Type-2 Fuzzy Mixture of Gaussians Model: Application to Background Modeling

Fida El Baf, Thierry Bouwmans, Bertrand Vachon

Type-2 Fuzzy Mixture of Gaussians Model: Application to Background Modeling

Fida El Baf, Thierry Bouwmans, Bertrand Vachon

Department of Mathematics, Images and Applications
University of La Rochelle, France
felbaf@univ-lr.fr

Abstract. Background modeling is a key step of background subtraction methods used in the context of static camera. The goal is to obtain a clean background and then detect moving objects by comparing it with the current frame. Mixture of Gaussians Model [1] is the most popular technique and presents some limitations when dynamic changes occur in the scene like camera jitter, illumination changes and movement in the background. Furthermore, the MGM is 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. To take into account this uncertainty, we propose to use a Type-2 Fuzzy Mixture of Gaussians Model. Results show the relevance of the proposed approach in presence of camera jitter, waving trees and water rippling.

1 Introduction

The common approach for discriminating moving objects from the background is the background subtraction which is used in the field of video surveillance [2], optical motion capture [3–5] and multimedia applications [6]. In this context, background modeling is the first key step to obtain a clean background. The simplest way to model the background is to acquire a background image which doesn't include any moving object. In some environments, the background isn't available and can always be changed under critical situations like camera jitter, illumination changes, objects being introduced or removed from the scene. To take into account these problems of robustness and adaptation, many background modeling methods have been developed and the most recent surveys can be found in [2, 7, 8]. These background modeling methods can be classified in the following categories: Basic Background Modeling [9–11], Statistical Background Modeling [12, 1, 13], Fuzzy Background Modeling [14, 15] and Background Estimation [16–18]. The models the most used are the statistical ones: 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 [9]. 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 Models (MGM) has been used to model dynamic backgrounds [1]. 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 was

developed for estimating background probabilities at each pixel from many recent samples over time using Kernel density estimation [13] but it is time consuming. Finally, due to a good compromise between robustness and time/memory requirements MGM are the most used.

In the MGM initialization, an expectation-maximization (EM) algorithm is used and allows to estimate MGM parameters from a training sequence according to the maximum-likelihood (ML) criterion. The MGM is completely certain once its parameters are specified. However, because of insufficient or noisy data in training sequence, the MGM may not accurately reflect the underlying distribution of the observations according to the ML estimation. It may seem problematical to use likelihoods that are themselves precise real numbers to evaluate MGM with uncertain parameters. To solve this problem, we propose to model the background by using a Type-2 Fuzzy Mixture of Gaussians Model (T2 FMGM). T2 FMGM was recently developed by Zeng et al. [19] to introduce descriptions of uncertain parameters in the MGM. T2 FMGM has proved their superiority in pattern classification [19].

The rest of this paper is organized as follows: In the section 2, we presented the basis of T2 FMGM. Then, the T2 FMGM is used for background modeling in section 3. In the section 4, experiments on indoor and outdoor scenes show that T2 FMGM outperform the crisp MGM when dynamic changes occurs.

2 Type-2 Fuzzy Mixture of Gaussians Model

This section reviews the principle of T2 FMGM developed in [19]. Mixture of Gaussians Models (MGMs) have been widely used in density modelling and clustering. They have universal approximation ability because they can model any density function closely provided that they contain enough mixture components. The MGM is composed of K mixture components of multivariate Gaussian as follows:

$$P(o) = \sum_{i=1}^{k} \omega_i \eta(o, \mu_i, \sum_i) \tag{1}$$

where $\sum_{i=1}^{k} \omega_i = 1$ and $\omega_i > 0$. The multivariate Gaussian distribution is:

$$\eta(o, \mu, \sum) = \frac{1}{(2\pi)^{\frac{d}{2}} |\sum|^{\frac{1}{2}}} exp\left(-\frac{1}{2}(o-\mu)^T \sum^{-1} (o-\mu)\right) \tag{2}$$

where μ is the mean vector, $\sum$ is the covariance matrix and d is the dimensionality of o. For the sake of simplicity, $\sum$ is considered as a diagonal covariance matrix. So, the MGM (Equation 1) is expressed as a linear combination of multivariate Gaussian distribution.

To take into account the uncertainty, Zeng et al. [19] proposed recently T2 membership functions to represent multivariate Gaussian with uncertain mean vector or covariance matrix, and replace the corresponding parts in (Equation 1) to produce the T2 FMGM with uncertain mean vector (T2 FMGM-UM) or uncertain variance (T2 FMGM-UV). Given a $d-$dimensional observation vector o, the mean vector μ, and the diagonal

covariance matrix $\sum = diag\left(\sigma_1^2, \ldots, \sigma_d^2\right)$, the multivariate Gaussian with uncertain mean vector is:

$$\eta\left(o, \tilde{\mu}, \sum\right) = \frac{1}{(2\pi)^{\frac{d}{2}} |\sum|^{\frac{1}{2}}} exp\left[-\frac{1}{2}\left(\frac{o_1 - \mu_1}{\sigma_1}\right)^2\right] \ldots exp\left[-\frac{1}{2}\left(\frac{o_d - \mu_d}{\sigma_d}\right)^2\right] \quad (3)$$
$$\mu_1 \in \left[\underline{\mu}_1, \overline{\mu}_1\right], \ldots, \mu_d \in \left[\underline{\mu}_d, \overline{\mu}_d\right]$$

In the case of uncertain covariance matrix, it's defined as:

$$\eta\left(o, \mu, \tilde{\sum}\right) = \frac{1}{(2\pi)^{\frac{d}{2}} |\sum|^{\frac{1}{2}}} exp\left[-\frac{1}{2}\left(\frac{o_1 - \mu_1}{\sigma_1}\right)^2\right] \ldots exp\left[-\frac{1}{2}\left(\frac{o_d - \mu_d}{\sigma_d}\right)^2\right] \quad (4)$$
$$\sigma_1 \in \left[\underline{\sigma}_1, \overline{\sigma}_1\right], \ldots, \mu_d \in \left[\underline{\sigma}_d, \overline{\sigma}_d\right]$$

where $\tilde{\mu}$ 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. [19] assume that the mean and standard deviation vary within intervals with uniform possibilities, i.e., $\mu \in \left[\underline{\mu}, \overline{\mu}\right]$ or $\sigma \in \left[\underline{\sigma}, \overline{\sigma}\right]$. Each exponential component in Equation 3 and Equation 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:

$$\overline{h}\left(o\right) = \begin{cases} f\left(o; \underline{\mu}, \sigma\right), & \text{if } o < \underline{\mu} \\ 1, & \text{if } \underline{\mu} \leq o < \overline{\mu} \\ f\left(o; \overline{\mu}, \sigma\right), & \text{if } o > \overline{\mu} \end{cases} \quad (5)$$

where $f\left(o; \underline{\mu}, \sigma\right) = exp\left[-\frac{1}{2}\left(\frac{o-\underline{\mu}}{\sigma}\right)^2\right]$ and $f\left(o; \overline{\mu}, \sigma\right) = exp\left[-\frac{1}{2}\left(\frac{o-\overline{\mu}}{\sigma}\right)^2\right]$.
The lower MF is:

$$\underline{h}\left(o\right) = \begin{cases} f\left(o; \overline{\mu}, \sigma\right), & \text{if } o \leq \frac{\underline{\mu}+\overline{\mu}}{2} \\ f\left(o; \underline{\mu}, \sigma\right), & \text{if } o > \frac{\underline{\mu}+\overline{\mu}}{2} \end{cases} \quad (6)$$

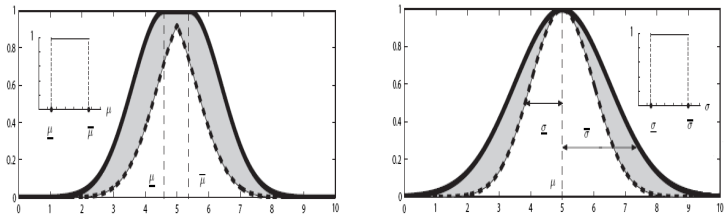
In the Gaussian primary MF with uncertain standard deviation, the upper MF is $\underline{h}\left(o\right) = f\left(o; \mu, \overline{\sigma}\right)$ and the lower MF is $\overline{h}\left(o\right) = f\left(o; \mu, \underline{\sigma}\right)$.

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

$$\underline{\mu} = \mu - k_m\sigma, \quad \overline{\mu} = \mu + k_m\sigma, \quad k_m \in \left[0, 3\right], \quad (7)$$

$$\underline{\sigma} = k_\nu\sigma, \quad \overline{\sigma} = \frac{1}{k_\nu}\sigma, \quad k_\nu \in \left[0.3, 1\right]. \quad (8)$$

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. [19] constrain $k_m \in [0, 3]$ and $k_\nu \in [0.3, 1]$. These factors also control the area of the FOU. The bigger k_m or the smaller k_ν, the larger the FOU, which implies the greater uncertainty.



a) b)

Fig. 1. a): At the left, the Gaussian primary MF with uncertain mean. b): At the right, the Gaussian primary MF with uncertain std having uniform possibilities. The shaded region is the FOU. The thick solid and dashed lines denote the lower and upper MFs. Picture from [19]

3 Application to Background Modeling

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 MGM is composed of K mixture components of multivariate Gaussian as follows:

$$P\left(X_t\right) = \sum_{i=1}^{k} \omega_{i,t} \eta\left(X_t, \mu_{i,t}, \sum_{i,t}\right) \tag{9}$$

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 $\sum_{i,t}$. η is a Gaussian probability density function:

$$\eta\left(X_t, \mu, \sum\right) = \frac{1}{(2\pi)^{\frac{3}{2}} |\sum|^{\frac{1}{2}}} exp\left(-\frac{1}{2}\left(X_t - \mu\right)^T \sum^{-1} \left(X_t - \mu\right)\right) \tag{10}$$

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

$$\eta\left(X_t, \tilde{\mu}, \sum\right) = \frac{1}{(2\pi)^{\frac{3}{2}} |\sum|^{\frac{1}{2}}} \prod exp\left[-\frac{1}{2}\left(\frac{X_{t,c} - \mu_c}{\sigma_c}\right)^2\right] \tag{11}$$

with $\mu_c \in \left[\underline{\mu}_c, \overline{\mu}_c\right]$ and $c \in \{R, G, B\}$.
For the T2 FMGM-UV, the multivariate Gaussian with uncertain variance vector is:

$$\eta\left(X_t, \mu, \tilde{\sum}\right) = \frac{1}{(2\pi)^{\frac{3}{2}} |\sum|^{\frac{1}{2}}} \prod exp\left[-\frac{1}{2}\left(\frac{X_{t,c} - \mu_c}{\sigma_c}\right)^2\right] \tag{12}$$

where $\sigma_c \in [\underline{\sigma}_c, \overline{\sigma}_c]$ and $c \in \{R, G, B\}$.
Both the T2 FMGM-UM and T2 FMGM-UV can be used to model the background and we can expect that the T2 FMGM-UM will be more robust than the T2 FMGM-UV. Indeed, only the means are estimated and tracked correctly over time in the MGM maintenance. The variance and the weights are unstable and unreliable as explained by Greiffenhagen et al. [20].

3.1 Training

Training a T2 FMGM consists to estimate the parameters μ, $\sum$ and the factor k_m or k_ν. Zeng et al. [19] set the factors k_m and k_ν as constants according to prior knowledge. In our work, they are fixed depending to the video (see Section 4). Thus, parameters estimation of T2 FMGM includes three steps:

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

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

3.2 Foreground Detection

Foreground detection consists in classifying the current pixel as background or foreground. By using the ratio $= r_j = \omega_j/\sigma_j$, we firstly ordered the K Gaussians as in [1]. 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:

$$B = argmin_b \left(\sum_{i=1}^{b} \omega_{i,t} > T \right) \tag{13}$$

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., $H(X_t) = \left| ln\left(\underline{h}(X_t)\right) - ln\left(\overline{h}(X_t)\right) \right|$. In Fig 1.a), the Gaussian primary MF with uncertain mean has:

$$H(X_t) = \begin{cases} \frac{2k_m|X_t-\mu|}{\sigma}, & \text{if } X_t \leq \mu - k_m\sigma \text{ or } X_t \geq \mu + k_m\sigma \\ \frac{|X_t-\mu|}{2\sigma^2} + \frac{k_m|X_t-\mu|}{\sigma} + \frac{k_m^2}{2}, & \text{if } \mu - k_m\sigma < X_t < \mu + k_m\sigma \end{cases} \tag{14}$$

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

$$H(X_t) = \left(\frac{1}{1/k_\nu^2 - k_\nu^2} \right) \frac{|X_t - \mu|^2}{2\sigma^2} \tag{15}$$

μ and σ are the mean and the std of the original certain T1 MF without uncertainty. Both (14) and (15) are increasing functions in terms of the deviation $|X_t - \mu|$. 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 relation-ship 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:

$$H(X_t) < k\sigma \tag{16}$$

where k is a constant threshold determined experimentally and 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.

3.3 Maintenance

The T2 FMGM Maintenance is made as in the original MGM [1] as follows:

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

$$\omega_{i,t+1} = (1-\alpha)\,\omega_{i,t} + \alpha \tag{17}$$

where α is a constant learning rate.

$$\mu_{i,t+1} = (1-\rho)\,\mu_{i,t} + \rho X_{t+1} \tag{18}$$

$$\sigma^2_{i,t+1} = (1-\rho)\,\sigma^2_{i,t} + \rho\,(X_{t+1}-\mu_{i,t+1})\,(X_{t+1}-\mu_{i,t+1})^T \tag{19}$$

where $\rho = \alpha\eta\,(X_{t+1}, \mu_i, \sum_i)$.
For the unmatched components, μ and σ are unchanged, only the weight is replaced by $\omega_{j,t+1} = (1-\alpha)\,\omega_{j,t}$.

– 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:

$$\omega_{k,t+1} = \text{Low Prior Weight} \tag{20}$$

$$\mu_{k,t+1} = X_{t+1} \tag{21}$$

$$\sigma^2_{k,t+1} = \text{Large Initial Variance} \tag{22}$$

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

4 Experimental Results

We have applied the T2 FMGM-UM and T2 FMGM-UV algorithms to indoor and outdoor videos where different critical situations occur like camera jitter, movement in the background, illuminations change and shadows. The experiments are conducted to compare the results of T2 FMGM with the crisp MGM [1]. Note that the best results were obtained with the values 2 and 0.9, respectively for the factors k_m and k_ν.

4.1 Indoor Scene Videos

PETS 2006 dataset [21] provides several video presenting indoor sequence in video surveillance context. In these video sequences, there are illumination changes and shadows. Fig.2 presents the results obtained by the MOG [1], the T2 FMGM-UM and the T2 FMGM-UV. It is noticed that the results obtained using the T2 FMGM-UM and the T2 FMGM-UV are better than using the crisp MOG. The silhouettes are well detected with the T2 FMGM-UM. T2 FMGM-UV is more sensitive because the variance is more unstable over time.

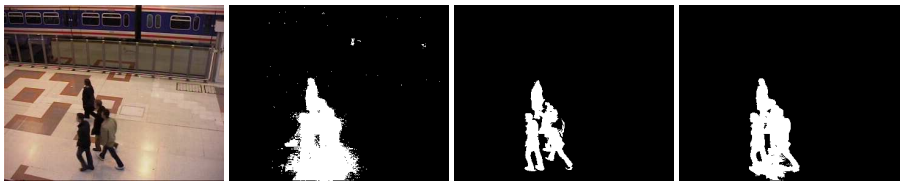


Fig. 2. Background subtraction with illumination changes (PETS 2006 dataset). From left to right respectively, the original image (Frame 165), the segmented image obtained by the MGM, the result obtained using the T2 FMGM-UM and the T2 FMGM-UV

4.2 Outdoor Scene videos

We have chosen three videos presenting different dynamic backgrounds: camera jitter, waving trees and water rippling. The first outdoor sequence that was tested involved a camera mounted on a tall tripod and comes from [22]. The wind caused the tripod to sway back and forth causing nominal motion in the scene. In Fig.3, the first row shows different current image. The second row and the third one show respectively the ground truth and the results obtained by the MGM proposed in [1]. It is evident that the motion causes substantial degradation in performance. The fourth and fith rows show respectively the results obtained by the T2 FMGM-UM and the T2 FMGM-UV. As for indoor scene, the T2 FMGM-UM and T2 FMGM-UV give better results than the crisp MOG. Numerical evaluation has been done in term of false positive and false negative. In Table.1 we can see that the T2 FMGM-UM and T2 FMGM-UV give less total error than the MGM. Furthermore, as shows in Fig.4, the T2 FMGM-UM generates the least false positive which is important in the context of target detection. We have also tested our method to the sequences Campus and Water Surface which come from [23]. Fig.5 shows the robustness of T2 FMGM-UM against waving trees and water rippling.

5 Conclusion

Type 2 Fuzzy Mixture of Gaussians Model (T2 FMGM) is an elegant technique to model the background and allows to handle critical situations like camera jitter, illumination changes, movement in the background and shadow. The T2 FMGM-UM is more

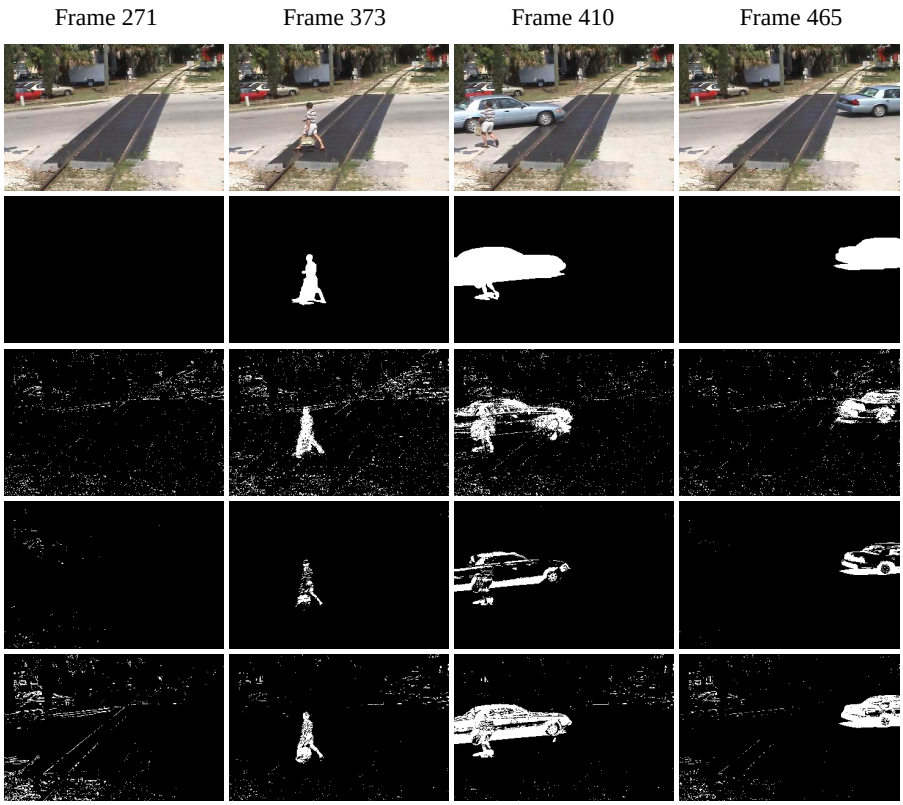


Fig. 3. The first row are the original images, the second row are the ground truth images, the third row are the results obtained by using the MGM, the fourth row are the result obtained using the T2 FMGM-UM and the firth row are the result obtained by using the T2 FMGM-UV

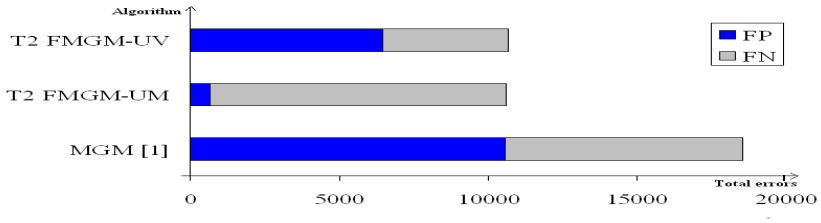


Fig. 4. Overall performance

robust than the T2 FMGM-UV due to a better estimation of the mean than the variance. One future direction of this work is an adaptive version of the proposed method which allows to determine dynamically the optimal number of Gaussians.

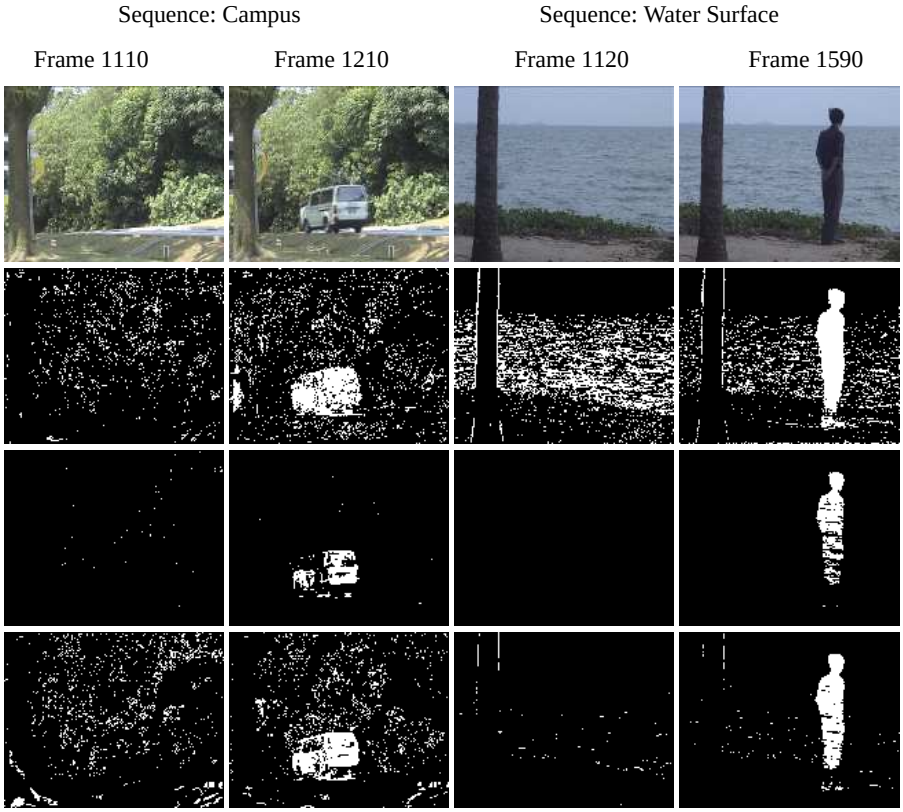


Fig. 5. Background subtraction with dynamic background. The first row shows the original frames for Campus and Water Surface sequences. The second row presents the segmented images obtained by the MGM. The Third and the fourth rows illustrate the result obtained using the T2 FMGM-UM and the T2 FMGM-UV respectively

References

1. Stauffer, C.: Adaptive background mixture models for real-time tracking. Computer Vision and Pattern Recognition **1** (1999) 246–252
2. Cheung, S., Kamath, C.: Robust background subtraction with foreground validation for urban traffic video. EURASIP 2005 (2005)
3. Carranza, J., Theobalt, C., Magnor, M., Seidel, H.: Free-viewpoint video of human actors. ACM Transactions on Graphics **22** (2003) 569–577

Table 1. Performance analysis

Method	Error Type	Frame 271	Frame 373	Frame 410	Frame 465	Total Error
MGM[1]	False neg.	0	1120	4818	2050	
	False pos.	2093	4124	2782	1589	18576
T2 FMGM-UM	False neg.	0	1414	6043	2520	
	False pos.	203	153	252	46	10631
T2 FMGM-UV	False neg.	0	957	2217	1069	
	False pos.	3069	1081	1119	1158	10670

4. Mikic, I., Trivedi, M., Hunter, E., Cosman, P.: Human body model acquisition and tracking using voxel data. International Journal of Computer Vision **1** (2003) 199–223
5. Horprasert, T., Haritaoglu, I., Wren, C., Harwood, D., Davis, L., Pentland, A.: Real-time 3d motion capture. Workshop on Perceptual User Interfaces (PUI 1998) **1** (1998) 87–90
6. El Baf, F., Bouwmans, T., Vachon, B.: Comparison of background subtraction methods for a multimedia learning space. SIGMAP 2007 (2007)
7. Piccardi, M.: Background subtraction techniques: a review. SMC 2004 **1** (2004) 3199–3104
8. Elhabian, S., El-Sayed, K., Ahmed, S.: Moving object detection in spatial domain using background removal techniques. Recent Patents on Computer Science **1** (2008) 32–54
9. Lee, B., Hedley, M.: Background estimation for video surveillance. IVCNZ 2002 **1** (2002) 315–320
10. McFarlane, N., Schofield, C.: Segmentation and tracking of piglets in images. BMVA 1995 **1** (1995) 187–193
11. Zheng, J., Wang, Y., Nihan, N., Hallenbeck, E.: Extracting roadway background image: A mode based approach. Journal of Transportation Research Report **1** (2006) 82–88
12. Wren, C., Azarbayejani, A., Darrell, T., Pentland, A.: Pfinder : Real-time tracking of the human body. PAMI 1997 **19** (1997) 780 –785
13. Elgammal, A., Harwood, D., Davis, L.: Non-parametric model for background subtraction. ECCV 2000 **1** (2000) 751–767
14. Sigari, M., Mozayani, N., Pourreza, H.: Fuzzy running average and fuzzy background subtraction: Concepts and application. International Journal of Computer Science and Network Security **8** (2008) 138–143
15. El Baf, F., Bouwmans, T., Vachon, B.: Fuzzy integral for moving object detection. IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2008) (2008)
16. Toyama, K., Krumm, J., Brumitt, B., Meyers, B.: Wallflower: Principles and practice of background maintenance. ICCV 1999 (1999)
17. Messelodi, S., Modena, C., Segata, N., Zanin, M.: A kalman filter based background updating algorithm robust to sharp illumination changes. ICIAP 2005 **3617** (2005) 163–170
18. Chang, R., Ghandi, T., Trivedi, M.: Vision modules for a multi sensory bridge monitoring approach. ITSC 2004 **1** (2004) 971–976
19. Zeng, J., Xie, L., Liu, Z.: Type-2 fuzzy gaussian mixture models. Pattern Recognition, available online 13 June 2008
20. Greiffenhagen, M., Ramesh, V., Niemann, H.: The systematic design and analysis cycle of a vision system: A case study in video surveillance. IEEE CVPR 2001 **2** (2001) 704
21. $http://www.cvg.rdg.ac.uk/PETS2006/data.html$
22. $http://www.cs.ucf.edu/\ yaser/backgroundsub.htm$
23. $http://perception.i2r.astar.edu.sg/bk\_model/bk\_index.html$