

# Special Issue on Background Modeling for Foreground Detection in Real-World Dynamic Scenes

Thierry Bouwmans · Jordi González · Caifeng Shan · Massimo Piccardi · Larry Davis

Received: date / Accepted: date

Although background modeling and foreground detection are not mandatory steps for computer vision applications, they may prove useful as they separate the primal objects usually called "foreground" from the remaining part of the scene called "background", and permits different algorithmic treatment in the video processing field such as video-surveillance, optical motion capture, multimedia applications, teleconferencing and human-computer interfaces. Conventional background modeling methods exploit the temporal variation of each pixel to model the background and the foreground detection is made by using change detection. The last decade witnessed very significant publications on background modeling but recently new applications in which background is not static, such as recordings taken from mobile devices or Internet videos, need new developments to detect robustly moving objects in challenging environments. Thus, effective methods for robustness to deal both with dynamic backgrounds, illumination

changes in real scenes with fixed cameras or mobile devices are needed and so different strategies may be used such as automatic feature selection, model selection or hierarchical models. Another feature of background modeling methods is that the use of advanced models has to be computed in real-time and with low memory requirements. Algorithms may need to be redesigned to meet these requirements. Thus, the readers can find 1) new methods to model the background, 2) recent strategies to improve foreground detection to tackle challenges such as dynamic backgrounds and illumination changes, and 3) adaptive and incremental algorithms to achieve real-time applications.

First, Shah et al. [10] adopt the Mixture of Gaussians (MOG) [12] as the basic framework for their complete system. A new online and self-adaptive method permits an automatic selection of the parameters for the GMM. In a second time, they introduce several new solutions to address key challenges such as sudden illumination changes and ghosts. Indeed, a novel hierarchical SURF feature matching algorithm suppresses ghosts in the foreground mask. Moreover, a voting based scheme is used to exploit spatial and temporal information to refine the foreground mask. Finally, temporal and spatial history of foreground blobs is used to detect and handle paused objects. The proposed model shows significant robustness in presence of illumination changes and ghosts.

Shimada et al. [11] propose a novel framework for the GMM to reduce the memory requirement without loss of accuracy. This "case-based background modeling" creates or removes a background model only when necessary. Furthermore, a case-by-case model is shared by some of the pixels. Finally, pixel features are divided into two groups, one for model selection and the other for modeling. This complete approach realizes a low-

---

T. Bouwmans  
Lab. MIA, Univ. of La Rochelle, France  
Tel.: +33-0546457202  
Fax: +33-0546458242  
E-mail: tbouwman@univ-lr.fr

J. González  
Computer Vision Center, Univ. Autònoma de Barcelona, Spain  
E-mail: jordi.gonzalez@uab.cat

C. Shan  
Philips Research, The Netherlands  
E-mail: caifeng.shan@philips.com

M. Piccardi  
Univ. of Technology, Sydney, Australia  
E-mail: massimo.piccardi@uts.edu.au

L. Davis  
CV Lab, Univ. of Maryland, USA  
E-mail: lsd@umiacs.umd.edu

cost and high accurate background model. The memory usage and the computational cost could be reduced by half of the traditional GMM with better accuracy.

Alvar et al. [1] present an algorithm called Mixture of Merged Gaussian Algorithm (MMGA) to reduce drastically the execution time to reach real time implementation, without altering the reliability and accuracy. The algorithm is based on a combination of the probabilistic model of the Mixture of Gaussians (MOG) [12], and the learning process of Real Time Dynamic Ellipsoidal Neural Network (RTDENN) model. Results show that the MMGA achieves a very significant reduction of execution time compared to the MOG with a higher degree of robustness against noise and illumination changes.

Modeling the background using the Gaussian mixture is based on the assumption that the background and foreground distributions are Gaussians which isn't always the case for most environments. Furthermore, it is unable to distinguish between moving shadows and moving objects. In this context, Elguebaly and Bouguila [4] propose a mixture of asymmetric Gaussians to enhance the robustness and flexibility of mixture modeling, and a shadow detection scheme to remove unwanted shadows from the scene.

Narayana and Learned-Miller [8] use a simply use Bayes' rule to classify pixels arguing. Then, their approach uses a background likelihood, a foreground likelihood, and a prior at each pixel. Then, they describe a model for the likelihoods that is built by using not only the past observations at a given pixel location, but by also including observations in a spatial neighborhood around the location. This allows them to model the influence between neighboring pixels. Although similar in spirit to the joint domain-range model, their model overcomes certain deficiencies in that model.

Hernandez-Lopez and Rivera [7] adopt a change detection method to achieve real-time performance. This approach implements a probabilistic segmentation based on the Quadratic Markov Measure Fields (QMMF) model. This framework regularizes the likelihood of each pixel belonging to each one of the classes, that is background or foreground). A likelihood that takes into account two cases. The first one is when the background is static and the foreground might be static or moving. The second one is when the background is unstable and the foreground is moving. Moreover, this likelihood is robust to illumination changes, cast shadows and camouflage situations. Furthermore, the algorithm was implemented in CUDA using a NVIDIA GPU in order to fulfill real-time execution requirement.

Camplani et al. [3] develop a Bayesian framework that is able to accurately segment foreground objects in

RGB-D imagery. In particular, the final segmentation is obtained by considering a prediction of the foreground regions, carried out by a novel Bayesian network with a depth-based dynamic model, and, by considering two independent depth and color-based GMM background models. As a result, more compact segmentations, and refined foreground object silhouettes are obtained.

In another way, Fernandez-Sanchez et al. [5] propose a depth-extended Codebook model which fuses range and color information, as well as a post-processing mask fusion stage to get the best of each feature. Results are presented with a complete dataset of stereo images.

Seidel et al. [9] adopt a Robust PCA model to separate the sparse foreground objects from the background. While many RPCA algorithms use the  $l_1$ -norm as a convex relaxation, their approach uses a smoothed  $l_p$ -quasi-norm Robust Online Subspace Tracking (pROST). The algorithm is based on alternating minimization on manifolds. The implementation on a graphics processing unit (GPU) achieves realtime performance at a resolution of  $160 \times 120$ . Experimental results show that the method succeeds in a variety of challenges such as camera jitter and dynamic backgrounds.

Hagege [6] describes a scene appearance model as a function of the behavior of static illumination sources, within or beyond the scene, and arbitrary three dimensional configurations of patches and their reflectances distributions. Then, a spatial prediction technique was developed to predict the appearance of the scene, given a few measurements within it. The scene appearance model and the prediction technique were developed analytically and tested empirically. Results shows that this scene appearance model permit to detect changes that are not the result of illumination changes at the resolution of single pixels, despite sudden and complex illumination changes, and to do so independently of the texture of the region in the neighborhood of the pixel.

Maritime environment represents a challenging application due to the complexity of the observed scene (waves on the water surface, boat wakes, weather issues). In this context, Bloisi et al. [2] present a method for creating a discretization of an unknown distribution that can model highly dynamic background such as water background with varying light and weather conditions. A quantitative evaluation carried out on the recent MAR datasets demonstrates the effectiveness of this approach.

Zhang et al. [13] propose an effective mosaic algorithm which Combined SIFT and Dynamic Programming (CSDP) for image mosaic which is a useful pre-processing step for background subtraction in videos recorded by a moving camera. To deal with the ghosting effect and mosaic failure, this algorithm uses an im-

proved optimal seam searching criterion that provides a protection mechanism for moving objects with an edge-enhanced weighting intensity difference operator. Furthermore, it addresses the ghosts and incomplete effect induced by moving objects. Experimental results show the effectiveness in the presence of huge exposure difference and big parallax between adjacent images.

**Acknowledgments** We thank all the reviewers for their valuable comments that ensure the high quality of the special issue, and all the contributing authors for their interesting and innovative work. We would also like to thank the current Editor-in-Chief Prof. Mubarak Shah for sharing our vision and providing guidance. The editorial staff of MVA, especially Cherry Place and Shradha Menon, have been extremely supportive, helpful, and patient throughout the entire process.

## References

1. M. Alvar, A. Rodriguez-Calvo, A. Sanchez-Miralles, and A. Arranz. Mixture of merged gaussian algorithm using RT-DENN. *Machine Vision and Applications*, 2013.
2. D. Bloisi, A. Pennisi, and L. Iocchi. Background modelling in the maritime domain. *Machine Vision and Applications*, 2013.
3. M. Camplani, C. Del Blanco, L. Salgad, N. Garca, and F. Jaureguizar. Advanced background modeling with rgb-d sensors through classifiers combination and inter-frame foreground prediction. *Machine Vision and Applications*, 2013.
4. T. Elguebaly and N. Bouguila. Background subtraction using finite mixtures of asymmetric gaussian distributions and shadow detection. *Machine Vision and Applications*, 2013.
5. E. Fernandez-Sanchez, J. Diaz, and E. Ros. Background subtraction model based on color and depth cues. *Machine Vision and Applications*, 2013.
6. R. Hagege. Scene appearance model based on spatial prediction. *Machine Vision and Applications*, 2013.
7. F. Hernandez-Lopez and M. Rivera. Change detection by probabilistic segmentation from monocular view. *Machine Vision and Applications*, 2013.
8. M. Narayana and E. Learned-Miller. Background subtraction - separating the modeling and the inference. *Machine Vision and Applications*, 2013.
9. F. Seidel, C. Hage, and M. Kleinsteuber. pROST : A smoothed lp-norm robust online subspace tracking method for realtime background subtraction in video. *Machine Vision and Applications*, 2013.
10. M. Shah, J. Deng, and B. Woodford. Video background modeling: Recent approaches, issues and our solutions. *Machine Vision and Applications*, 2013.
11. A. Shimada, Y. Nonaka, H. Nagahara, and R. Taniguchi. Case-based background modeling -towards low-cost and high-performance background model. *Machine Vision and Applications*, 2013.
12. C. Stauffer and E. Grimson. Adaptive background mixture models for real-time tracking. *IEEE Conference on Computer Vision and Pattern Recognition, CVPR*, pages 246–252, 1999.
13. L. Zeng, S. Zhang, and Y. Zhang. Dynamic image mosaic via SIFT and dynamic programming. *Machine Vision and Applications*, 2013.