Improved Parameters Updating Algorithm for the Detection of Moving Objects

Brahim Farou[1,2], Hamid Seridi[2], and Herman Akdag[3]

[1] Computer Science Department, Badji Mokhtar-Annaba University, P.O.B 12, 23000 Annaba, Algeria
farou@ymail.com
[2] LabSTIC, Guelma University, POB 401, 24000 Guelma, Algeria
seridihamid@yahoo.fr
[3] LIASD, Paris 8 University, 93526 Saint-Denis, France
Herman.akdag@ai.univ-paris8.fr

Abstract. The presence of dynamic scene is a challenging problem in video surveillance systems tasks. Mixture of Gaussian (MOG) is the most appropriate method to model dynamic background. However, local variations and the instant variations in the brightness decrease the performance of the later. We present in this paper a novel and efficient method that will significantly reduce MOG drawbacks by an improved parameters updating algorithm. Starting from a normalization step, we divide each extracted frame into several blocks. Then, we apply an improved updating algorithm for each block to control local variation. When a significant environment changes are detected in one or more blocs, the parameters of MOG assigned to these blocks are updated and the parameters of the rest remain the same. Experimental results demonstrate that the proposed approach is effective and efficient compared with state-of-the-art background subtraction methods.

Keywords: Background subtraction, motion detection, MOG, Machine Vision, Videosurveillance

1 Introduction

The detection of moving object is the key step in many computer vision applications such as video surveillance, control applications, human machine interaction, and motion analysis. The challenge in such systems is to achieve high sensitivity in the detection of moving objects while maintaining a good discrimination rates and low processing time. The intrinsic nature of environment with illumination changes, shadows, waving flags, dust, bootstrapping and ghosts make tasks even more difficult. Recently, important efforts in this field have been focused on developing theories, methods and systems to deal with this problems and the most widely adopted techniques for handling these issues are optical flow, frame differencing and background subtraction. Background subtraction process is usually used with the assumption that the im-ages extracted form

video without any additional objects follow a fixed behavior and can be well described by a statistical model. In this case, the appearance of a new object in background will make this part inappropriate with the building model. The main idea in such approach is to model each pixel separately by a probability density function. Works done in [31] showed that GMM provides a good compromise between quality and execution time compared to other methods. The first use of GMM for modeling the background was proposed by Friedman and Russell [11].However Stauffer and Grimson [26] proposed the standard algorithm with an efficient update equations. Some extensions are given by [20, 12, 14] to improve the model adaptation speed. Other GMM algorithms were also proposed [27, 34] to remove GMM drawback. Unfortunately, local variations and instant changes in brightness remains the major problem of GMM [33, 13]. In the last decade, several studies have attempted to improve the performance of GMM in environments with multiple dimming and high condensation background. Initial ideas focused on substitution of using color characteristics [2] Setiawan et al. [24]or infrared camera [23]. Hybrid models such as GMM and K-means [3], GMM and fuzzy logic [1], Markov Random Fields [22],GMM and adaptive background [9, 25], have been proposed to overcome GMM drawbacks. Other works have focused on improving the learning speed [15, 28] through an adaptive learning rate [29], Better settings White and Shah [32] and the execution time [17] by using real parallel operations on multi-processor machines. Other systems use two backgrounds [4] to solve the problem of change in brightness between day and night or use Multi-level approaches [5–7]. Despite many algorithms have been pro-posed in the literature, the detection of moving objects in complex and dynamic environments is still far from being completely solved. In this paper, we will focus on the detection of moving objects in video surveillance through a fixed camera. To overcome the problems mentioned before, we propose a new and efficient background subtraction method based on GMM and local background monitoring. To cover all sections, the rest of the paper is organized as follows. The Preprocessing task is presented in Section 2. Section 3 is devoted to the similarity measurement. The background subtraction method and the proposed algorithm is presented in section 4 and section 6. We present in section 5 a local monitoring method used to update the MOG parameters. Results and discussion are presented in Section 6. Section 7 concludes the paper.

2   Preprocessing

The objective of the preprocessing task is to make the images more appropriate to apply algorithms in any system component allowing improvement in the success rates. In this phase, we start by transforming the captured video into a set of images. Then, we apply median filter to remove noise from the image. The extracted images from the video is done in the RGB color space, but this representation is not adequate because of the influence of light on the description of objects [28]. For this reason, we made a transfer to HSL model recognized to be one of the closest model of human perception and it provides a direct control of

chromaticity. The following are supplementary preprocessing techniques applied in our system.

2.1   Histogram Equalization

Histogram Equalization is an illumination normalization technique that uses the distribution of the original image to generate an image with uniform histogram. The objective of histogram equalization is to minimize the contrast in areas that are too light or too dark for an image.

2.2   Contour Detection

Contour correspond to the local variations in the intensity of the image pixel values. It is applied to preserve local features despite the influence of brightness. There are numerous contour detection techniques, but the context of real time processing lead to use a fast contour detection algorithm with inherent smoothing properties that can be adapted to different conditions of noise and artifacts. We used Soblel filter reported to be the best filter under real time consideration.

2.3   Splitting

This operation is only used in the initialization step. We divide the first frame into N equal size blocks to minimize local variations and to simplify the monitoring task. We noticed that the number of areas greatly influences on system quality. A large number of areas lead us to the starting point (pixel-based approach). In case where the number of areas is small (the size of the area is large), local variations accumulated in the same area force the system to consider the latter as an intense variation. In this way, all pixels belonging to the area will be updated. However, the number of blocks may change in processing time to improve system performance.

3   Similarity Measurement

The similarity between two sequences of measurement is a measure that quantifies the dependency between them. The use of similarity measure requires solving three major problems. The first one is to find the saved image that best matches the observed image. The second problem involves locating an object of interest in an observed image. The last one is the presence of rotational and scaling differences between the stored and observed image. In our case, the two first problems are similar and resolved by using contour detection algorithm. Indeed, the original image is divided into a set of blocks and the similarity is applied, not to detect any type of object, but to measure the blocks dependence at the same position between the reference image and the observed image. The use of a binary image containing only contours, reduces the brightness change effect since the contours are invariant to the latter. The third problem is not probable since

the camera is static and it has no zoom effect. Various similarity measures have been proposed in the literature. However, each measure has its own strengths and weaknesses and a measure that performs well on one type of images may not work on another types. In this paper we use Pearsons correlation coefficient which is reported in the literature as the best similarity measure on various image types.

4   Mixture of Gaussians

MoG is a statistical model that assumes the data where originates from a weighted sum of several Gaussian distributions. Stauffer and Grimson [26] presented an adaptive GMM method to model a dynamic background in image sequences. If K Gaussian distributions are used to describe the history of a pixel, the observation of the given pixel will be in one of the K states at one time [3]. K determines the multimodality of the background and the selection of K is generally based on the available memory and computing power. Stauffer and Grimson [26] proposed to set K from 3 to 5. First, each pixel is characterized by its intensity in the HSL color space. Then, the probability of observing the current pixel value is given by the following equation in the multidimensional case:

$$P\left(P_t\right) = \sum_{i=1}^{k} w_{i,t} \cdot \eta\left(P_t, \mu_{i,t}, \Sigma_{i,t}\right) \tag{1}$$

Where: k is the number of associated Gaussians to each pixel, $w_{i,t}$ is the calculated weight, $\mu_{i,t}$ is the mean and $\Sigma_{i,t}$ is the covariance matrix that are respectively evaluated for the ith Gaussian at time t. η is a Gaussian probability density function:

$$\eta\left(P_t, \mu, \Sigma\right) = \frac{1}{2\pi^{\frac{n}{2}} \left|\Sigma\right|^{\frac{1}{2}}} \exp^{\frac{1}{2}\left(P_t-\mu\right)\Sigma^{-1}\left(P_t-\mu\right)} \tag{2}$$

For real time consideration, the update of the model is carried out by using an online K-Means approximation algorithm [3], [8]. After the parameters initialization, a first foreground detection can be made and the parameters are updated. When the new frame incomes, each pixel value is checked through the existing k Gaussian distributions, until a match is found. A pixel matched a Gaussian distribution if the pixel value is within 2.5 standard deviations of distribution according to Eq. 3.

$$\frac{\left|P_t - \mu_i\right|}{\sigma_i} < 2.5 \tag{3}$$

When a match is found with one of the k Gaussian, we look for the Gaussian distribution classification. If the Gaussian distribution is identified as a background, the pixel is classified as background. Otherwise, the pixel is classified as

foreground. The prior weights of the K distributions are updated according to Eq. 4:

$$W_{k,t} = (1 - \alpha) \cdot W_{k,t-1} + \alpha M_{k,t} \tag{4}$$

Where: α is the learning coefficient which determines the model adaptation speed and $M_{k,t}$ is equal to 1 for the distribution which satisfy 3and 0 for others. After updating weights, a normalization step is carried out to ensure that the sum of the weights is always equal to 1. For the unmatched components, μ and σ parameters remain unchanged. The parameters of the distribution which matches the new observation are updated using the following equations:

$$\mu_{k,t} = (1 - \varphi_k) \cdot \mu_{k,t-1} + \varphi_k \cdot P_t \tag{5}$$

$$\sigma_{k,t}^2 = (1 - \varphi_k) \cdot \varphi_{k,t-1}^2 + \varphi_k \left(P_t - \mu_{k,t}\right)^T \left(P_t - \mu_{k,t}\right) \tag{6}$$

With

$$\varphi_t = \alpha\eta\left(P_t/\mu_k\sigma_k\right) \tag{7}$$

If none of the distributions satisfy the Eq. 3, then the pixel is associated with first plan and the parameters of the least probable distribution is replaced by a new Gaussian with the current value as its mean value, an initially high variance, and a low prior weight parameter according to Eq. 8,Eq. 9 and Eq. 10 described below:

$$\sigma_{k,t}^2 = Large \quad Initial \quad Variance \tag{8}$$

$$W_{k,t} = Low \quad Prior \quad Weight \tag{9}$$

$$\mu_{k,t} = P_t \tag{10}$$

W is the initial weight value for the new Gaussian. If w is higher, the distribution is chosen as the background model for a long time. To decide if P_t is included in the background distributions, the distributions are ordered according to the value of $W_{k,t}/\sigma_{k,t}$. This ordering use the assumption that a background pixel corresponds to a high weight with a weak variance for the reason that the background is practically constant and it is more present than moving objects. The first β distributions that verify the Eq. 11 are selected to represent the background.

$$\beta = \arg\min\left(\Sigma_{k=1}^b W_{k,t} > B\right) \tag{11}$$

The threshold B represents the minimum portion of the total weight given to background model. If a small value for B is chosen, then the background becomes unimodal. If B is higher, a multi-modal distribution caused by a repetitive background motion could result from a variety of background component that allows the background to accept more than one Gaussian distribution. The use of unique threshold B for GMM implies a miss classification especially when scene contains both dynamic and static area. A higher threshold can achieve correct classifications in a dynamic background but makes incorrect detection of moving objects in stationary background.

5   Adaptive Local Monitoring

Methods based on MOG use the pixel value for detecting a probable change in the background based on the assumption that a moving object is a set of pixels in movement. This vision is very useful because it requires no a priori knowledge of objects and their trajectories. However, the natural environment is far from perfect. The presence of dust, the change in brightness, rain, wind,etc. influence on pixel value making unwanted local variation and leading to a misdetection of motion. The false pixel detection induces the system to make errors in the following steps, either by the deformation of the moving objects or by signaling a false movement. To overcome these problems, we proposed an adaptive local monitoring algorithm for each block to control local variation. From the start of the process of detecting moving objects, the monitoring task is enabled by assigning an observer to each block. The role of the latter is to monitor and report the presence of any activity that may be a movement. The decision of the presence or absence of movement is ensured by calculating the similarity between two states of the same block. Indeed, after assigning a block for each monitor, it stores the initial state which contains only contour. The first state is taken without the presence of moving objects. The second state represents the image in process. For the convenience of the update algorithm, each pixel has been labeled with the block number to which it belongs. It is used to provide updates to the concerned pixels only. So if a significant activity in a block of the image is detected, the parameters of the Gaussian assigned to all pixels of this block decide whether there has been any motion, and will be updated according to the proposed model. The parameters of the Gaussian assigned to other blocks will not undergo any change. This process will eliminate local variation, because only blocks with significant change will be considered by the system.

6   Results and interpretation

The system presented in this paper is implemented in Java on a computer with an Intel Core i5 2.67 GHz and a 4GB memory capacity. In this section, we shall present results of our method while challenging real-world situations. We take in addition to our database, three publicly available Benchmark Dataset Collection. The first one (BDC1) has six sequences in the Dataset (campus, highway I, highway I2, highway II intelligent room and laboratory) [21]. The second (BDC2) has nine sequences (bootstrap, a campus, a curtain, an escalator, a fountain, a hall, a lobby, a shopping mall and a water surface) [18].The last one (BDC3) has two sequences (highway and hallway) [16]. Our database (BDC4) has four sequences (campus, a hallway, a highway and a public park). In BDC4, The outdoor videos are recorded in a random situation and without any assumption on the observed scene where a group of clouds is passing in the sky, causing sudden illumination changes. For measuring accuracy we used different metrics, namely Precision and Recall.
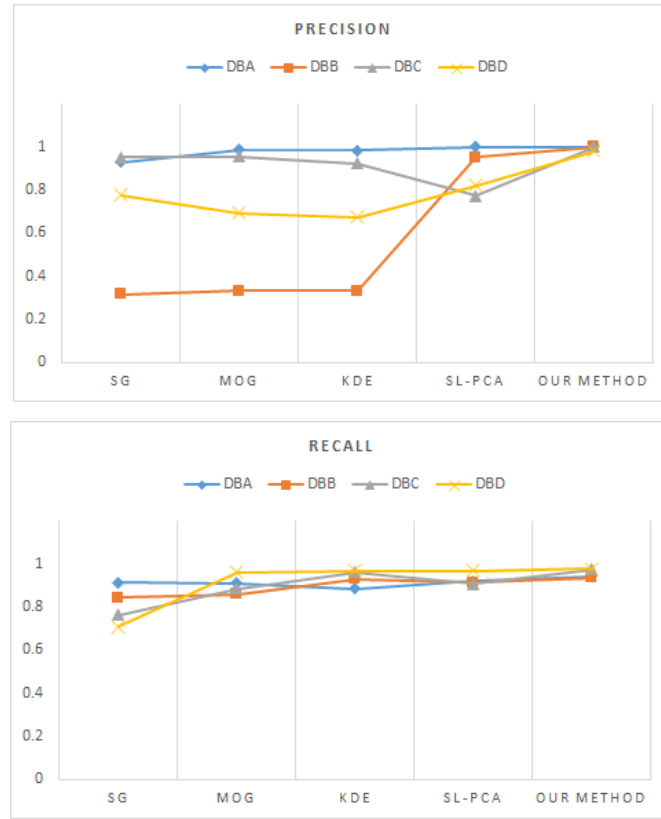
Fig. 1. Precision and recall results for MOG [26], SL-PCA [19], SG [30], KDE [10], and our method

Fig. 2. Background subtraction results in personnel video database in both indoor and outdoor environments

Frame 407 - Hallway

Frame 966 - Hallway

Frame 114 - Campus

Frame 225 – Intelligent Room

Frame 337 - Laboratory

Frame 26 – Highway II

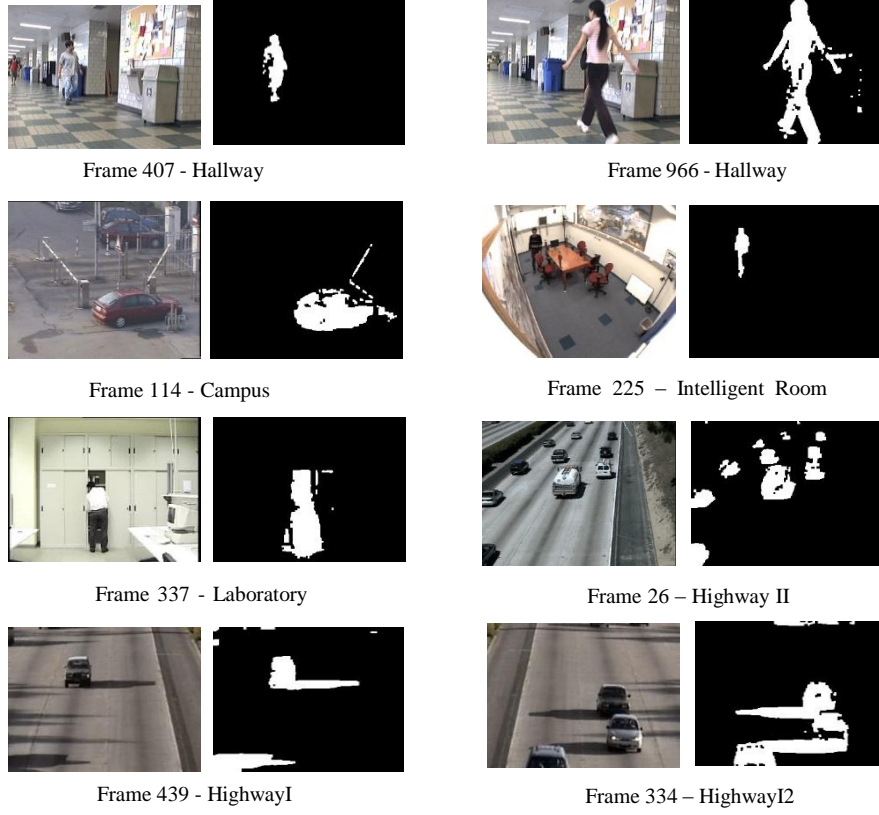Frame 439 - HighwayI

Frame 334 – HighwayI2

Fig. 3. Background subtraction results in public video database

Recall gives the percentage of corrected pixels classified as background when compared with the total number of background pixels in the ground truth. Precision gives the percentage of corrected pixels classified as background as compared at the total pixels classified as background by the method.

$$Recall = \frac{TP}{TP + FN} \tag{12}$$

$$Precision = \frac{TP}{TP + FP} \tag{13}$$

FP and FN refer to pixels misclassified as foreground (FP) or background (FN) while TP account for accurately classified pixels respectively as foreground. A good performance is obtained when the detection rate is high without altering the precision.

Figure.1 clearly shows that our method outperform the well-known background subtraction methods in term of precision and recall. Figure.2 and Figure.3

show some qualitative results on both public and personal databases. They show that our system is able to give a very good subtraction in environment without any assumption on lighting condition. They also show the effectiveness of the proposed system in outdoor and indoor environment.

7   Conclusion

In this paper, we proposed a background subtraction system for image sequences extracted from fixed camera using an improved parameters updating algorithm for mixture of Gaussians. To overcome the brightness and local variation we first made a transition from RGB to HSL color space. Then we divided the image into N areas and assigned to each one a local monitoring algorithm that allows selecting regions with a very large change using Pearsons Correlation Coefficient. Transfer to HSL color space has significantly decreased light effect on the system behavior through accumulating all brightness variations in a single component (L). While segmenting the image into regions have eliminated local variations caused mainly by the presence of dust. Tests conducted on databases show that our system has a good sensitivity, more accuracy compared with well-known methods. In future work, our algorithm will be adjusted by dividing the image into homogenous regions and solving the problem of shadow and color similarity between moving objects and background.

References

1. BAF, F.E., BOUWMANS, T., VACHON, B.: Fuzzy statistical modeling of dynamic backgrounds for moving object detection in infrared videos. In: Computer Vision and Pattern Recognition (2009)
2. Caseiro, R., Henriques, J.F., Batista, J.: Foreground Segmentation via Background Modeling on Riemannian Manifolds. In: International Conference on Pattern Recognition. pp. 3570–3574 (2010)
3. Charoenpong, T., Supasuteekul, A., Nuthong, C.: Adaptive background modeling from an image sequence by using K-Means clustering (2010)
4. Cheng, F.C., Huang, S.C., Ruan, S.J.: Illumination-Sensitive Background Modeling Approach for Accurate Moving Object Detection. IEEE Transactions on Broadcasting 57, 794–801 (2011)
5. Cristani, M., Bicegi, M., Murino, V.: Integrated Region-and Pixel-based Approach to Background Modeling (2002)
6. Cristani, M., Murino, V.: A spatial sampling mechanism for effective background subtraction. In: Computer Vision Theory and Applications. pp. 403–412 (2007)
7. Cristani, M., Murino, V.: Background Subtraction with Adaptive Spatio-Temporal Neighborhood Analysis. In: Computer Vision Theory and Applications. pp. 484–489 (2008)
8. Djouadi, A., Snorrason, ., Garber, F.D.: The Quality of Training Sample Estimates of the Bhattacharyya Coefficient. IEEE Transactions on Pattern Analysis and Machine Intelligence 12, 92–97 (1990)

9. Doulamis, A., Kalisperakis, I., Stentoumis, C., Matsatsinis, N.: Self Adaptive background modeling for identifying persons' falls. In: International Workshop on Semantic Media Adaptation and Personalization (2010)

10. Elgammal, A.M., Harwood, D., Davis, L.S.: Non-parametric Model for Background Subtraction (2000)

11. Friedman, N., Russell, S.J.: Image Segmentation in Video Sequences: A Probabilistic Approach. In: Uncertainty in Artificial Intelligence. pp. 175–181 (1997)

12. Hayman, E., olof Eklundh, J.: Statistical Background Subtraction for a Mobile Observer. In: International Conference on Computer Vision. pp. 67–74 (2003)

13. Hedayati, M., Zaki, W.M.D.W., Hussain, A.: Real-time background subtraction for video surveillance: From research to reality. In: International Colloquium on Signal Processing & Its Applications (2010)

14. Kaewtrakulpong, P., Bowden, R.: An improved adaptive background mixture model for real time tracking with shadow dectection (2001)

15. Kan, J., Li, K., Tang, J., Du, X.: Background modeling method based on improved multi-Gaussian distribution. In: International Conference on Computer Application and System Modeling (2010)

16. Li, L., Huang, W., Gu, I.Y.H., Tian, Q.: Statistical modeling of complex backgrounds for foreground object detection. IEEE Transactions on Image Processing 13, 1459–1472 (2004)

17. Li, X., Jing, X.: FPGA based mixture Gaussian background modeling and motion detection 4, 2078–2081 (2011)

18. Martel-brisson, N., Zaccarin, A.: Learning and Removing Cast Shadows through a Multidistribution Approach. IEEE Transactions on Pattern Analysis and Machine Intelligence 29, 1133–1146 (2007)

19. 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, 831–843 (2000)

20. Power, P.W., Schoonees, J.A.: Understanding Background Mixture Models for Foreground Segmentation (2002)

21. Prati, A., c, I.M.., Trivedi, M.M., Cucchiara, R.: Detecting Moving Shadows: Formulation, Algorithms and Evaluation

22. Schindler, K., Wang, H.: Smooth Foreground-Background Segmentation for Video Processing (2006)

23. 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)

24. Setiawan, N.A., ju Hong, S., woon Kim, J., woo Lee, C.: Gaussian Mixture Model in Improved HLS Color Space for Human Silhouette Extraction (2006)

25. Sheng, Z.B., Cui, X.Y.: An adaptive learning rate GMM for background extraction. Optoelectronics Letters 4, 460–463 (2008)

26. Stauffer, C., Grimson, W.E.L.: Adaptive Background Mixture Models for Real-Time Tracking. In: Computer Vision and Pattern Recognition. vol. 2, pp. 2246–2252 (1999)

27. Stenger, B., Ramesh, V., Paragios, N., Coetzee, F., Buhmann, J.M.: Topology Free Hidden Markov Models: Application to Background Modeling. In: International Conference on Computer Vision. pp. 294–301 (2001)

28. Suo, P., Wang, Y.: An improved adaptive background modeling algorithm based on Gaussian Mixture Model. In: International Conference on Signal Processing Proceedings (2008)

29. Wang, H., Suter, D.: A re-evaluation of mixture of Gaussian background modeling [video signal processing applications]. In: International Conference on Acoustics, Speech, and Signal Processing. vol. 2 (2005)
30. Wren, C.R., Azarbayejani, A., Darrell, T., Pentland, A.: Pfinder: Real-Time Tracking of the Human Body. IEEE Transactions on Pattern Analysis and Machine Intelligence 19, 780–785 (1997)
31. Yu, J., Zhou, X., Qian, F.: Object kinematic model: A novel approach of adaptive background mixture models for video segmentation (2010)
32. Zang, Q., Klette, R.: Evaluation of an Adaptive Composite Gaussian Model in Video Surveillance (2003)
33. Zhang, L., Liang, Y.: Motion Human Detection Based on Background Subtraction. In: International Workshop on Education Technology and Computer Science (2010)
34. Zivkovic, Z., Heijden, F.V.D.: Recursive Unsupervised Learning of Finite Mixture Models. IEEE Transactions on Pattern Analysis and Machine Intelligence 26, 651–656 (2004)