GMM with dynamic management of the number of gaussians based on AIRS

Nebili Wafa[1], Farou Brahim[2], and Seridi Hamid[3]

[1] Labstic Laboratory ,Computer Science department, 8 Mai 1945 Guelma University, BP 401 Guelma, Algeria
nebili.wafa@univ-guelma.dz
[2] Labstic Laboratory ,Computer Science department, 8 Mai 1945 Guelma University, BP 401 Guelma, Algeria
farou.brahim@univ-guelma.dz
[3] Labstic Laboratory ,Computer Science department, 8 Mai 1945 Guelma University, BP 401 Guelma, Algeria
seridi.hamid@univ-guelma.dz

Abstract. Background subtraction is an essential step in the process of monitoring videos. Several works have proposed models to differentiate the background pixels from the foreground pixels. Mixtures of Gaussian (GMM) are among the most popular models for a such problem. However, they suffer from certain inconveniences related to the light variations and complex scene due to the use of a fixed number of Gaussians. In this paper, we will propose an improvement of the GMM based on the use of the bio-inspired algorithm AIRS (Artificial Immune Recognition System) to generate and introduce new Gaussian instead of using a fixed number of Gaussians. Our approach is to exploit the robustness of the mutation function in the generation phase of the new ARBs to create new Gaussians. These Gaussians are then filtered into the resource competition phase in order to keep only ones that best represent the background. The system implemented and tested on the Wallflower database has proven its effectiveness against other state-of-art methods.

keywords: GMM · Video surveillance · AIRS · background subtraction.

1 Introduction

Various applications of video surveillance such as the detection and tracking of moving objects begin with background subtraction phase. Background subtraction is a binary classification operation that gives each pixel of a video sequence a label [8], for example: the pixels of the moving objects (foreground) take the value 1 and the pixels of the static objects are labeled by 0.In the real environment, the variations of pixels are very fast, which requires a robust and adaptable method to these variations. GMM is one of the most popular methods that has achieved considerable success in detecting changes in videos. However, this method has failed in problems related to: lighting changes and hidden areas. Several studies showed that the number of Gaussians in GMM influence on the results quality. The contribution of our work is to manage dynamically the number of Gaussians based on the AIRS algorithm instead of fixing them a priori by the user. The proposed system stars with a learning phase using the GMM algorithm and creates several background models for each pixel. These models are filtered according to resource competition and memory cell development process of the AIRS algorithm to select only the best models.

2 Related work

Subtracting the background from video sequences captured from fixed or non-fixed cameras remains a crucial problem due to the diversity of scenes that represent the background.

In recent years, various approaches, methods and systems have been proposed and developed to inspect dynamic regions and static regions. One of the most intuitive approaches is to compute the absolute difference (Δ_t) either between two successive frames [8], or between a reference image I_R, without any moving object, and the current image. In order to determine the objects in motion, a bit mask is applied according to a predefined threshold on the pixels of the resulting image.

Another way to subtract the background is to describe the history of the last n pixel values by a Gaussian probability distribution [39]. However, modeling by a single Gaussian is sensitive to fast pixel variations. Indeed, a single Gaussian can not memorize the old states of the pixel. This requires migration to more robust and multi-modal approach. The authors in [16] propose the first model which describes the variance of the recent values of each pixel by a mixture of the Gaussians. In this model, the Expectation Maximization (EM) algorithm is used to initialize and estimate the parameters of each Gaussian. In [13] authors estimate the probability density function of the recent N values of each pixel by a kernel estimator (KDE).

[17] Provides a nonparametric estimation of the background pattern. He uses the concept of a visual dictionary words to model the pixels of the background. Indeed, each pixel of the image is represented by a set of three values (visual word) which describes its current state. these values are initially estimated during the learning phase and are updated regularly over time to build a robust modeling.

Several works have taken spatial information into consideration. [25] Proposes a sub-spatial learning based on PCA (SL-PCA). The idea in [25] is to make a learning of the N background images by the PCA. Moving objects are identified according to the input image and the reconstructed image from its projection in the reduced dimension space. Authors in [34] provide a quick schema (SL-ICA) for background subtraction with Independent Component Analysis (ICA). Another work [5] presents a decomposition of video content by an incremental non-negative matrix factorization (NMF). Other methods [1] [20] [19] [22] [41] focalised on the selection and combination of good characteristics (the colors , texture, outlines) to improve the result quality.

Recently, some research works have introduced the fuzzy concept to develop more efficient and robust methods for modeling the background [10] [32] [4] [11] [12] [42].

Works done in [40] showed that the GMM offers a good compromise between quality and execution time compared to other methods. The first GMM model was proposed by [16], however Stauffer and Grimson [33] offers a GMM standard with efficient update equations. Several works and contributions have been proposed to improve the quality of GMM. Among these methods are those that are focused on improving the model adaptation speed [27] [18] [21]. Others are interested in hybrid models such as GMM and K-means [7], GMM and fuzzy logic [12], GMM and adaptive background [9],GMM and Block matching [15], boosted Gaussian Mixture Model [24], Markov Random Fields [28], GMM with PSO [38] to overcome GMM problems. There are also several works that are invested in the type of characteristics [6] [30] or in the acquisition material [29]. In addition to spatio-temporal methods [31], some researchers have used local contextual information around a pixel, such as the region [14] [26], the block [36] and the cluster [3] [35].

There are also many methods that used deep learning for subtracting the background FgSegNet_S (FPM) [23], Cascade CNN [37], DeepBS [2]However, deep leaning methods require a large number of simples and needs more time for training.

3 Proposition

The in-depth study made on Gaussian mixtures shows the important role of the number of Gaussians in describing the pixel variations. Following this principle, we propose a novel mechanism to produce new Gaussians based on the AIRS algorithm in order to be as faithful as possible to the background model. Indeed, the idea is to pass from a static model where the number of Gaussians is fixed empirically for all pixels towards a model dynamic and adaptive according to the environment and the background complexity.

First, we create a set of Gaussian (g_i) representing the background for the pixel P_t (at time t) that vitrifies:

$$Set_{background} = \{g_i, \frac{P_t - u_i}{\sigma_i} < 2.5\} \tag{1}$$

Each gaussian g_i is represented by: the pixel value P_i, the average u_i, the variance σ_i and the weight w_i .

After creating the background model, we choose the $Gaussian(mc_{match})$ that has the closest distance to the value of the current pixel.

$$mc_{match} = min\_Gaussian(Set_{background}) \tag{2}$$

mc_{match} is mutated in the ARBs generation phase. At the end of this phase, new Gaussian (clones) is created. the number of clones is calculated by the following equation:

$$Num\_Clones = clonalRate \times hyperClonalRate \times distance(P_t, mc_{match}) \quad (3)$$

$$Set\_Clones = \{g_{clone_1}, g_{clone_2}, .., g_{clone_{Num\_Clones}}\} \quad (4)$$

Such that :

$$g_{clone_i} = Mutation(mc_{match}) \quad (5)$$

The *clonalRate* and the *hyperClonalRate* are two integer values chosen by the user.

New clones must be filtered through the Resource Competition (AIRS) process, keeping only the best and the correct Gaussians. The filtering operation uses the condition (6) to remove the least representative Gaussians.

$$\frac{P_i - u_i}{\sigma_i} < 2.5 \quad (6)$$

The last step of the AIRS algorithm is to introduce the memory cells mc from the previous set (Set_Clones). This operation consists of choosing the most representative Gaussians among the new Gaussians and adding them to the background model according to the following equation:

$$distance(P_t, g_{clone_i}) < distance(P_t, mc_{match}) \quad (7)$$

If the previous condition is verified, we compare the average distance of mc_{match} and g_{clone_i} with the affinity threshold AT multiplied by the scalar affinity threshold ATS:

$$mean_{distance}(mc_{match}, g_{clone_i}) < AT \times ATS \quad (8)$$

With : AT : the average distance of all background models generated in the learning phase.
If the equation (8) is satisfied the mc_{match} will be deleted from the set of MC. After these steps and to determine whether the pixel belongs to the background or foreground, the Gaussians are ordered according to the value of $w_{k,t}/\sigma_{k,t}$. The Gaussians that represent the state of P_t is the first distribution that satisfies the following equation:

$$\beta = argmin(\sum_{k=1}^{b} w_{k,t} > B) \quad (9)$$

Where B determines the minimum part of the data corresponding to the background. $w_{k,t}$ is the weight of the K distribution. Regarding the learning phase, we applied the same principle of classical GMM.

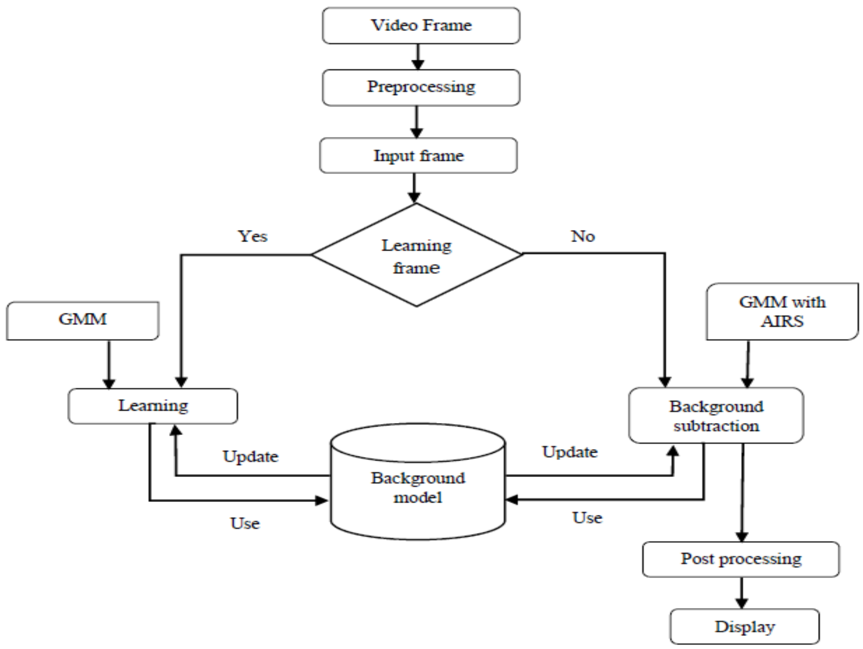

Fig. 1. The global architecture of the proposed system.

4 Tests and results

Our approach is implemented and tested on some videos from the Wallflower database. After several empirical tests, the learning rate α, the minimum part of the data corresponding to the background β, clonalRate, hyperClonalRate, mutationRate, ATS) are respectively fixed to 0.01, 0.3, 10, 2, 0.1, 0.2.
The results obtained are compared with the most referenced state of art methods in the modeling of the background (see Figure 2).

Our system achieved good results in Foreground Aperture, Camoflage, Bootstarp, Waving trees videos, it ranks in the 1st position compared to the other state of the art methods, but they have some false detection. However, our system failed to detect objetcs in scenes that have a large change in illumination.

The obtained results clearly show that our system exceeds other sate of the art methods in videos with small variations in the background. However, our system is sensitive when the scene contains high illumination. This is due to the nature of the method that uses a pixel-based approach to detect moving objects.

5 Conclusion

In this paper, we have proposed a new approach which allows to reduce the inconvenience of GMM for background subtraction. The idea is to introduce new Gaussians using Artificial Immune Recognition System. This allows to move from a static to dynamic approach that can easily adapt the model to nature of the environment. Results obtained on several videos from a public benchmark showed the effectiveness of this new process with small variations in the background. However, our system is sensitive when the scene contains high illumination. This is due to the nature of the method that uses a pixel-based approach to detect moving objects.

Fig. 2. Results obtained on Wallflower dataset.

References

1. Azab, M.M., Shedeed, H.A., Hussein, A.S.: A new technique for background modeling and subtraction for motion detection in real-time videos. In: Image Processing (ICIP), 2010 17th IEEE International Conference on. pp. 3453–3456. IEEE (2010)

2. Babaee, M., Dinh, D.T., Rigoll, G.: A deep convolutional neural network for video sequence background subtraction. Pattern Recognition **76**, 635–649 (2018)
3. Bhaskar, H., Mihaylova, L., Maskell, S.: Automatic target detection based on background modeling using adaptive cluster density estimation (2007)
4. Bouwmans, T., El Baf, F.: Modeling of dynamic backgrounds by type-2 fuzzy gaussians mixture models. MASAUM Journal of of Basic and Applied Sciences **1**(2), 265–276 (2010)
5. Bucak, S.S., Gunsel, B.: Video content representation by incremental non-negative matrix factorization. In: Image Processing, 2007. ICIP 2007. IEEE International Conference on. vol. 2, pp. II–113. IEEE (2007)
6. Caseiro, R., Henriques, J.F., Batista, J.: Foreground segmentation via background modeling on riemannian manifolds. In: Pattern Recognition (ICPR), 2010 20th International Conference on. pp. 3570–3574. IEEE (2010)
7. Charoenpong, T., Supasuteekul, A., Nuthong, C.: Adaptive background modeling from an image sequence by using k-means clustering. In: Electrical Engineering/Electronics Computer Telecommunications and Information Technology (ECTI-CON), 2010 International Conference on. pp. 880–883. IEEE (2010)
8. Collins, R.T., Lipton, A.J., Kanade, T., Fujiyoshi, H., Duggins, D., Tsin, Y., Tolliver, D., Enomoto, N., Hasegawa, O., Burt, P., et al.: A system for video surveillance and monitoring. VSAM final report pp. 1–68 (2000)
9. Doulamis, A., Kalisperakis, I., Stentoumis, C., Matsatsinis, N.: Self adaptive background modeling for identifying persons' falls. In: Semantic Media Adaptation and Personalization (SMAP), 2010 5th International Workshop on. pp. 57–63. IEEE (2010)
10. El Baf, F., Bouwmans, T., Vachon, B.: Fuzzy integral for moving object detection. In: Fuzzy Systems, 2008. FUZZ-IEEE 2008.(IEEE World Congress on Computational Intelligence). IEEE International Conference on. pp. 1729–1736. IEEE (2008)
11. El Baf, F., Bouwmans, T., Vachon, B.: Type-2 fuzzy mixture of gaussians model: application to background modeling. In: International Symposium on Visual Computing. pp. 772–781. Springer (2008)
12. El Baf, F., Bouwmans, T., Vachon, B.: Fuzzy statistical modeling of dynamic backgrounds for moving object detection in infrared videos. In: Computer Vision and Pattern Recognition Workshops, 2009. CVPR Workshops 2009. IEEE Computer Society Conference on. pp. 60–65. IEEE (2009)
13. Elgammal, A., Harwood, D., Davis, L.: Non-parametric model for background subtraction. In: European conference on computer vision. pp. 751–767. Springer (2000)
14. Fang, X., Xiong, W., Hu, B., Wang, L.: A moving object detection algorithm based on color information. In: Journal of Physics: Conference Series. vol. 48, p. 384. IOP Publishing (2006)
15. Farou, B., Kouahla, M.N., Seridi, H., Akdag, H.: Efficient local monitoring approach for the task of background subtraction. Engineering Applications of Artificial Intelligence **64**, 1–12 (2017)
16. Friedman, N., Russell, S.: Image segmentation in video sequences: A probabilistic approach. In: Proceedings of the Thirteenth conference on Uncertainty in artificial intelligence. pp. 175–181. Morgan Kaufmann Publishers Inc. (1997)
17. Haritaoglu, I., Harwood, D., Davis, L.S.: W4: Real-time surveillance of people and their activities. IEEE Transactions on Pattern Analysis & Machine Intelligence (8), 809–830 (2000)

18. Hayman, E., Eklundh, J.O.: Statistical background subtraction for a mobile observer. In: null. p. 67. IEEE (2003)
19. Jain, V., Kimia, B.B., Mundy, J.L.: Background modeling based on subpixel edges. In: Image Processing, 2007. ICIP 2007. IEEE International Conference on. vol. 6, pp. VI–321. IEEE (2007)
20. Jian, X., Xiao-qing, D., Sheng-jin, W., You-shou, W.: Background subtraction based on a combination of texture, color and intensity. In: Signal Processing, 2008. ICSP 2008. 9th International Conference on. pp. 1400–1405. IEEE (2008)
21. KaewTraKulPong, P., Bowden, R.: An improved adaptive background mixture model for real-time tracking with shadow detection. In: Video-based surveillance systems, pp. 135–144. Springer (2002)
22. Kristensen, F., Nilsson, P., Öwall, V.: Background segmentation beyond rgb. In: Asian COnference on Computer Vision. pp. 602–612. Springer (2006)
23. Lim, L.A., Keles, H.Y.: Foreground segmentation using convolutional neural networks for multiscale feature encoding. Pattern Recognition Letters **112**, 256–262 (2018)
24. Martins, I., Carvalho, P., Corte-Real, L., Alba-Castro, J.L.: Bmog: boosted gaussian mixture model with controlled complexity for background subtraction. Pattern Analysis and Applications **21**(3), 641–654 (2018)
25. Oliver, N.M., Rosario, B., Pentland, A.P.: A bayesian computer vision system for modeling human interactions. IEEE transactions on pattern analysis and machine intelligence **22**(8), 831–843 (2000)
26. Pokrajac, D., Latecki, L.J.: Spatiotemporal blocks-based moving objects identification and tracking. IEEE Visual Surveillance and Performance Evaluation of Tracking and Surveillance (VS-PETS) pp. 70–77 (2003)
27. Power, P.W., Schoonees, J.A.: Understanding background mixture models for foreground segmentation. In: Proceedings image and vision computing New Zealand. vol. 2002 (2002)
28. Schindler, K., Wang, H.: Smooth foreground-background segmentation for video processing. In: Asian Conference on Computer Vision. pp. 581–590. Springer (2006)
29. Seki, M., Okuda, H., Hashimoto, M., Hirata, N.: Object modeling using gaussian mixture model for infrared image and its application to vehicle detection. Journal of Robotics and Mechatronics **18**(6), 738 (2006)
30. Setiawan, N.A., Seok-Ju, H., Jang-Woon, K., Chil-Woo, L.: Gaussian mixture model in improved hls color space for human silhouette extraction. In: Advances in Artificial Reality and Tele-Existence, pp. 732–741. Springer (2006)
31. Shimada, A., Nonaka, Y., Nagahara, H., Taniguchi, R.i.: Case-based background modeling: associative background database towards low-cost and high-performance change detection. Machine vision and applications **25**(5), 1121–1131 (2014)
32. Sigari, M.H., Mozayani, N., Pourreza, H.: Fuzzy running average and fuzzy background subtraction: concepts and application. International Journal of Computer Science and Network Security **8**(2), 138–143 (2008)
33. Stauffer, C., Grimson, W.E.L.: Adaptive background mixture models for real-time tracking. In: cvpr. p. 2246. IEEE (1999)
34. Tsai, D.M., Lai, S.C.: Independent component analysis-based background subtraction for indoor surveillance. IEEE Transactions on image processing **18**(1), 158–167 (2009)
35. Valentine, B., Apewokin, S., Wills, L., Wills, S.: An efficient, chromatic clustering-based background model for embedded vision platforms. Computer Vision and Image Understanding **114**(11), 1152–1163 (2010)

36. Varadarajan, S., Miller, P., Zhou, H.: Spatial mixture of gaussians for dynamic background modelling. In: Advanced Video and Signal Based Surveillance (AVSS), 2013 10th IEEE International Conference on. pp. 63–68. IEEE (2013)
37. Wang, Y., Luo, Z., Jodoin, P.M.: Interactive deep learning method for segmenting moving objects. Pattern Recognition Letters **96**, 66–75 (2017)
38. White, B., Shah, M.: Automatically tuning background subtraction parameters using particle swarm optimization. In: Multimedia and Expo, 2007 IEEE International Conference on. pp. 1826–1829. IEEE (2007)
39. Wren, C.R., Azarbayejani, A., Darrell, T., Pentland, A.P.: Pfinder: Real-time tracking of the human body. IEEE Transactions on pattern analysis and machine intelligence **19**(7), 780–785 (1997)
40. Yu, J., Zhou, X., Qian, F.: Object kinematic model: A novel approach of adaptive background mixture models for video segmentation. In: Intelligent Control and Automation (WCICA), 2010 8th World Congress on. pp. 6225–6228. IEEE (2010)
41. Zhang, H., Xu, D.: Fusing color and texture features for background model. In: Fuzzy Systems and Knowledge Discovery: Third International Conference, FSKD 2006, Xi'an, China, September 24-28, 2006. Proceedings 3. pp. 887–893. Springer (2006)
42. Zhao, Z., Bouwmans, T., Zhang, X., Fang, Y.: A fuzzy background modeling approach for motion detection in dynamic backgrounds. In: Multimedia and signal processing, pp. 177–185. Springer (2012)