Modeling of Dynamic Backgrounds by Type-2 Fuzzy Gaussians Mixture Models

Thierry Bouwmans, Fida El Baf

Modeling of Dynamic Backgrounds by Type-2 Fuzzy Gaussian Mixture Models

T. Bouwmans, F. El Baf

Abstract—**Gaussian Mixture Models (GMMs) are the most popular techniques in background modeling but present some limitations when some dynamic changes occur like camera jitter, illumination changes, movement in the background. Furthermore, the GMM are initialized using a training sequence which may be noisy and/or insufficient to model correctly the background. All these critical situations generate false classification in the foreground detection mask due to the related uncertainty. In this context, we propose to model the background by using a Type-2 Fuzzy Gaussian Mixture Models. The interest is to introduce descriptions of uncertain parameters in the GMM. Experimental validation of the proposed method is performed and presented on a diverse set of RGB and infrared videos. Results show the relevance of the proposed approach.**

Index Terms—**Background Modeling, Gaussian Mixture Models, Foreground Detection.**

I. Introduction

MANY video surveillance systems in visible spectrum [1], [2], [3] or infrared (IR) [4], [5], [6] need in the first step to detect moving objects in the scene. The basic operation used is the separation of the moving objects called foreground from the static information called the background. The process is called the background subtraction. In the literature, many background modeling methods have been developed and the most recent surveys can be found in [7], [8]. These background modeling methods can be classified in the following categories: Basic Background Modeling [9], [10], [11], Statistical Background Modeling [12], [13], [14] and Background Estimation [15], [16], [17]. Reading the literature, two remarks can be made: The most used models are the statistical ones due to their robustness to the critical situations. The first way to represent statistically the background is to assume that the history over time of intensity values of a pixel can be modeled by a single Gaussian (SG) [12]. However, a unimodal model cannot handle dynamic backgrounds when there are waving trees, water rippling or moving algae. To solve this problem, the Mixture of Gaussians (MOG) (or Gaussian Mixture Model (GMM)) has been used to model dynamic backgrounds [13]. This model has some disadvantages. Background having fast variations cannot be accurately modeled with just a few Gaussians (usually 3 to 5), causing problems for sensitive detection. So, a non-parametric technique [14] was developed for estimating background probabilities at each pixel from many recent samples over time using Kernel density estimation (KDE) but it is time consuming. In [18], Subspace Learning using Principal Component Analysis (SL-PCA) is applied on N images to construct a background model, which is represented by the mean image and the projection matrix comprising the first p significant eigenvectors of PCA. In this way, foreground segmentation is accomplished by computing the difference between the input image and its reconstruction. These four models define the first category using basic statistical model. The second category uses more sophisticated statistical models as Support Vector Machines (SVM) [19], Support Vector Regression (SVR) [20] and Support Vector Data Description (SVDD) [21]. The third category generalizes the models of the first category as the single general Gaussian (SGG) [22], the mixture of general Gaussians (MOGG) [23] and subspace learning using Independent Component Analysis (SL-ICA) [24], [25] or using Incremental Non-negative Matrix Factorization (SL-INMF) [26], [27] or using Incremental Rank-(R1,R2,R3) Tensor (SL-IRT) [28]. The Table I shows an overview of the statistical background modeling. The first column indicates the categories and the second column the name of each method. Their corresponding acronym is indicated in the first parenthesis and the number of papers counted for each method in the second parenthesis. The third column gives the name of authors and the dates of the first related publication. We can see that the MOG (or GMM) with 100 papers is the most used and improved due to a good compromise between robustness and time/memory requirements. In the GMM initialization, an expectation-maximization (EM) algorithm is used and allows to estimate GMM parameters from a training sequence according to the maximum likelihood (ML) criterion. The GMM is completely certain once its parameters are specified. However, because of insufficient or noisy data in training sequence, the GMM may not accurately reflect the underlying distribution of the observations according to the ML estimation. It is problematical to use likelihoods that are themselves precise real numbers to evaluate GMM with uncertain parameters.

Manuscript received September 9, 2009

T. Bouwmans is a member of the laboratory MIA at the University of La Rochelle, France (phone: (33).05.46.45.72.02; fax: (33) 05.46.45.82.40; e-mail: tbouwman@univ-lr.fr).

F. El Baf is a member of the laboratory MIA at the University of La Rochelle, France. (e-mail: felbaf@univ-lr.fr).

TABLE I
STATISTICAL BACKGROUND MODELING: AN OVERVIEW

Categories	Methods	Authors - Dates
First Category	Single Gaussian (SG) (5) Mixture of Gaussians (MOG - GMM) (~100) Kernel Density Estimation (KDE) (21) Principal Components Analysis (SL-PCA) (15)	Wren *et al.* (1997) [12] Stauffer and Grimson (1999) [13] Elgammal *et al.* (2000) [14] Oliver *et al.* (1999) [18]
Second Category	Support Vector Machine (SVM) (3) Support Vector Regression (SVR) (2) Support Vector Data Description (SVDD) (5)	Lin *et al.* (2002) [19] Wang *et al.* (2006) [20] Tavakkoli *et al.* (2006) [21]
Third Category	Single General Gaussian (SGG) (3) Mixture of General Gaussians (MOGG) (3) Independent Component Analysis (SL-ICA) (2) Incremental Non Negative Matrix Factorization (SL-INMF) (2) Incremental Rank-(R_1,R_2,R_3) Tensor (SL-IRT) (1)	Kim *et al.* (2007) [22] Allili *et al.* (2007) [23] Yamazaki *et al.* (2006) [24] Bucak *et al.* (2007) [26] Li *et al.* (2008) [28]

To solve this problem, we propose to model the background by using a Type-2 Fuzzy Gaussians Mixture Model (T2-FGMM) recently developed by Zeng et al. [29] to introduce descriptions of uncertain parameters in the GMM.

The rest of this paper is organized as follows: In the section II, we present briefly related works on GMM's improvements. In the section III, the T2-FGMM is used for background modeling. In the section IV, qualitative and quantitative experiments on RGB videos and infrared videos show that T2-FGMM outperforms the crisp GMM when dynamic changes occur. Finally, we present conclusions and perspectives in Section V.

II. RELATED WORKS

The original GMM for background modeling was proposed by Stauffer and Grimson [13] and present several advantages. Indeed, it can work without having to store an important set of input data in the running process. The multimodality of the model allows to tackle multimodal backgrounds and gradual illumination changes. Despite it, this model present some disadvantages: the number of Gaussians must be predetermined, the need for good initializations, the dependence of the results on the true distribution law which can be non-Gaussian and slow recovery from failures. Others limitations are the needs for a series of training frames absent of moving objects and the amount of memory required in this step. To alleviate these limitations, numerous improvements have been proposed over the recent years as shown by the different acronyms found like AKGMM [30], TLGMM [31], STGMM [32], SKMGM [33], TAPPMOG [34] and S-TAPPMOG [35]. All the developed improvements can be classified following the strategies used to be more robust to the critical situations met in video sequences [15]. The first strategies called intrinsic strategies consist to be more rigorous in the statistical sense or to introduce spatial and/or temporal constraint in the different step of the model. For example, some authors propose to determine automatically and dynamically the number of Gaussians to be more robust to dynamic backgrounds [36], [37], [38], [39], [40], [41]. Other approaches use another algorithm for the initialization [42], [43] and allow presence of foreground objects in the training sequence [44]. For the maintenance, the learning rates are better set [45], [46] or adapt over time [47], [48]. For the foreground detection, some authors use a different measure for the matching test [42] or use a foreground model [49]. Another way to improve the efficiency and robustness of the original GMM consist in using extrinsic strategies. Some authors used Markov Random Fields [50], hierarchical approaches [51], multi-level approaches [52], multiple backgrounds [53], graph cuts [54], multilayer approaches [55] or specific post-processing [56]. A recent complete survey of these improvements can be found in [57]. However, none of these improvements consider the uncertainty related to insufficient or noisy data in training sequence. Nevertheless, due to this uncertainty, the GMM may not accurately reflect the underlying distribution of the observations according to the ML estimation. One way to take into account this uncertainty is to use fuzzy concepts with the GMM. A first approach developed by [58] consists in the fuzzy GMM (FGMM) that estimates its parameters based on the modified fuzzy cmeans algorithm. So, the FGMM focuses on the precise parameter estimation of GMMs using fuzzy approaches rather than modeling GMMs uncertain parameters. On the other hand, Type-2 fuzzy sets (T2-FSs) [59] provide a theoretically well-founded framework to handle GMMs uncertain parameters. Their recent success achieved in pattern recognition has been largely attributed to their three-dimensional membership functions (MFs) for modeling uncertainties. Recently, Zeng et al. [29] introduce the Type-2 fuzzy sets in the GMM and called it Type-2 Fuzzy Gaussian Mixture Model (T2-FGMM). Experimental validations made in [29] show the superiority of T2-FGMM in pattern classification. In this context, we propose to apply the T2-FGMM for background modeling to take into account the uncertainty.

III. BACKGROUND MODELING USING TYPE-2 FGMM

In this section, we apply the Type-2 FGMM to background modeling and explain it in the case of RGB videos but it can be apply to infrared videos too by reducing the dimension of the observation to one dimension.

A. Principle

Each pixel is characterized by its intensity in the RGB color space. So, the observation o is a vector X_t in the RGB space and $d=3$. Then, the GMM is composed of K mixture components of multivariate Gaussian as follows:

$$P(X_t)=\sum_{i=1}^{K}\omega_{i,t}.\eta(X_t,\mu_{i,t},\Sigma_{i,t}) \quad (1)$$

where the parameters are K is the number of distributions, $\omega_{i,t}$ is a weight associated to the ith Gaussian at time t with mean $\mu_{i,t}$ and standard deviation $\Sigma_{i,t}$. η is a Gaussian probability density function:

$$\eta(X_t,\mu,\Sigma)=\frac{1}{(2\pi)^{3/2}|\Sigma|^{1/2}}e^{-\frac{1}{2}(X_t-\mu)\Sigma^{-1}(X_t-\mu)} \quad (2)$$

For the T2-FGMM-UM, the multivariate Gaussian with uncertain mean vector is:

$$\eta(X_t,\tilde{\mu},\Sigma)=\frac{1}{(2\pi)^{3/2}|\Sigma|^{1/2}}\exp(\prod -\frac{1}{2}\left(\frac{X_{t,c}-\mu_c}{\sigma_c}\right)^2) \quad (3)$$

with $\tilde{\mu}_c\in[\underline{\mu_c},\overline{\mu_c}]$ and $c\in\{R,G,B\}$.

For the T2-FGMM-UV, the multivariate Gaussian with uncertain variance vector is:

$$\eta(X_t,\mu,\tilde{\Sigma})=\frac{1}{(2\pi)^{3/2}|\Sigma|^{1/2}}\exp(\prod -\frac{1}{2}\left(\frac{X_{t,c}-\mu_c}{\sigma_c}\right)^2) \quad (4)$$

with $\tilde{\sigma}_c\in[\underline{\sigma_c},\overline{\sigma_c}]$ and $c\in\{R,G,B\}$.

$\tilde{\mu}$ and $\tilde{\Sigma}$ denote uncertain mean vector and covariance matrix respectively. Because, there is no prior knowledge about the parameter uncertainty, practically Zeng et al. [29] assume that the mean and standard deviation vary within intervals with uniform possibilities, i.e., $\tilde{\mu}\in[\underline{\mu},\overline{\mu}]$ or $\tilde{\sigma}\in[\underline{\sigma},\overline{\sigma}]$. Each exponential component in (3) and (4) is the Gaussian primary membership function (MF) with uncertain mean or standard deviation as shown in Fig. 1. The shaded region is the footprint of uncertainty (FOU). The thick solid and dashed lines denote the lower and upper MFs. In the Gaussian primary MF with uncertain mean, the upper MF is:

$$\begin{aligned}
&\overline{h}(o)=f(o;\underline{\mu};\sigma) && \text{if } o<\underline{\mu}\\
&\overline{h}(o)=1 && \text{if } \underline{\mu}\le o\le\overline{\mu}\\
&\overline{h}(o)=f(o;\overline{\mu};\sigma) && \text{if } o>\overline{\mu}
\end{aligned} \quad (5)$$

where

$$f(o,\underline{\mu},\sigma)=\exp-\frac{1}{2}\left(\frac{o-\underline{\mu}}{\sigma}\right)^2$$

and

$$f(o,\overline{\mu},\sigma)=\exp-\frac{1}{2}\left(\frac{o-\overline{\mu}}{\sigma}\right)^2$$

The lower MF is:

$$\begin{aligned}
&\underline{h}(o)=f(o;\overline{\mu};\sigma) && \text{if } o\le\frac{\underline{\mu}+\overline{\mu}}{2}\\
&\underline{h}(o)=f(o;\underline{\mu};\sigma) && \text{if } o>\frac{\underline{\mu}+\overline{\mu}}{2}
\end{aligned} \quad (6)$$

In the Gaussian primary MF with uncertain standard deviation, the upper MF is $\underline{h}(o)=f(o;\mu;\overline{\sigma})$ and lower MF is $\overline{h}(o)=f(o;\mu;\underline{\sigma})$.

The factor k_m and k_v control the intervals in which the parameter vary as follows:

$$\underline{\mu}=\mu-k_m\sigma,\ \overline{\mu}=\mu+k_m\sigma \quad (7)$$

$$\underline{\sigma}=k_v\sigma,\ \overline{\sigma}=(1/k_v)\sigma \quad (8)$$

Because a one-dimensional gaussian has 99.7% of its probability mass in the range of $[\mu-3\sigma,\mu+3\sigma]$, Zeng et al [29] constrain $k_m\in[0,3]$ and $k_v\in[0.3,1]$. These factors also control the area of the FOU. The bigger k_m or k_v, the larger the FOU which implies the greater uncertainty.

Both the T2-FGMM-UM and T2-FGMM-UV can be used to model the background and we can expect that the T2-FGMM-UM will be more robust than the T2-FGMM-UV. Indeed, in the GMM maintenance, only the means are estimated and tracked correctly. The variance and the weights are unstable and unreliable as explained by Greiffenhagen et al. [60].

B. Training

Training the T2-FGMM consist to estimate the parameters μ , Σ and the factor k_m or k_v. Zeng et al. [29] set the factors k_m or k_v as constants according to prior knowledge and then in our application are fixed depending to the video. Thus, parameters estimation of T2 FGMM includes three steps:

- Step 1: Choose K between 3 and 5.
- Step 2: Estimate GMM parameters by an EM algorithm
- Step 3: Add the factor k_m or k_v to GMM to produce T2- FGMM-UM or T2-FGMM-UV.

Once the training is made, a first foreground detection can be processed.

C. Foreground Detection

Foreground detection consist to classify current pixel as background or foreground. By using the ratio $r_j = \omega_j / \sigma_j$, we firstly ordered the K Gaussians as in [13]. This ordering supposes that a background pixel corresponds to a high weight with a weak variance due to the fact that the background is more present than moving objects and that its value is practically constant. The first B Gaussian distributions which exceed certain threshold T are retained for a background distribution:

$$B = \arg\min_b \left(\sum_{i=1}^{b} \omega_{i,t} > T \right) \qquad (9)$$

The other distributions are considered to represent a foreground distribution. When the new frame incomes at times t+1, a match test is made for each pixel. For this, we use the log-likelihood, and thus we are only concerned with the length between two bounds of the log-likelihood interval, i.e., $H(X_t) = \left| \ln(\underline{h}(X_t)) - \ln(\overline{h}(X_t)) \right|$. In Fig 1. (a), the Gaussian primary MF with uncertain mean has:

$$\begin{aligned} & H(X_t) = 2k_m \left| X_t - \mu \right| / \sigma \\ & \text{if } X_t \le \mu - k_m\sigma \text{ or } X_t \ge \mu + k_m\sigma \\ & H(X_t) = \left| X_t - \mu \right|^2 / 2\sigma^2 + k_m \left| X_t - \mu \right| / \sigma + k_m^2 / 2 \\ & \text{if } \mu - k_m\sigma < X_t < \mu + k_m\sigma \end{aligned} \qquad (10)$$

In Fig 1.(b), the Gaussian primary MF with uncertain standard deviation has

$$H(X_t) = (1/k_v^2 - k_v^2) \left| X_t - \mu \right|^2 / 2\sigma^2 \qquad (11)$$

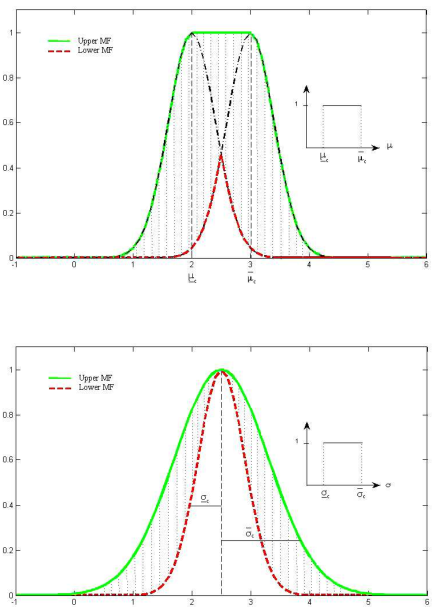

Fig. 1. The upper image shows the Gaussian primary MF with uncertain mean and the lower image shows the Gaussian primary MF with uncertain std. Both of them have uniform possibilities. The hatched region is the FOU. The thick solid and dashed lines denote the upper and the lower MFs respectively.

μ and σ are the mean end the std of the original certain T1 MF without uncertainty. Both (10) and (11) are increasing functions in terms of the deviation $\left| X_t - \mu \right|$. For example, given a fixed k_m, the farther the X_t deviates from μ , the larger $H(X_t)$ is in (12), which reflects a higher extent of the likelihood uncertainty. This relationship accords with the outlier analysis. If the outlier X_t deviates farther from the center of the class-conditional distribution, it has a larger $H(X_t)$ showing its greater uncertainty to the class model. So, a pixel is ascribed to a Gaussian if:

$$H(X_t) < k\sigma \qquad (12)$$

where k is a constant threshold equal to 2.5. Then, two cases can occurs: (1) A match is found with one of the K Gaussians. In this case, if the Gaussian distribution is identified as a background one, the pixel is classified as background else the pixel is classified as foreground. (2) No match is found with any of the K Gaussians. In this case, the pixel is classified as foreground. At this step, a binary mask is obtained. Then, to make the next foreground detection, the parameters must be updated.

D. Maintenance

The T2-FGMM Maintenance is made as in the original GMM [13] as follows:

- Case 1: A match is found with one of the K Gaussians. For the matched component, the update is done as follows:

$$\omega_{i,t+1} = (1-\alpha)\omega_{i,t} + \alpha \tag{13}$$

where α is a constant learning rate.

$$\mu_{i,t+1} = (1-\rho)\mu_{i,t} + \rho.X_{t+1} \tag{14}$$

$$\sigma^2_{i,t+1} = (1-\rho)\sigma^2_{i,t} + \rho(X_{t+1}-\mu_{i,t+1}).(X_{t+1}-\mu_{i,t+1})^T \tag{15}$$

where $\rho = \alpha.\eta(X_{t+1}, \mu_i, \Sigma_i)$ (16)

For the unmatched components, μ and Σ are unchanged, only the weight is replaced by $\omega_{j,t+1} = (1-\alpha)\omega_{j,t}$ (17).

- Case 2: No match is found with any of the K Gaussians. In this case, the least probable distribution k is replaced with a new one with parameters:

$$\omega_{k,t+1} = \text{Low Prior Weight} \tag{18}$$

$$\mu_{k,t+1} = X_{t+1} \tag{19}$$

$$\sigma^2_{k,t+1} = \text{Large Initial Variance} \tag{20}$$

Once a background maintenance is made, another foreground detection can be processed.

IV. Results And Discussion

We have applied the T2-FGMM-UM and T2-FGMM-UV algorithms in different dynamic backgrounds in the visible and beyond the visible spectrum. In these scenes, the four main types of dynamic backgrounds appear: camera jitter, waving trees, water rippling and water surface. On a 3 GHz Intel Pentium Duo processor with 2 GB RAM, an optimized implementation of the proposed approach can process about 11 fps for a frame size of 240*360. All the algorithms were implemented under Microsoft Visual C++ using the OpenCV library. Firstly, qualitative comparative results with the original mixture of Gaussians method [13] are shown then a quantitative evaluation is presented. For the proposed algorithms, the best results are obtained by setting $k_m = 2$ and $k_v = 0.9$.

A. Qualitative Analysis

Qualitative results on six sequences of dynamic scenes are presented in this section. Four sequences concern outdoor scenes in RGB videos [62] [64] and two sequences in infrared videos[65].

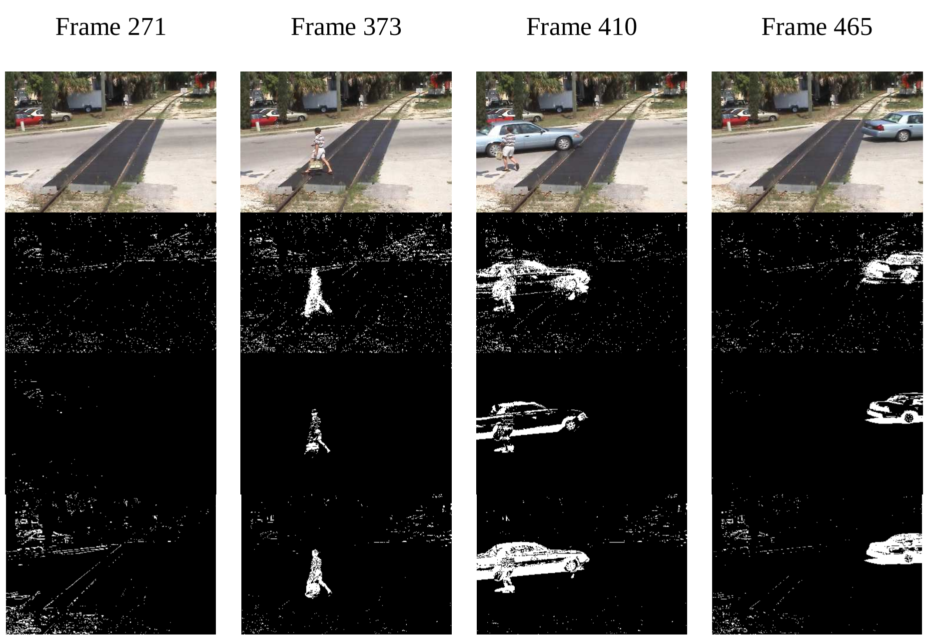

Fig. 2. Camera jitter: The first row are the original images, the second row are the results obtained by using the GMM [13], the third row are the result obtained using the T2-FGMM-UM and the fourth row are the result obtained by using the T2-FGMM-UV.

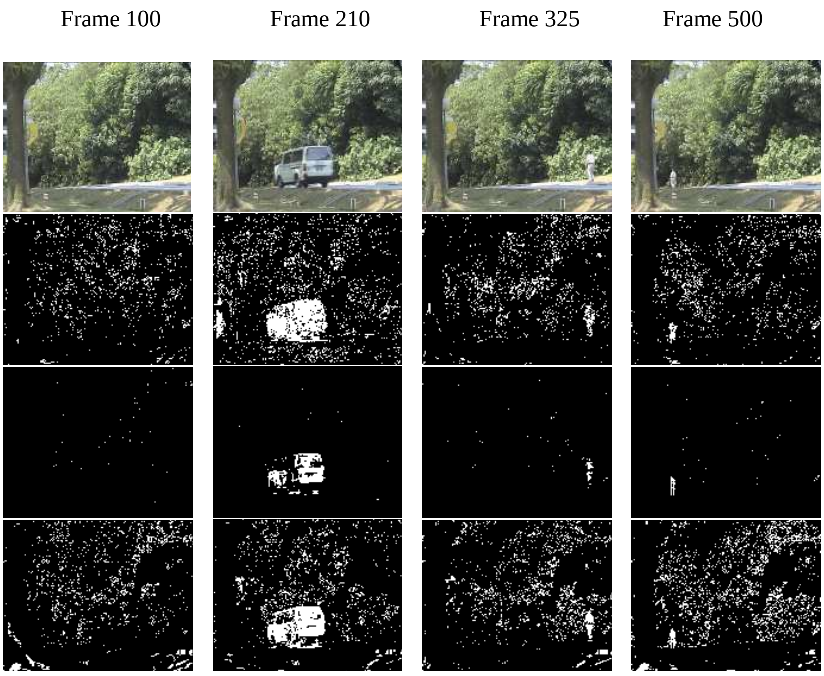

Fig. 3. The first row shows the original frames for Campus sequences. The second row presents the segmented images obtained by the GMM [13]. The third and the fourth rows illustrate the result obtained using the T2-FGMM-UM and the T2-FGMM-UV respectively.

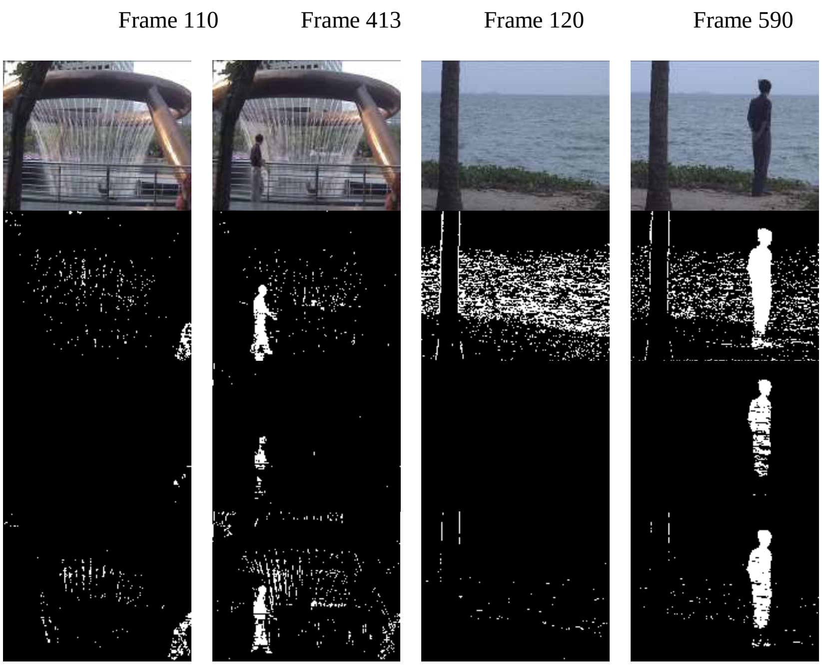

Fig. 4. The first row shows the original frames for Fountain and Water Surface sequences. The second row presents the segmented images obtained by the GMM [13]. The third and the fourth rows illustrate the result obtained using the T2-FGMM-UM and the T2-FGMM-UV respectively.

1) RGB Videos

We have tested our method on the sequence Camera Jitter which comes from [61] downloadable at [62]. We have also tested our method on the sequences Campus, Fountain and Water Surface which come from [63] downloadable at [64]. Fig. 2 shows the robustness of T2-FGMM-UM against camera jitter and respectively Fig. 3 for the waving trees (Campus sequence) and Fig. 4 for the water rippling (Fountain sequence) and water surface (Water Surface sequence). In each case, the T2-FGMM-UM gives the best result followed by the T2-FGMM-UV and the GMM. These different experiments confirm that to take into account the uncertainty using T2-FGMM performs the GMM. Furthermore, the T2-FGMM-UM is more robust than the T2F-GMM-UV like supposed in Section III.

2) Infrared Videos

We have tested the proposed algorithm on the Terravic datasets [65]. We have choosen the two sequences called Uneventful Background Motion because they present dynamic backgrounds as waving vegetations. In this sequence, nothing must be detected. The Fig. 5 shows the result obtained using the GMM, the T2-FGMM-UM and the T2-FGMM-UV on the frame 150 of the sequence IRTR01. The Fig. 6 shows the same experiments on the frame 150 of the sequence IRTR02.

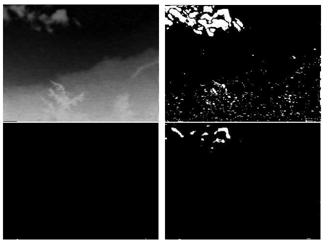

Fig. 5. Sequence IRTR01 - First row: The current image, Result with the GMM [47 Second Row: Result with T2-FGMM-UM, Result with T2-FGMM-UV

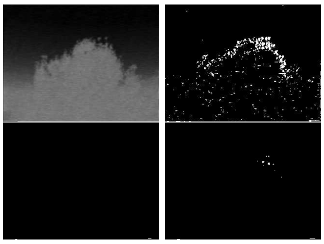

Fig. 6. Sequence IRTR02 - First row: The current image, Result with the GMM [47]. Second Row: Result with T2-FGMM-UM, Result with T2-FGMM-UV

The motion causes substantial false positive detection in the GMM. The more robust is the T2-FGMM-UM followed by the T2-FGMM-UV. These results confirm the robustness of the proposed method in the presence of dynamic backgrounds in infrared videos.

B. Quantitative Analysis

1) RGB Videos

In order to provide a quantitative perspective about the quality of foreground detection with our approach, we have used a test sequence and the corresponding ground-truth segmentation from [61]. This outdoor sequence involved a camera mounted on a tall tripod and was available. The wind caused the tripod to sway back and forth causing nominal motion in the scene. To be more complete, we have compared our method to the GMM, and two improved versions: The first one was developed by Bowden et al. [47] and performed the initialization and maintenance of the GMM's parameters. The second one was proposed by Zivkovic et al. [36] and allowed to adapt the number of Gaussian K overtime. In Fig. 7, the first row shows different current image and the second row shows the corresponding ground truth. The third row shows the results obtained by the Bowden's GMM. The fourth and the fifth rows show respectively the results obtained by the corresponding T2-FGMM-UM and T2-FGMM-UV versions. The sixth row shows the results obtained by the Zivkovic's GMM. The seventh and eigth rows show respectively the results obtained by the corresponding T2-FGMM-UM and T2-FGMM-UV versions. Table II, Table III and Table IV show the performance in term on False Positive (FP), False Negative (FN) and Total Error (TE) and Fig. 8 show graphically these results. In each case, the T2-FGMM-UM and T2-FGMM-UV give less error than the corresponding original crisp version and the T2FGMM-UM appears the best algorithm for object/target detection.

2) Infrared Videos

For the evaluation beyond the visible spectrum, we have used the Dataset 01: OSU Thermal Pedestrian Database which comes from the OTCBVS 2009 dataset [66]. The Fig. 9 shows the results obtained on the Sequence 1 using the GMM [47], the T2-FGMM-UM and the T2-FGMM-UV on the frame 27. Silhouettes are well detected by the three algorithms but the T2-FGMM-UM gives less false detection followed by the T2-FGMM-UV and the crisp GMM. Then, to evaluate quantitatively our method, we have used the similarity measure derived by Li et al. [63]. Let A be a detected region and B be the corresponding ground truth, the similarity between A and B can be defined as:

$$S(A,B) = \frac{A \cap B}{A \cup B} \tag{21}$$

If A and B are the same, S (A,B) approaches 1, otherwise 0 i.e. A and B have the least similarity. The ground truth is marked manually. Table V shows similarity value obtained for this experiment. It confirms the qualitative evaluation.

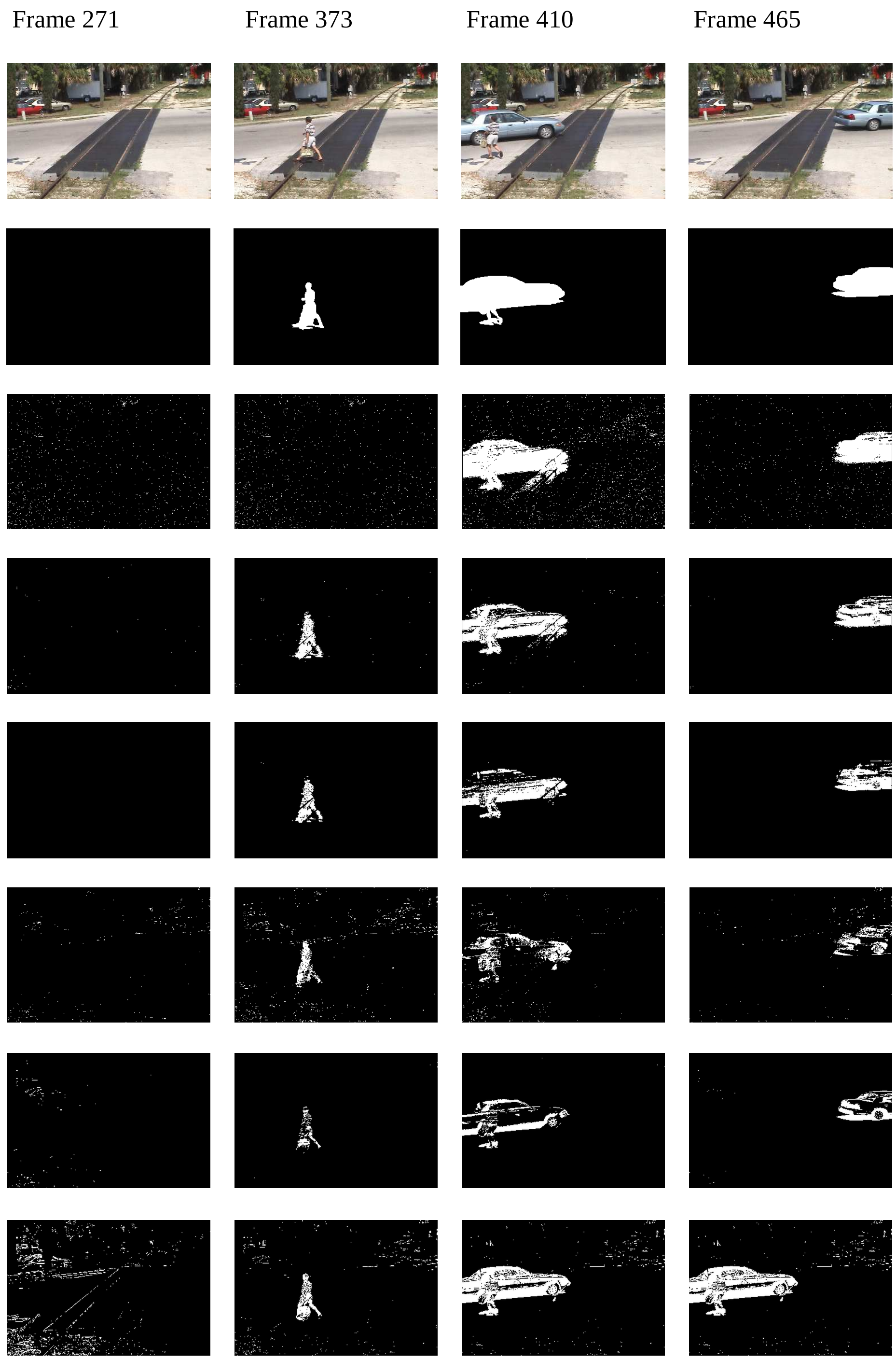

Fig. 7. Camera jitter. The first row shows the original images and the second row the corresponding ground truth. In the third, the fourth and the fifth rows respectively, the segmented images by the Bowden's GMM [47], the T2-FGMM-UM and T2-FGMM-UV versions are shown. The sixth, the seventh and the eigth rows present respectively the segmented images by the Zivkovic's GMM [36], the T2 FGMM-UM and T2 FGMM-UV versions.

TABLE II
PERFORMANCE ANALYSIS WITH GMM FROM STAUFFER AND GRIMSON [13]

Method	Error Type	Fr. 271	Fr. 373	Fr 410	Fr 465	Total Error
GMM Stauffer [13]	FN	0	1120	4818	2050	
	FP	2093	4124	2782	1589	18576
T2-GMM-UM	FN	0	1414	6043	2520	
	FP	203	153	252	46	10631
T2-GMM-UV	FN	0	957	2217	1069	
	FP	3069	1081	1119	1158	10670

TABLE III
PERFORMANCE ANALYSIS WITH THE MODIFIED GMM FROM BOWDEN ET AL. [47]

Method	Error Type	Fr. 271	Fr. 373	Fr 410	Fr 465	Total Error
GMM Bowden et al.[47]	FN	0	265	637	413	
	FP	1034	1359	3308	814	7830
T2-GMM-UM	FN	0	522	2179	1251	
	FP	37	287	787	122	5185
T2-GMM-UV	FN	0	757	4130	1818	
	FP	0	162	252	21	7140

TABLE IV
PERFORMANCE ANALYSIS WITH THE MODIFIED GMM FROM ZIVKOVIC [36]

Method	Error Type	Fr. 271	Fr. 373	Fr 410	Fr 465	Total Error
GMM Zivkovic [36]	FN	0	1152	6688	3009	
	FP	341	1404	1077	310	13981
T2-GMM-UM	FN	8	1414	6043	2520	
	FP	204	154	252	48	10635
T2-GMM-UV	FN	0	957	2216	1068	
	FP	3072	1082	1119	1166	10680

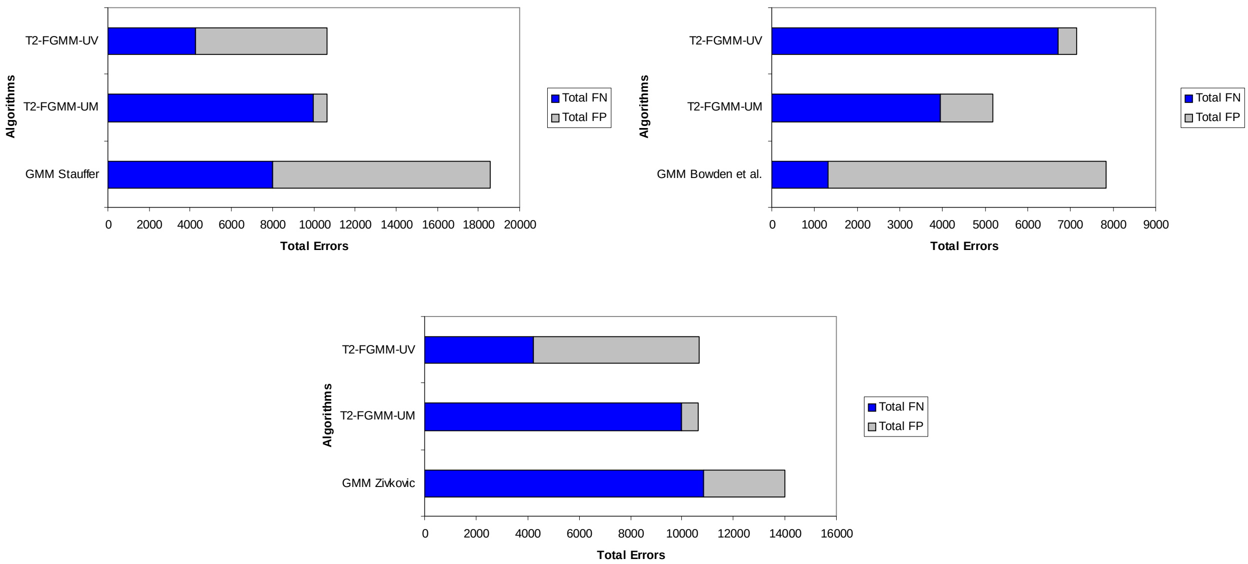

Fig. 8. Performance analysis The upper left image shows the performance of the GMM [13] and their corresponding T2-FGMM-UM and T2-FGMM-UV versions. The upper right image shows the performance of the Bowden's GMM [47], the T2-FGMM-UM and T2-FGMM-UV versions. The lower image the shows the performance of Zivkovic's GMM [36], the T2 FGMM-UM and T2 FGMM-UV versions.

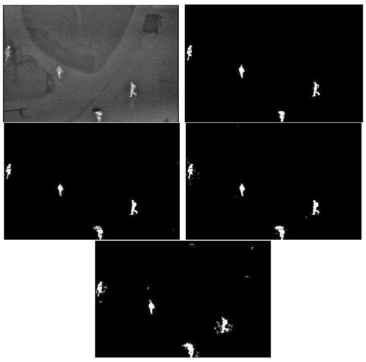

Fig. 9. Sequence OSU - First row: The current image, the ground truth. Second row: Results with T2-FGMM-UM and the T2-FGMM-UV respectively. Third row: Result with the GMM [47].

TABLE V
PERFORMANCE ANALYSIS IN INFRARED VIDEOS

Method	T2-FGMM-UM	T2-FGMM-UV	GMM [47]
$S(A,B)$	48%	43%	36%

V. CONCLUSION

In this work, we have modeled the background by using a Type-2 Fuzzy Gaussian Mixture Models. Experimental validations on RGB and infrared videos show very satisfactory performance and more robustness than the crisp GMM in difficult environments. The proposed approach also addresses most of the four main dynamic backgrounds: Camera jitter, waving trees, water rippling and water surface. The T2-FGMM-UM is more robust than the T2-FGMM-UV due to a better estimation of the mean than the variance.

This work confirms the pertinence of the fuzzy concepts in the field of background subtraction. Future works concern applications of fuzzy concepts in other steps of background subtraction like in the background initialization, the background maintenance and the foreground detection.

REFERENCES

[1] S. Cheung and C. Kamath, “Robust Background Subtraction With Foreground Validation for Urban Traffic Video”, EURASIP 2005, New York, USA, 2005.

[2] J. Carranza, C. Theobalt. M. Magnor and H. Seidel, “Free-Viewpoint Video of Human Actors”, ACM Transactions on Graphics, Vol. 22, Issue3, pp. 569-577, 2003.

[3] T. Horprasert, I. Haritaoglu, C. Wren, D. Harwood, L. Davis and A. Pentland ,”Real-time 3D Motion Capture”, Workshop on Perceptual User Interfaces, PUI 1998, San Francisco, California, pp. 87-90, Nov. 1998.

[4] J. Davis and V. Sharma, “Background-Subtraction in Thermal Imagery Using Contour Saliency”, International Journal of Computer Vision, 2006.

[5] L. Latecki, R. Miezianko and D. Pokrajac, “Tracking Motion Objects in Infrared Videos”, IEEE International Conference on Advanced Video and Signal Based Surveillance, AVSS 2005, Como, Italy, Sept. 2005.

[6] B. Bhanu and I. Pavlidis, “Computer Vision beyond the visible spectrum”, Springer Verlag, Nov. 2004.

[7] M. Piccardi, “Background subtraction techniques: a review”, the International Conference on Systems, Man and Cybernetics, SMC 2004, pp. 3199-3104, The Hague, Oct. 2004.

[8] S. Elhabian, K. El-Sayed and S. Ahmed, “Moving Object Detection in Spatial Domain using Background Removal Techniques - State-of-Art”, Recent Patents on Computer Science, Vol. 1, Number 1, pp. 32-54, Jan. 2008.

[9] B. Lee and M. Hedley, “Background Estimation for Video Surveillance”, IVCNZ 2002, pp. 315-320, 2002.

[10] N. McFarlane and C. Schofield, “Segmentation and tracking of piglets in images”, BMVA 1995, pp. 187-193, 1995.

[11] J. Zheng, Y. Wang, N. Nihan and E. Hallenbeck, “Extracting Roadway Background Image: A mode based approach”, Journal of Transportation Research Report, No 1944, pp. 82-88, Mar. 2006.

[12] C. Wren, A. Azarbayejani, T. Darrell and A. Pentland, “Pfinder : Real-Time Tracking of the Human Body”, PAMI 1997, Vol. 19, No. 7, pp. 780 –785 , Jul. 1997.

[13] C. Stauffer and W. Grimson, “Adaptive background mixture models for real-time tracking”, IEEE Conference on Computer Vision and Pattern Recognition, CVPR 1999, pp. 246-252, 1999.

[14] A. Elgammal, D. Harwood and L. Davis, “Non-parametric Model for Background Subtraction”, ECCV 2000, pp. 751-767, Dublin, Ireland, Jun. 2000.

[15] K. Toyama, J. Krumm, B. Brumitt and B. Meyers, ”Wallflower: Principles and Practice of Background Maintenance”, International Conference on Computer Vision, pp. 255-261, Corfu, Greece, Sept. 1999.

[16] S. Messelodi, C. Modena, N. Segata and M. Zanin, “A Kalman filter based background updating algorithm robust to sharp illumination changes”, ICIAP 2005, Vol. 3617, pp. 163-170, Cagliari, Italy, Sept. 2005.

[17] R. Chang, T. Ghandi and M. Trivedi, “Vision modules for a multi sensory bridge monitoring approach”, ITSC 2004, pp. 971-976, Oct. 2004.

[18] N. Oliver, B. Rosario and A. Pentland, “A Bayesian Computer Vision System for Modeling Human Interactions”, International Conference on Vision Systems, ICVS 1999, Gran Canaria, Spain, Jan. 1999.

[19] H. Lin, T. Liu and J. Chuang, “A probabilistic SVM approach for background scene initialization”, ICIP 2002, Vol. 3, pp. 893-896, Rochester, New York, Sept. 2002.

[20] J. Wang, G. Bebis, and R. Miller, “Robust Video-Based Surveillance by Integrating Target Detection with Tracking”, IEEE Workshop on Object Tracking and Classification Beyond the Visible Spectrum in conjunction with CVPR 2006, New York, NY, Jun. 2006.

[21] A. Tavakkoli, M. Nicolescu and G. Bebis, “A Novelty Detection Approach for Foreground Region Detection in Videos with Quasi-stationary Backgrounds”, ISVC 2006, pp. 40-49, Lake Tahoe, USA, Nov. 2006.

[22] H. Kim, R. Sakamoto, I. Kitahara, T. Toriyama and K. Kogure, “Robust Silhouette Extraction Technique Using Background Subtraction”, 10th Meeting on Image Recognition and Understand, MIRU 2007, Hiroshima, Japan, Jul. 2007.

[23] M. Allili, N. Bouguila and D. Ziou, “A Robust Video Foreground Segmentation by Using Generalized Gaussian Mixture Modeling”, Fourth Canadian Conference on Computer and Robot Vision, CRV 2007, pp. 503-509, 2007.

[24] M. Yamazaki, G. Xu and Y. Chen, “Detection of Moving Objects by Independent Component Analysis”, ACCV 2006, pp. 467-478, 2006.
[25] D. Tsai and C. Lai, “Independent Component Analysis-Based Background Subtraction for Indoor Surveillance”, IEEE Transactions on Image Processing, IP 2009, Vol. 18, Issue 1, 2009.
[26] S. Bucak, B. Gunsel and O. Gursoy, “Incremental Non-negative Matrix Factorization for Dynamic background Modelling”, International Workshop on Pattern Recognition in Information Systems, Funchal, Portugal, Jun. 2007
[27] S. Bucak and B. Gunsel, “Incremental Subspace Learning and Generating Sparse Representations via Non-negative Matrix Factorization”, Pattern Recognition, Vol. 42, Issue 5, pp. 788-797, 2009.
[28] X. Li, W. Hu, Z. Zhang and X. Zhang, “Robust Foreground Segmentation Based on Two Effective Background Models”, MIR 2008, pp. 223-228, Vancouver, Canada, Oct. 2008.
[29] J. Zeng, L. Xie and Z. Liu, “Type-2 Fuzzy Gaussian Mixture”, Pattern Recognition, Vol. 41, Issue 2, pp. 3636-3643, Dec. 2008.
[30] B. Han and X. Lin, “Update the GMMs via adaptive Kalman filtering”, Proceedings of SPIE - The International Society for Optical Engineering, Vol. 5960, Issue 3, pp. 1506-1515, 2005.
[31] H. Yang, Y. Tan, J. Tian and J. Liu, “Accurate dynamic scene model for moving object detection”, International Conference on Image Processing , ICIP 2007, Vol. VI, pp. 157-160, 2007.
[32] W. Zhang , X. Fang, X. Yang and Q. Wu, “Spatiotemporal Gaussian mixture model to detect moving objects in dynamic scenes”, Journal of Electronic Imaging, Vol. 16, Issue 2, April 2007.
[33] P. Tang, L. Gao and Z. Liu, “Salient Moving Object Detection Using Stochastic Approach Filtering”, Fourth International Conference on Image and Graphics, ICIG 2007, pp. 530-535, 2007.
[34] M. Harville, “A framework for high-level feedback to adaptive, per-pixel, mixture-of-Gaussian background models”, 7th European Conference on Computer Vision, ECCV 2002, pp. 543 -560, Copenhagen, Denmark, May 2002.
[35] M. Cristani and V. Murino, “A spatial sampling mechanism for effective background subtraction”, VISAPP 2007, Vol. 2, pp. 403-410, Barcelona, Spain, Mar. 2007.
[36] Z. Zivkovic, “Improved adaptive Gaussian mixture model for background subtraction”, International Conference Pattern Recognition, ICPR 2004, Vol. 2, pp. 28-31, 2007.
[37] J. Cheng , J. Yang, Y. Zhou and Y. Cui, “Flexible background mixture models for foreground segmentation”, Journal of Image and Vision Computing, Vol. 24, pp. 473-482, 2006.
[38] R. Tan, H. Huo, J. Qian and T. Fang , “Traffic Video Segmentation using Adaptive-K Gaussian Mixture Model”, The International Workshop on Intelligent Computing , IWICPAS 2006, Xi'An, China, August 2006.
[39] L. Carminati and J. Benois-Pinau, “Gaussian Mixture Classification for Moving Object Detection in Video Surveillance Environment“, IEEE International Conference on Image Processing, ICIP 2005, pp. 113-116, Sept. 2005.
[40] C. Cuevas, L. Salgado and N. Garcia, “A new strategy based on adaptive mixture of Gaussians for real-time moving objects segmentation”, Real Time image Processing, SPIE 2008, Vol. 6811, Jan. 2008.
[41] Y. Wang, Y. Tan and J. Tian, “Video segmentation algorithm with Gaussian Mixture Model and shadow removal”, Journal of Opto Electronic Engineering, Guangdian Gongcheng, Vol. 35, Issue 3, pp. 21-25, Mar. 2008
[42] V. Morellas, I. Pavlidis and P. Tsiamyrtzis, “DETER: detection of events for threat evaluation and recognition”, Machine Vision and Applications, Vol. 15, pp. 29-45, Jun. 2003.
[43] D. Lee, “Online Adaptive Gaussian Mixture Learning for Video Applications”, Workshop on Statistical Methods for Video Processing, SMVP 2004, pp. 105-116, Prague, May 2004.
[44] Y. Zhang, Z. Liang, Z. Hou, H. Wang and M. Tan, “An Adaptive Mixture Gaussian Background Model with Online Background Reconstruction and Adjustable Foreground Mergence Time for Motion Segmentation”, ICIT 2005, pp. 23-27, Dec. 2005.
[45] Q. Zang and R. Klette, “Evaluation of an Adaptive Composite Gaussian Model in Video Surveillance”, CITR Technical Report 114, Auckland University, Aug. 2002.
[46] B. White and M. Shah, “Automatically Tuning Background Subtraction Parameters Using Particle Swarm Optimization”, IEEE International Conference on Multimedia & Expo, ICME 2007, Beijing, China, 2007.
[47] P. KaewTraKulPong and R. Bowden, “A Real-Time Adaptive Visual Surveillance System for Tracking Low Resolution Color Targets In Dynamically Changing Scenes”, Journal of Image and Vision Computing, Vol. 21, Issue 10, pp. 913-929, Sept. 2003.
[48] D. Lee, “Improved Adaptive Mixture Learning for Robust Video Background Modeling”, IAPR Workshop on Machine Vision for Applications, Nara, Japan, pages 443-446, December 2002.
[49] P. Withagen, F. Groen and K. Schutte, “EMswitch: a multi-hypothesis approach to EM background modeling”, Proceedings of the IEEE Advanced Concepts for Intelligent Vision Systems, ACIVS 2003, pp. 199-206 Sept. 2003.
[50] P. Kumar and K. Sengupta, “Foreground background segmentation using temporal and spatial markov processes”, Department of Electrical and Computer Engineering, National University of Singapore, Nov. 2000.
[51] Y. Sun and B. Yuan, “Hierarchical GMM to handle sharp changes in moving object detection”, Electronic Letters, Vol. 40, No 13, Jun. 2004.
[52] M. Cristani, M. Bicego and V. Murino, “Integrated Region- and Pixel-based Approach to Background Modeling”, IEEE Workshop on Motion and Video Computing, MOTION 2002, pp. 3-8, 2002
[53] J. Hu and T. Su, “Robust Background Subtraction with Shadow and Highlight Removal for Indoor Surveillance”, Journal on Advances in Signal Processing, EURASIP 2007, Vol. 2007, pp. 1-14, 2007.
[54] Y. Sun, “Better Foreground Segmentation for Static Cameras via New Energy Form and Dynamic Graph-cut”, 18th International Conference on Pattern Recognition, ICPR 2006, pp. 49-52, 2006
[55] F. Porikli and O. Tuzel, “Bayesian Background Modeling for Foreground Detection”, ACM International Workshop on Video Surveillance and Sensor Networks, VSSN 2005, pp. 55-28, Nov. 2005.
[56] D. Parks and S. Fels, “Evaluation of Background Subtraction Algorithms with Post-processing”, IEEE International Conference on Advanced Video and Signal-based Surveillance, AVSS 2008, Sept. 2008.
[57] T. Bouwmans, F. El Baf and B. Vachon, “Background Modeling using Mixture of Gaussians for Foreground Detection – A survey”, Recent Patents on Computer Science, Vol. 1, No 3, pp. 219-237, Nov. 2008.
[58] D. Tran, T. Van Le and M. Wagner, “Fuzzy Gaussian mixture models for speaker recognition”, International Conference Spoken Language Processing, pp. 759–762, 1998.

[59] J. Mendel, "Type-2 fuzzy sets: Some questions and answers", IEEE Connections, Newsletter of the IEEE Neural Networks Society, pp. 10–13, 2003.
[60] M. Greiffenhagen, V. Ramesh and H. Niemann, "The Systematic Design and Analysis Cycle of a Vision System: A Case Study in Video Surveillance", IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2001, Vol. 2, p. 704, 2001.
[61] Y. Sheikh and M. Shah, "Bayesian Modeling of Dynamic Scenes for Object Detection", IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI 2005, Vol. 27, No. 11, pp. 1778-1792, Nov. 2005.
[62] http://www.cs.ucf.edu/~yaser/backgroundsub.htm
[63] L. Li, W. Huang, I. Gu and Q.Tian, "Statistical Modeling of Complex Background for Foreground Object Detection", IEEE transaction image processing, Vol. 13, Issue 11, pp. 1459-1472, Nov. 2004.
[64] http://perception.i2r.a-star.edu.sg/bk_model/bk_index.html
[65] R. Miezianko. IEEE OTCBVS WS series bench. terravic researchinfrared database.
[66] http://www.cse.ohio-state.edu/otcbvs-bench/

T. Bouwmans received his PhD in Image Processing from the University of Littoral Côte d'Opale, Calais, France in 1997. From 2000 and 2007, he was a member of the Laboratory L3i at the University of la Rochelle. During this period, he was a member of the Aqu@theque project. Since 2008, he is a member of the laboratory MIA (Mathematic, Images and Applications). Currently, his research concern moving object detection in videos sequences by using fuzzy concepts in background subtraction.

F. El Baf received her Master in image and calculus from the University of La Rochelle, La Rochelle, France in 2005 and her PhD in Mathematics applied to Image Processing in June 2009. Her research concerns the application of fuzzy concepts in the field of background subtraction.