmentation by applying Expectation Maximization (EM) algorithm, the same technique cannot be followed in video segmentation due to exceptional increase in computational complexity. Thus, in spite of being a more heavily-tailed distribution compared to Gaussian, Student's t mixture model remained unexplored for video segmentation. In this work, a novel and effective recursive filter based formulation has been introduced to update the mixture model with new observations. Our analysis and experimental results show that real-time, robust and improved video segmentation can be performed using Student's t mixture model compared to the conventional mixture models.

Keywords: Video segmentation, Student's t-distribution, Mixture models

1. Introduction

Detection of moving objects from a video sequence is a popular research area due to its vast application areas such as human-computer interaction, traffic monitoring, surveillance, content-based video compression and gesture recognition. The most common approach to identifying moving objects is to determine the background model, and then subtract each frame from the background to yield the foreground. The pixels that deviate from the background by a significant amount, are foreground pixels. The process, although simple to understand, poses a number of difficulties such as - slow foreground, shadows cast by foreground objects, nonstationary background (e.g. illumination variances, background movement due to wind etc.) and most importantly, real-time computation. Background subtraction has been thoroughly researched by different researchers [1, 2, 3]. Unfortunately, increase in speed reduces the robustness of algorithm, while increase in robustness prevents a real-time implementation.

Email address: `mukherjd@uwindsor.ca` (Dibyendu Mukherjee)

doi:10.1016/j.procs.2012.06.023

Modeling of spatiotemporal information is a major concern in foreground/background subtraction. Different approaches have been proposed [4, 5, 6] to model this information effectively. Among the approaches, Gaussian Mixture Model (GMM) has been successfully used to model the recent history of pixel intensities [5, 7, 8, 9]. The approach has proven to be robust and real-time. However, data segmentation requires that the GMM be trained by some variant of an EM algorithm that iteratively converges to an optimal solution. But, due to the constraint of real-time and dynamic nature of a video sequence, any iterative solution is not directly implementable. The application demands an approximate solution that can learn incrementally with each new video data, and yield a temporal model for the video sequence. Stauffer and Grimson [7] have successfully provided a recursive filter formulation to train the GMM. However, the approach suffered from slow adaptivity and susceptance to noise. Extensive research has been carried out to find alternative or improved methods that can counter these problems. Among the others, Effective GMM [10] has shown a faster convergence. Also, Conditional Random Field has been proposed [11] to take into account spatial information that reduces the effects of noise by high amount, with a compensation by reduction in computational speed. In recent years, a number of researchers proposed effective methods for background/foreground segmentation [12, 13]. But, the constraint of balancing the computational speed and accuracy still remains an open area.

In this work, we propose a new mixture model that is based on Student's t-distribution (STMM) for video segmentation. Student's t-mixture model has been successfully employed in image segmentation [14] and it has proven to be very robust against outliers (e.g. noises) due to its more heavily-tailed nature compared to Gaussian mixture model. For image segmentation, authors have used the abbreviation SMM for their model. Thus, we chose to use STMM to signify the difference in application area. Till now, STMM has not been applied to video processing in real-time, because EM algorithm cannot be directly applied to the process. The huge increase in complexity would prevent a real-time implementation. Thus, we propose a new real-time recursive filter approach to update the parameters of the distribution effectively. The method can be used to segment the background and foreground with high accuracy. Experimental results show that the method is very fast and robust against slow foreground and nonstationary background processes.

The conventional GMM model is discussed in Section 2. The proposed algorithm is detailed in Section 3. A number of experimental results and comparisons with other state-of-the-art methods are provided in Section 4 and conclusions are drawn in Section 5.

2. Conventional Gaussian Mixture Model

Video segmentation is fundamentally different from image segmentation, because, an image consists of only spatial distribution, while a video sequence consists of a set of temporally distributed pixels and at each time instant, the pixels also follow a spatial distribution. In conventional GMM, the values of a particular pixel over time is termed as "pixel process". Thus, the pixel process is a set that consists of scalar gray values for gray scale images, or vector of color values for color images. At time t, the history of a single pixel at position (x, y) consists of the set

$$\{X_1, ..., X_t\} \text{ with } X_i = I(x, y, i), \tag{1}$$

where, $I(x, y, i)$ denotes the D-dimensional pixel intensity (gray scale or color) at position (x, y) and time $i \in [1, t]$. Pixel processes are discussed in detail in [7]. Due to their dynamic nature, they need an adaptive mixture model for the effective representation. The recent history of a pixel can be modeled as a mixture of K Gaussians, as

$$f(X_t) = \sum_{j=1}^{K} w_{j,t} * \Phi(X_t; \mu_{j,t}, \Sigma_{j,t}). \tag{2}$$

Here, $w_{j,t}$ is the weight (or prior distribution) of the j^{th} Gaussian in the mixture at time t, and $\Phi(\cdot)$ is the Gaussian probability density function with mean $\mu_{j,t}$ and variance $\Sigma_{j,t}$ for j^{th} distribution at time t as follows

$$\Phi(X_t; \mu_{j,t}, \Sigma_{j,t}) = \frac{1}{(2\pi)^{\frac{D}{2}} \mid \Sigma \mid^{\frac{1}{2}}} \exp\left\{-\frac{1}{2}(X_t - \mu_{j,t})^T {\Sigma_{j,t}}^{-1}(X_t - \mu_{j,t})\right\}. \tag{3}$$

For reduction in computation, covariance matrix $\Sigma_{j,t}$ is assumed to be of the form $(\sigma_{j,t})^2 I$. This assumes independence among the different color channels with every channel having the same variance. This assumption, although not correct, avoids costly matrix inversion at the cost of slight decrease in accuracy.

After the model is constituted, the parameters of the model are updated to fit it to the new observations. Among the distributions of the model, some represent the background while the others represent the foreground and the shadows.

The representation of a pixel process using a GMM is very practical as a new pixel value would, in general, follow at least one distribution. However, the conventional GMM suffers from some of the inherent problems of video segmentation e.g. slow foreground, nonstationary background, background noises etc. Thus, following the basis of the conventional model, we propose our STMM model.

3. Proposed Algorithm

Real-time video segmentation is a complex problem considering the time limit and computational complexity needed. Till now, very few real-time robust algorithms have been proposed. In existing theory and application of image processing, Student's t-distribution has proven itself better than GMM for segmentation. But, due to the constraints of video segmentation, it has never been exploited for this application. Our algorithm proposes the application of Student's t-distribution to video segmentation with complete update formulation for the parameters of the distribution. This section has been divided in two parts for convenience. Section 3.1 discusses the distribution and formulation for mean and variance updates. The complete explanation and update equations for degrees of freedom is kept to Section 3.2 to highlight the importance and novelty.

3.1. Student's t Mixture Model

In this work, we propose the use of Student's t-distribution for the modeling of pixel processes. Here, the current pixel value X_t is represented as a mixture of Student's t-distributions as

$$f(X_t) = \sum_{j=1}^{K} w_{j,t} * \Psi(X_t; \mu_{j,t}, \Sigma_{j,t}, v_{j,t}), \tag{4}$$

where, $\Psi(\cdot)$ represents the Student's t probability density function with mean $\mu_{j,t}$, variance $\Sigma_{j,t}$ and degrees of freedom $v_{j,t}$ for j^{th} distribution at time t as

$$\Psi(X_t; \mu_{j,t}, \Sigma_{j,t}, v_{j,t}) = \frac{\Gamma\left(\frac{v_{j,t}+D}{2}\right) \mid \Sigma \mid^{-\frac{1}{2}}}{(\pi v_{j,t})^{\frac{D}{2}} \Gamma\left(\frac{v_{j,t}}{2}\right) [1 + v_{j,t}{}^{-1}(X_t - \mu_{j,t})^T \Sigma_{j,t}{}^{-1}(X_t - \mu_{j,t})]^{\frac{v_{j,t}+D}{2}}}. \tag{5}$$

The degrees of freedom $v_{j,t}$ control the shape and tail of the density function. It can be shown that for $v_{j,t} \longrightarrow \infty$, the distribution becomes a Gaussian distribution with covariance $\Sigma_{j,t}$. Also, for $v_{j,t} > 1$, $\mu_{j,t}$ is the mean of X_t. The parameter $\Sigma_{j,t}$ is treated in the same way as for GMM and is considered as of the form $(\sigma_{j,t})^2 I$. The independence of color channels is again assumed to avoid the matrix inversion.

For any mixture model, the parameters need to be updated at each time frame to model the new observations. However, as there exists a mixture model for every pixel in a video frame, an EM algorithm for optimization is highly intensive in computations. Also, new observations and changes reduce the dependency on the history. Thus, updating the parameters in each time step is a necessary requirement.

To address this problem, an online recursive filter based GMM was proposed in [7]. Following that model, every new pixel is compared against the K Gaussian means. A match is found if the new pixel value X_t is within a multiple of standard deviation from the mean. Mathematically, it can be written as

$$X_t \in \Phi(X_t; \mu_{j,t}, \Sigma_{j,t}) \text{ if } \mid X_t - \mu_{j,t} \mid < T\sigma_{j,t}, \tag{6}$$

where, T is a constant that denotes the multiple of standard deviation, normally lying in 2.5 − 3.5. For the matched distribution(s) (there may be more than one matched distribution), $w_{j,t}$, $\mu_{j,t}$ and $\sigma_{j,t}$ are updated according to the recursive filter equation as follows

$$w_{j,t} = (1-\alpha)w_{j,t-1} + \alpha;\ \mu_{j,t} = (1-\rho)\mu_{j,t-1} + \rho X_t;\ {\sigma_{j,t}}^2 = (1-\rho){\sigma_{j,t-1}}^2 + \rho(X_t - \mu_{j,t})^T(X_t - \mu_{j,t}) \quad (7)$$

where, α is the learning rate and ρ is the learning factor both lying in 0.01 − 0.1 range.

The formulation for $v_{j,t}$ is not that simple compared to $\mu_{j,t}$ and $\sigma_{j,t}$. At this point, we consider that $v_{j,t}$ is updated accordingly, and provide a dedicated section 3.2 for the explanation of the recursive formulation for $v_{j,t}$.

For unmatched distributions, $\mu_{j,t}$, $\sigma_{j,t}$ and $v_{j,t}$ remain same, while the prior weight is updated as the $(1-\alpha)$ fraction of the weight at previous time instant. If none of the distributions match the current pixel value, the distribution with lowest weight is replaced by a distribution with initially low weight, X_t as mean, a high variance and an initial value for the degrees of freedom.

Next, we need to determine which of distribution(s) represent the background. The background, in general, does not change rapidly, thus having a low variance. Also, background should be the highest probable value for any pixel over a sufficient period of time. In case of ego-motion, the variance may not stay at a low range, but it would still be lower compared to a moving foreground. Thus, it can be assumed that the distribution(s) with highest weight(s) and lowest variance(s) should constitute the background. Thus, the distributions are ordered by the value of w/σ [7]. Then, the first B distributions are chosen as the background for which the following holds

$$B = \arg\min_b \left(\sum_{j=1}^{b} w_{j,t} > Th \right), \quad (8)$$

where, Th is a threshold that determines the minimum amount of data constituting the background. If a single distribution is chosen, the mean of the distribution would represent the background intensity value. Otherwise, a sum of B means weighted according to their prior weights, would represent the background intensity. Here, a small discussion of the effect of Th on segmentation is necessary. Increasing Th would imply increasing the number of mixtures that constitute the background. Thus, a low value indicates an improperly constructed background, while a too high value would indicate an over-complete background where, foreground variations are also included as part of background.

3.2. Online Update of Degrees of Freedom

A pixel X_t follows a Student's t-distribution with mean μ_t, covariance matrix Σ_t and degrees of freedom v_t if, given a weight u_t, the pixel X_t follows a Gaussian distribution with mean μ_t and covariance Σ_t/u_t [14]. The weight u_t follows a Gamma distribution parameterized by v_t as follows

$$P(u_t; v_t) = Gamma(v_t/2, v_t/2) = (\frac{v_t}{2})^{\frac{v_t}{2}} \frac{1}{\Gamma(\frac{v_t}{2})} {u_t}^{\frac{v_t}{2}-1} e^{-\frac{v_t}{2}u_t} \quad (9)$$

Integrating out the weights from the joint density function would yield Eq. 5. Here, we left out the distribution index j to simplify the equation.

Before going into the approach of recursive filter for the degrees of freedom, a brief discussion of EM algorithm is necessary. For EM algorithm based iterative update of the parameters, the weights u_i are updated as follows [14]

$$u_i = \frac{v_{i-1} + D}{v_{i-1} + \delta_{i-1}} \quad (10)$$

where, $\delta_i = (X_t - \mu_i)^T {\Sigma_i}^{-1} (X_t - \mu_i)$ represents the Mahalanobis squared distance at i^{th} iteration. The following can be proven using the definition of δ_i

$$E[\delta_i] = D \quad (11)$$

Thus, $\delta_i = E[\delta_i]$ implies $u_i = 1$ and u_i inversely changes with the change in δ_i from its expected position. Also, for EM algorithm, the degrees of freedom are updated as the solution to the following equation

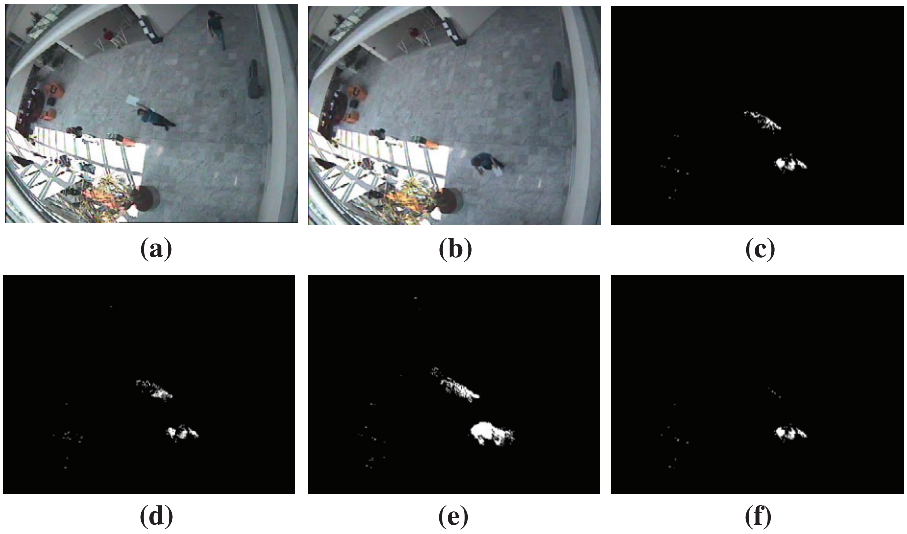


Fig. 1. First experiment - ghost in motion: (a) Original video frame 1, (b) Original video frame 52, (c) GMM output, (d) EGMM output, (e) CRF output, (e) STMM output

$$log(\frac{v_i}{2}) - \frac{\partial(\log(\Gamma(\frac{v_i}{2})))}{\partial(\frac{v_i}{2})} + 1 - \log(\frac{v_{i-1}+D}{2}) + \frac{\sum_{j=1}^{t} z_t(\log(u_{i-1} - u_{i-1}))}{\sum_{j=1}^{t} z_t} + \frac{\partial(\log(\Gamma(\frac{v_{i-1}+D}{2})))}{\partial(\frac{v_{i-1}+D}{2})} = 0, \quad (12)$$

where, z_t is the posterior weight at t^{th} time, and $\frac{\partial(\log(\Gamma(x)))}{\partial(x)}$ is the Digamma function.

From the above, a recursive formulation for degrees of freedom is not readily available. Thus, we propose an approach that divides the computation of $v_{j,t}$ for j^{th} distribution in two steps. First, we compute an approximation of $u_{j,t}$. The computation is based on Eq. 10, but we have removed the degrees of freedom term from the numerator and denominator. This is because, we only need the amount of deviation of u_t from unity and not the exact value of it. Also partly because, we cannot use an exact iterative method like EM and the original formulation cannot be directly applied.

$$u_{j,t} = \frac{D}{(X_t - \mu_{j,t})^T \Sigma_{j,t}^{-1} (X_t - \mu_{j,t})}. \quad (13)$$

Next, we define an index $\xi_{j,t}$ as follows

$$\xi_{j,t} = \begin{cases} -1 & \text{if } |\log(u_{j,t})| > |\log(u_{j,t-1})|; \\ 1 & \text{otherwise.} \end{cases} \quad (14)$$

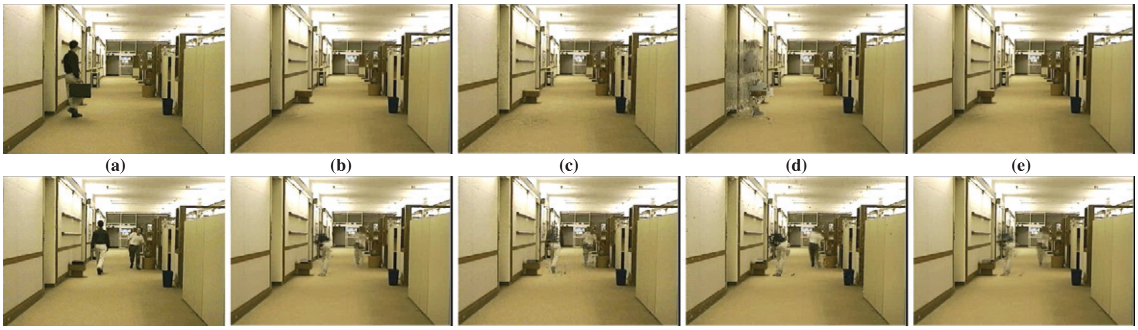
$\xi_{j,t}$ determines how far the weights $u_{j,t}$ have shifted from unity. The $\log(\cdot)$ is used to shift the weights to origin. The formulation for updating $v_{j,t}$ is as follows

$$v_{j,t} = v_{j,t-1} + f * \rho * \xi_{j,t}, \quad (15)$$

where, f is a multiplication factor that determines the increment size of $v_{j,t}$. The insensitiveness against outliers is quite dependent on f.

Heuristically, a decrease in $\delta_{j,t}$ would mean that the new observation is getting closer to the mean and a distribution with less "spread" should be able to model it, while an increase in $\delta_{j,t}$ would need more "spreading". Also it can be seen by plotting the solution of $v_{j,t}$ from Eq. 12 as a function of $u_{j,t}$ while keeping other parameter values constant, that $v_{j,t}$ drops with the deviation of $u_{j,t}$ from unity. Eq. 15 recursively updates $v_{j,t}$ accordingly.

This formulation does not guarantee an optimal solution for $v_{j,t}$ at each time instant. But, we provide experimental results and loglikelihood information that shows the formulation is quite effective.



(a) (b) (c) (d) (e) (f) (g) (h) (i) (j)

Fig. 2. Second experiment - radial motion: (a) Original video frame 34, corresponding (b) GMM output, (c) EGMM output, (d) CRF output, (e) STMM output, (f) Output video frame 160, corresponding (g) GMM output, (h) EGMM output, (i) CRF output, (j) STMM output

4. Experimental Results

We have tested the proposed algorithm on the image sequence data from Caviar Database and also on different types of video data. Here, we provide some of the results of our experiments. We compare our results to conventional GMM [7], Effective GMM (EGMM) [10] and Conditional Random Field based method (CRF) [11]. All algorithms run on MATLAB on a computer with 3 GHz AMD Phenom II X6 Processor. The number of clusters has been kept constant at 5.

The experimental section has been divided into five sections. First, we show the robustness of the algorithm against slow foreground that keeps a "ghost" image on the background. Secondly, we show background extracted from radial motion in the "Hall" sequence. Radial motion is very hard for detection as the moving object occupies a portion of the same position in the frame while the size of the object changes. Third experiment provides an example of dynamic background using the "highway_qcif" video sequence that shows a road from a driver's point of view. In the fourth experiment, we show the loglikelihood for 120 frames in the "viptraffic" sequence (from MATLAB sample videos) for GMM, EGMM and STMM. Finally, in the fifth experiment, we simply compare the computational speeds of the algorithms by frame rates.

For the first experiment, we show two frames from "Fight_RunAway" sequence (size 384×288 pixels) in Fig. 1. Fig. 1(a) shows frame 1 consisting of a man standing while carrying a board in his hand. Due to his immobile posture, he is assimilated into the background. As it is the first frame, no algorithm would yield any output for this frame. Thus, the outputs are not shown. In the frame shown in Fig. 1(b) (frame 52), the man has moved considerably. The results of segmentation are shown for each algorithm. The moving person is well detected by all the algorithms. But, a ghost remains for GMM, EGMM and CRF. While the amount of foreground detected for CRF is more compared to others, the amount of ghost is also higher. For STMM, the ghost is reduced by a high amount because of the better adaptation of Student's t-distribution to the new observations.

The second experiment consists of background extraction in radial motion. A radial motion keeps the moving object occupying a part of the same pixels in a video sequence, thus creating a wrong belief of the object to be a part of the background. In Fig. 2, we show frame 34 and frame 160 of the Hall sequence (size 352×240 pixels) with the corresponding background detection by different algorithms. In frame 34, the man is moving parallel to the image plane and thus, the background is well detected by GMM and STMM. But, EGMM has some artifacts in the background that belong to the foreground while CRF is still highly affected. In frame 160, the man on the left is moving radially after staying at a place for some time. A close inspection of the outputs yields that GMM has the ghost of the position when the man stayed still. STMM also has the ghost, but, the intensity is lower compared to GMM. EGMM learns at a faster rate and can quickly adapt to the changed background. Thus, it quickly adapted to the radial motion of the man and considered it as part of the background contrary to other algorithms that considered the still position of the man as background. Lastly, CRF is again highly affected. The experiment shows that STMM is comparably

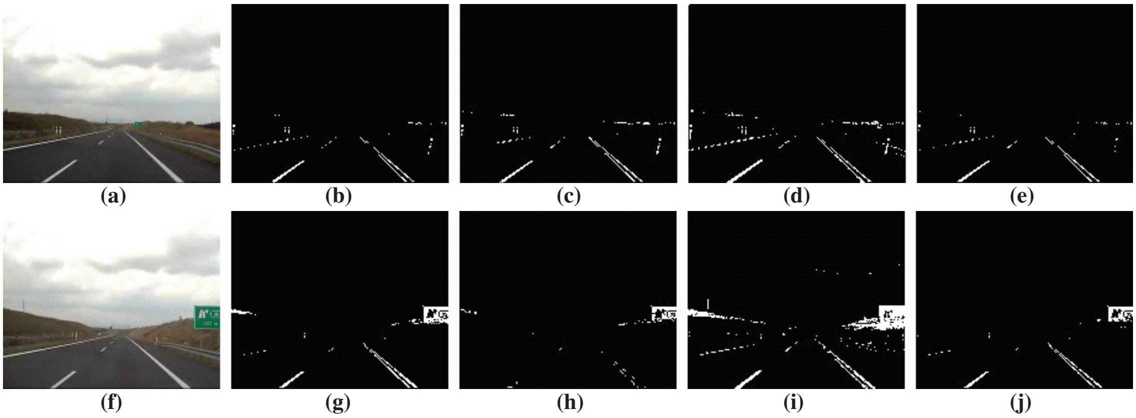


Fig. 3. Third experiment - dynamic background: (a) Original video frame 25, corresponding (b) GMM output, (c) EGMM output, (d) CRF output, (e) STMM output, (f) Output video frame 97, corresponding (g) GMM output, (h) EGMM output, (i) CRF output, (j) STMM output

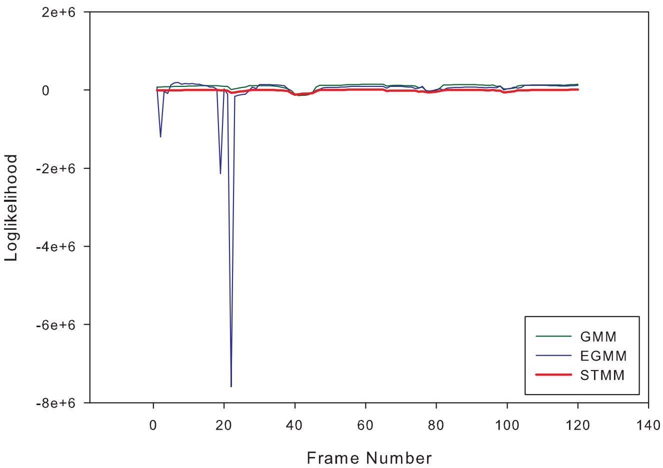


Fig. 4. Fourth experiment: The plot of loglikelihood of 120 video frame sequence for GMM, EGMM and STMM

robust in estimating the background.

The third experiment shows a case of dynamic background. If a camera is put on the front seat of a car while the car is moving, the background itself is nonstationary. If the mean texture of the background remains same as in the video sequence, we compare the robustness against the small changes in texture that can be considered as background noise. Fig. 3(a) and 3(f) show the frames 25 and 97 respectively. The background is very similar with some small texture changes and objects like the sign board. As can be seen from Fig. 3(b), 3(c), 3(g) and 3(h), GMM and EGMM work well but noise reduction is comparatively better for STMM in Fig. 3(e) and 3(j). As already shown, CRF has better detection percentage but is less robust against the noises.

In the fourth experiment, we plot the loglikelihood data of the entire viptraffic video sequence for GMM, EGMM and STMM in Fig. 4. For CRF, the computation of loglikelihood is complicated and not necessary. The plot shows that STMM consistently follows the loglikelihood of GMM. Thus, the experiment verifies that the recursive formulations for the parameters are consistent.

Finally, we compare the computational speeds of all the algorithms on a video sequence. We have used the "viptraffic" sequence (size 160×120 pixels) of 120 frames and run the algorithms in the previously mentioned computer with 3 clusters. Table 1 lists the algorithms compared with their total time taken and

Table 1. Comparison of Computational speed

Algorithm	GMM	EGMM	CRF	STMM
Total time (seconds)	6.14	4.5	41.2	6.8
Frames per second	19-20	25-30	2-4	17-19

frames per second. The variation in frame rate is due to the amount of foreground present in a frame. From the table, it is clear that STMM is comparable to GMM in terms of speed while EGMM is fastest and CRF is slowest. STMM provides an optimum performance in terms of the quality of the segmentation and number of frames processed per second.

5. Conclusion

In this work, we have proposed a video segmentation algorithm based on Student's t mixture model. The model is completely novel in the sense that, Student's t-distribution was not exploited for video segmentation before, due to its computation complexity and absence of EM algorithm. A new online recursive filter based update method is also provided for different parameters of the distributions. Experimental results show the improvement in performance compared to the conventional methods. Future work would include exploiting the spatial information to improve accuracy and searching for a better way to auto detect the number of clusters.

Acknowledgements

This research has been supported in part by the Canada Research Chair Program, AUTO21 NCE, and the NSERC Discovery grant. The authors would also like to acknowledge the EC Funded CAVIAR project/IST 2001 37540 (http://homepages.inf.ed.ac.uk/rbf/CAVIAR/) for the CAVIAR datasets.

References

[1] S. Toral, M. Vargas, F. Barrero, M. Ortega, Improved sigma-delta background estimation for vehicle detection, Electronics Letters 45 (1) (2009) 32 –34.

[2] I. Haritaoglu, D. Harwood, L. S. Davis, W4: Who? when? where? what? a real time system for detecting and tracking people, in: International Conference on Automatic Face and Gesture Recognition, 1998, pp. 222–227.

[3] S.-C. S. Cheung, C. Kamath, Robust background subtraction with foreground validation for urban traffic video, EURASIP Journal on Applied Signal Processing 2005 (2005) 2330–2340.

[4] H. Greenspan, J. Goldberger, A. Mayer, Probabilistic space-time video modeling via piecewise GMM, IEEE Transactions on Pattern Analysis and Machine Intelligence 26 (2004) 384–396.

[5] Y. Zhang, Z. Liang, Z. Hou, H. Wang, M. Tan, An adaptive mixture gaussian background model with online background reconstruction and adjustable foreground mergence time for motion segmentation, in: International Conference on Industrial Technology, 2005, pp. 23–27.

[6] P. KaewTraKulPong, R. Bowden, An improved adaptive background mixture model for real-time tracking with shadow detection, in: European Workshop on Advanced Video-Based Surveillance Systems, 2001, pp. 747–757.

[7] C. Stauffer, W. Grimson, Adaptive background mixture models for real-time tracking, IEEE Transactions on Pattern Analysis and Machine Intelligence 2 (1999) 246–252.

[8] I. Schiller, R. Koch, Improved video segmentation by adaptive combination of depth keying and mixture-of-gaussians, in: Scandinavian conference on Image analysis, 2011, pp. 59–68.

[9] Q. Zhu, Y. Xie, J. Gu, L. Wang, A new video object segmentation algorithm by fusion of spatio-temporal information based on GMM learning, Advances in Automation and Robotics,LNCS 123 (2012) 641–650.

[10] D.-S. Lee, Effective gaussian mixture learning for video background subtraction, IEEE Transactions on Pattern Analysis and Machine Intelligence 27 (5) (2005) 827–832.

[11] Y. Wang, Foreground and shadow detection based on conditional random field, in: the 12th international conference on Computer analysis of images and patterns, 2007, pp. 85–92.

[12] M. S. Allili, N. Bouguila, S. Boutemedjet, Image and video segmentation by combining unsupervised generalized gaussian mixture modeling and feature selection, IEEE Transactions on Circuits, systems and Video Technology 20 (2010) 1376–1377.

[13] X. Bai, J. Wang, G. Sapiro, Dynamic color flow: A motion-adaptive color model for object segmentation in video, in: European Conference on Computer Vision, 2010, pp. 617–630.

[14] G. Sfikas, C. Nikou, N. Galatsanos, Robust image segmentation with mixtures of student's t-distributions, in: IEEE International Conference on Image Processing, Vol. 1, 2007, pp. I–273–I–276.