The Importance of Features for Background Modeling and Foreground Detection

Jyoti

M.C. A. dept. M.D. University Rohtak

Abstract: **Background modeling has emerged as a popular foreground detection technique for various applica- tions in video surveillance. Background modeling methods have become increasing efficient in robustly modeling the background and hence detecting moving objects in any visual scene. Although several background subtraction and foreground detection have been proposed recently, no traditional algorithm today still seem to be able to simultaneously address all the key challenges of illumination variation, dynamic camera motion, cluttered background and occlusion. This limitation can be attributed to the lack of systematic investigation concerning the role and importance of features within background modeling and foreground detection. With the availability of a rather large set of invariant features, the challenge is in determining the best combination of features that would improve accuracy and robustness in detection. The purpose of this study is to initiate a rigorous and comprehensive survey of features used within background modeling and foreground detection. Further, this paper presents a systematic experimental and statistical analysis of techniques that provide valuable insight on the trends in background modeling and use it to draw meaningful recommendations for practitioners. In this paper, a preliminary review of the key characteristics of features based on the types and sizes is provided in addition to investigating their intrinsic spectral, spatial and temporal properties. Furthermore, improvements using statistical and fuzzy tools are examined and techniques based on multiple features are benchmarked against reliability and selection criterion. Finally, a description of the different resources available such as datasets and codes is provided.**

Keywords: **: Background modeling, Foreground detection Features, Local binary patterns**

I. INTRODUCTION

Background modeling and foreground detection are important steps for video processing applications in video-surveillance [1], optical motion capture [2], multimedia [3], teleconferencing and human–computer interface. The aim is to separate the moving objects, called ‘‘foreground’’, from the static information, called ‘‘background’’. For example, Fig. 1 shows an original frame of a sequence from the BMC 2012 dataset [4], the reconstructed background image and the moving objects mask obtained from a decomposition into the low-rank matrix and sparse matrix based model [5]. Conventional background modeling methods exploit the temporal variation of each pixel to model the background and hence use it in conjunction with change detection for foreground extraction. The last decade witnessed very significant contributions to this field [5–14]. Despite these works and advances to background modeling and foreground detection, the dynamic nature of visual scenes attributed by changing illumination conditions, occlusion, background clutter and noise have challenged the robustness of such techniques. Under this pretext, focus has shifted towards the investigation of features and their role in improving both the accuracy and robustness of background modeling and foreground detection. Although fundamental low-level features such as color, edge, texture, motion and stereo have reported reasonable success, recent visual applications using mobile devices and internet videos where the background is non-static, require more complex representations to guarantee robust moving ob- ject detection [15]. Furthermore, in order to generalize existing background modeling and foreground detection schemes to real-life scenes where dynamic variations are inevitable and the pose of the camera is little known, automatic feature selection, model selection and adaptation for such schemes are often desired.

Considering the needs and challenges aforementioned, in this paper, a comprehensive review of low-level and hand-crafted features used in background modeling and foreground detection is initiated for benchmarking them against the complexities of typical dynamic scenes. Thus, the aim of this survey is then to provide a first complete overview of the role and the importance of features in background modeling and foreground detection by reviewing both existing and new ideas for (1) novices who could be students or engineers beginning in the field of computer vision, (2) experts as we put forward the recent advances that need to be improved, and (3) reviewers to evaluate papers in journals, conferences, and workshops. In addition, this survey gives a complete overview Moreover, an accompanying website called the Features Website1 is provided. It allows the reader to have a quick access to the main resources, and codes in the field. So, this survey is intended to be a reference for researchers and developers in industries, as well as graduate students, interested in robust background modeling and foreground detection in challenging environments.

Some of the main contributions of this paper can be summa- rized as follows:

1 https://sites.google.com/site/featuresbackgroundforeground/.

– A review regarding feature concepts: A first complete overview of low-level and hand-crafted features used in background modeling and foreground detection over the last decade concerning more than 600 papers. After a pre- liminary overview on the key concepts in the field of fea- tures in Section 2, a survey of spectral features including color features are detailed in Section 4.

1.1.1. *Feature selection*

As seen in Section 2.2.1, there is not a unique feature that performs better than any other feature independently of the back- ground and foreground properties because each feature has its strengths and weaknesses against each challenge. Thus, a way to take advantage of the properties of each feature is to perform feature selection. The aim is to use the best feature or the best com- bination of features on a per-pixel [335–340] or per-block [282] basis. A set of feature could be composed of **(1)** homogeneous features that are feature from the same category, and then the idea is to reinforce the reliability for the concerned type of features, or
(2) heterogeneous features to complement the features each oth- ers [341]. Once the set of features is determined, ensemble learning methods such as the boosting classifier can be used for feature selection. Boosting algorithms usually generate a weighted linear combination of some weak classifiers that perform only a little better than random guess. So, weak classifiers can be learned from the feature values at a pixel and combined to perform better than the others alone. This combination produces a strong classifier. Thus, this method can effectively select different features at each pixel to distinguish foreground objects from the background.

Extensions of this conventional algorithm are available in the form of on-line boosting algorithms [335–337] which use several classifier pools, and each pool contains several weak classifiers. Once an input image is given, each classifier pool selects the best classifier for the given image. The selected classifiers form a strong classifier group, and the final classification is performed using those strong classifiers. At the same time, each classifier pool selects the worst classifier as well. The worst classifier is replaced with a randomly selected classifier so that a better classifier can be included in the classifier pool. Instead of selecting the best classifier from each classifier pool as the previous method does, an improvement according to [282] selects several good classifiers from each pool. While the previous method replaces the worst classifier in each pool, instead this improvement replaces several bad classifiers.

According to the literature in this area, feature selection has been less investigated in background modeling and foreground detection methods with only 9 papers. Practically, only five ap- proaches have so far been used in the literature: ***(1)*** Adaboost [342] used with the classifier-based background model [282,335–337],
(2) Realboost [343] used with the KDE model [338], ***(3)*** dynamic
thefeature selection [344] with OR-PCA model [345], ***(4)*** generic fea- ture selection [346] with the ViBe model [347], and ***(5)*** One-class SVM [339,340]. These different approaches and their characteris- tics are analyzed in Section 18.

II. FEATURE RELEVANCE AND LEARNING

To choose the most discriminative features in a multiple fea- tures or feature selection scheme, feature relevance may be ad- dressed. More generally, feature relevance can be determined in feature learning scheme which can be classified as developed in Zhong et al. [348]:

1. **Traditional feature learning:** This category includes linear algorithms and their kernel extension, and manifold learn- ing method. Practically, an learning algorithm can be linear or nonlinear, supervised or unsupervised, generative or dis- criminative, global or local. For example, Principal Compo- nent Analysis (PCA) is a linear, unsupervised, generative and global feature learning method, while Linear Discriminant Analysis (LDA) is a linear, supervised, discriminative and global method. Global methods aim to preserve the global information of data in the learned feature space, but local ones focus on preserving local similarity between data dur- ing learning the new representations. For instance, unlike PCA and LDA, Locally Linear Embedding (LLE) is a locality- based feature learning algorithm. Locality-based feature learning like LLE as manifold learning, since it is to discover the manifold structure hidden in the high dimensional data.
2. **Deep learning algorithms:** Deep learning models includes
models like Convolutional Neural Network (CNN) [349] and Recurrent neural network (RNN). A survey of deep learning models can be found in Schmidhuber [349].

Feature relevance has been less investigated in background modeling and foreground detection methods than manual im- age feature methods, such as Local Binary Patterns (LBP) [31], histogram of oriented gradients (HOG) [139], and Scale-Invariant Feature Transform (SIFT) [252]. For *traditional feature learning*, the one work which concerns feature relevance is the work of Molina-Giraldo et al. [350,351]. The feature relevance analysis is made through a Principal Component Analysis (PCA), searching for directions with greater variance to project the data. Thus, the relevance of the original features is quantified with weight- ing factors. Finally, Molina-Giraldo et al. [350,351] developed a background subtraction method based a multi-kernel learning in which the weight are selected from the feature relevance analysis. Experimental results [350,351] on the I2R dataset [109] show that the proposed Weighted Gaussian Kernel Video Segmentation (WGKVS) model outperforms SOBS [352]. For *deep learning algo- rithms*, the approaches available in literature can be classified as follows: *(1)* Deep Auto-encoder Networks (DAN) [16,353,354], *(2)* Convolutional Neural Networks (CNN) [17,18,355,356], *(3)* Neural Response Mixture (NeREM) [357].

III. FEATURES AND CHALLENGES

In this section, we grouped all the advantages and disadvan- tages of the different features in terms of robustness against the different challenges met in video and detailed in Bouwmans [9], and they can be summarized as follows:

- **Color features:** Although intensity and color features are often very discriminative features and allow basic fore- ground detection, they are not robust in challenges such as illumination changes,

foreground aperture, camouflage in color and shadows. However, intensity can be used in complementarity of color to deal with different color problems such as dark foreground and light foreground. Furthermore, this combination solves saturation problems and minimum intensity problems [358], and reduces the number of false negatives, false positives and increase true positives. But, the intensity as colors cannot work with intense shadows and highlight that often occur in indoor and outdoor scenes, and in presence of gradual or sudden illumination changes [359]. Then, different strategies can be found in literature to alle- viate the limitations of the basic color spaces: **(1)** the use of well-known color spaces which separate the luminance and the chrominance information such as HSV and YCrCb,
(2) the use of designed shape color space models such as the cylinder color model [235,237,360,361], the hybrid cone- cylinder [236,362], the ellipsoidal color model [238], the box-based color model [239], and the double-trapezium cylinder model [242], **(3)** the use of characteristics in ad- dition of the intensity or color value (mean, variance, min- imum, maximum, etc.) (see Section 16), **(4)** the use of de- signed illumination invariant intensity or color features ob- tained by normalization [205,219,243,363], **(5)** the use of illumination compensation methods [364–373], and **(6)** the addition of other features (see Section 17). Normalization based features sacrifice discriminability while texture fea- tures cannot operate on texture-less regions. Both types of features produce large missing regions in the foreground mask.

– **Edge features:** Edge features are obtained with edge de- tectors which operate on the difference between neigh- boring pixels, hence an edge detector should be reason- ably insensitive to global shifts in the mean level, i.e. to global illumination changes. Therefore it is interesting to run background/foreground separation algorithms on the output from edge detectors, hopefully reducing the effects of rapid illumination changes. So, the edge could handles the local illumination changes but also the ghost leaved when waking foreground objects begin to move. However, edge features are not sufficiently good to segment the foreground objects isolatedly. Indeed, edge features can sometimes han- dle dark and light camouflage problems and it is less sensi- tive to global illumination changes than color feature [111]. Nevertheless, problems like noise, false negative edges due to local illumination problems, foreground aperture and camouflage do not allow an accurate foreground detection. Furthermore, due to the fact that it is sometimes difficult to segment the foreground object borders, it is not possible to fill the objects, and solve the foreground aperture problem. Since it is not possible to handle dark and light camouflage problems only by using edges due to the foreground aper- ture difficulty, the brightness of color model is used to solve this problem and help to fill the foreground objects.
– **Texture features:** Texture features allow to be robust in presence of shadows and gradual illumination changes, and sometimes in dynamic backgrounds. Texture features can produce false detections due to textures induced by local illumination effects like in cast shadows. Furthermore, an algorithm based only on texture may cause detection errors in regions of blank texture and heterogeneous texture.
– **Motion features:** Motion features can handle irrelevant background motion and clutter such as waving trees and waves.
– **Stereo features:** Stereo features allow the model to deal with the camouflage in color but not in depth.

Thus, multiple features approaches with two, three or a set of features obtained from a bag-of features or by feature selection are suitable to address multiple challenges in the same video (see Section 17). A representative work developed by Li et al. [109] consists in a sets of features built following the type of background (static or dynamic) as follows:

– **Features for static background pixels:** For modeling pixels belonging to a stationary background object, the stable and most significant features are its color and local structure (gradient). As the gradient is less sensitive to illumination changes, the two types of feature vectors are integrated under the Bayes framework in the basic product formulation of the likelihoods.
– **Features for dynamic background pixels:** For modeling dynamic background pixels associated with non stationary objects, color co-occurrences are used as their dynamic features. This is because the color co-occurrence between consecutive frames has been found to be suitable to describe the dynamic features associated with non stationary back- ground objects, such as moving tree branches or a flickering screen.

IV. FEATURES AND STRATEGIES

There are several strategies in literature such as multi-scales strategies, multi-levels strategies, multi-resolutions strategies, multi-layers strategies, hierarchical strategies, and coarse-to-fine strategies (see Section 2.1). Practically, different features can be used following the scale, the level or the resolution. For example, a feature can be used at the block level (such as Haar-like features in [94]), and other features can be used at the pixel level (such as RGB in [94]). Thus, these strategies employed multiple features schemes. Please see Tables 9–11 for a quick overview.

1.1.2. *Features and similarities*

The foreground mask is usually obtained from a similarity/dissimilarity measure between *(1)* the direct value of the fea- ture in the background model and the current frame, or *(2)* a value computed from the direct value of the feature (mean, variance, probability, etc...) in the background model and the current frame. This value can be a scalar (intensity value, mean, probability, etc.), a vector (2D spatial vector, 3D spatiotemporal vector, etc...) or a his- togram (correlogram, etc.). Practically, comparison of features can be made by using similarity/dissimilarity measures obtained with **(1)** a crisp, statistical or fuzzy distance for scalar cases, **(2)** a ratio for scalar cases, **(3)** linear dependence measure for vector cases, and a intersection measure for histogram (correlogram) case. The choice of the suitable similarity/dissimilarity measure is guided by the properties and the distribution of the concerned features. Furthermore, spatial and temporal features such as LBP and LTP need also measures for their computing as follows: **(1)** a measure for the distance in the spatial neighborhood, and **(2)** a measure for the distance in the temporal neighborhood. Thus, for spatial and temporal features like texture, it needs to choose three distances. We list below the different similarity/dissimilarity measures used in the literature for foreground detection (see Table 8 for a quick overview):

(A) **Similarities for scalar case:** Scalar value is the most common case in the literature and the similarities used can be classified as follows:

– **Difference:** The difference computed in a pixel-wise man- ner between the feature in the background model and the current frame is the most measure used. So, the difference is obtained by a distance and then a threshold is used to classify the pixel as background or foreground as follows:

$$distance(B(x, y) - I(x, y)) < T \quad (4)$$

where $B(x,y)$ and $I(x,y)$ are the values of the feature in the background image and in the current image, respectively. *distance*(,) is a distance function. Several distance functions have been used in the literature and they can be classified as follows:

1. **Crisp distance:** The most common distance function used for intensity/color values is the absolute dis- tance [221,374]. Aach et al. [375] used a total least squares distance measure. In an other work, Yadav and Sing used a quasi-euclidean distance. To compare Spatiotemporal Condition Information (SCI), Wang et al. [38] designed a specific measure called Neigh- borhood Weighted Spatiotemporal Condition Infor- mation (NWSCI). Using compressive features [376], Yang et al. [377] developed a (Pixel-to-Model) P2M distance.
2. **Statistical distance:** To compare the K distribution in the original MOG, Stauffer and Grimson [20] used the Mahalanobis distance with the RGB features. An alter- native to the Mahalanobis distance is the Kullback– Leibler (KL) divergence used in Makantasis et al. [378] with the infrared features and Patwardhan et al. [379] with the RGB features. In a further work, Pavlidis et al. [380] claimed that the MOG algorithm needs a divergence measure between two distributions so that if the divergence measure between the new distribu- tion and one of the existing distributions is ‘‘too small’’, these two distributions could be merged together. Thus, Pavlidis et al. [380] used the Jeffreys divergence measure to check if the incoming pixel value can be ascribed to any of the existing K Gaussians. Experi- mental results presented by Pavlidis et al. [380] show that the false positives are reduced in comparison with the Mahalanobis distance and the KL divergence. In an other work, Santoyo-Morales and Hasimoto-Beltran used the Chi-2 distance with YUV features instead of the Mahalanobis distance. In a non parametric model based on KDE, Ko et al. [381] choose the Bhattacharyya distance due to its low computational cost. In an other work, Mukherjee et al. [83] developed a distance measure based on support weight to compare RGB features. St-Charles and Bilodeau [382] employed the Hamming distance to compare LSBPs.

Order-Consistency Measure: Xie et al. [189] used an explicit model for the camera response function, the camera noise model, and illumination prior. Assum- ing a monotone and nonlinear camera response func- tion, Xie et al. [189] show that the sign of the differ- ence between two pixel measurements is maintained across global illumination changes. Noise statistics are used to transform each frame into a confidence frame where each pixel is replaced by a probability that it is likely to keep its sign with respect to the most different pixel in its neighborhood. Hence, an order consistency measure is defined as a distance between two distri- butions. Xie et al. [189] used the Bhattacharyya dis- tance due to its properties to the Bayes error. Finally, an Illumination Invariant Change Detector via order consistency (IICD-OC) is developed. Experimental re- sults [189] on videos taken by an omni-directional camera show the robustness of IICD-OC against illu- mination changes. But, the ordinal measure required a reordering of blocks and it is computationally ex- pensive. To solve this problem, Singh et al. [383] ex- plicitly modeled noise under which rank-consistency is tested, and used a probabilistic generative model under which frame blocks are generated. The order-consistency is posed as a hypothesis validation prob- lem using fast significance testing based on PAV. In a further work, Parameswaran et al. [373] used the same order-consistency measure in an illumination compensation approach

V. CONCLUSION

In conclusion, this review on the role and the importance of features for background modeling and foreground detection high- lights the following points:

Features can be classified following their size, their type in a specific domain, their intrinsic properties and their math- ematical concepts. Each type of features presents different robustness against challenges met in videos taken by a fixed cameras. For the color feature, YCrCb color space seems to be the more suitable feature [105,384]. For the texture feature, Silva et al. [339] provided a study on the LBP and its variants that show that XCS-LBP is the best LBP feature for this application in presence of illumination changes and dynamic backgrounds. Although this study covered texture features, it is restricted to LBP features and then there is not a full study on the different texture features. For the depth feature, it needs to carefully used them following their properties as developed in Nghiem and Bremond [560]. Features in a domain transform are very useful to reduce computation times as in the case of compressive sensing features.

Several features have been used in other applications and none in background modeling and foreground detection such several variants of LBP (Multi-scale Region Perpen- dicular LBP (MRP-LBP) [299], Scale- and Orientation Adap- tive LBP (SOA-LBP) [300]). Furthermore, statistical or fuzzy version of crisp feature could be investigated such as his- tograms of fuzzy oriented gradients [202]. It would be in- teresting to evaluate them for this application.

Because each feature has its strengths and weaknesses against each challenge, multiple features schemes are used to combine the advantages of their different robustness. Most of the time, gradient, texture, motion and stereo fea- tures are used in addition to the color feature to deal with camouflage in color, illumination changes, dynamic back- grounds and shadows. Different fusion operators can be used to combine these different features but fuzzy integrals such as the Choquet integral [330] and interval-valued Cho- quet [188] seem the best way to aggregate different features because dependency between features can be taken into account.Because there is not a unique feature that performs better than any other feature independently of the background and foreground properties, feature selection allows to use the best feature or the best combination of features. Exper- imental results provided by the existing approaches show the pertinence of feature selection in background modeling and foreground detection. However, basic algorithms such as Adaboost and Realboost have been used most of the time. The most advanced scheme is the IWOC-SVM algorithm developed by Silva et al. [339], but more advanced selection schemes can be used such as statistical or fuzzy feature selection.

To summarize, the most interesting approach seems to fuse mul- tiple features with the intervalued fuzzy Choquet integral. The best set of features seems to be illumination invariant color fea- tures combined with spatio-temporal texture features and depth features. Future research should concern (1) a full evaluation of texture features, (2) a full comparison of feature fusion schemes, feature selection schemes and (4) reliability of features be- cause it has been less investigated. Finally, features learned by deep learning methods such as Stacked Denoising Auto-Encoder (SDAE) [16] and Convolutional Neural Networks (CNN) [17,18] are surely the features that will outperforms all the other features because deep learning methods have the sole ability of learning features that best fit a given set of data. Furthermore, unlike conventional hand-crafted features, learned features come from multiple layers which focus on various level of details in the video. Thus, learned feature representation allows to well capture the intrinsic structural properties of a scene and adaptively discover a set of filter patterns that are robust to complicated factors such as noise and illumination variation.

VI. REFERENCES

[1]. S. Cheung, C. Kamath, Robust background subtraction with foreground vali- dation for urban traffic video, EURASIP J. Appl. Signal Process. (2005).

[2]. J. Carranza, C. Theobalt, M. Magnor, H. Seidel, Free-viewpoint video of human actors, ACM Trans. Graph. 22 (3) (2003) 569–577.

[3]. F. El Baf, T. Bouwmans, B. Vachon, Comparison of background subtraction methods for a multimedia learning space, in :International Conference on Signal Processing and Multimedia, SIGMAP 2007, July 2007.

[4]. Vacavant, T. Chateau, A. Wilhelm, L. Lequievre, A benchmark dataset for foreground/background extraction, in: International Workshop on Back- ground Models Challenge, ACCV 2012, November 2012.

[5]. T. Bouwmans, A. Sobral, S. Javed, S. Jung, E. Zahzah, Decomposition into low- rank plus additive matrices for background/foreground separation: A review for a comparative evaluation with a large-scale dataset, Comput. Sci. Rev. (2016).

[6]. T. Bouwmans, F.E. Baf, B. Vachon, Background modeling using mixture of gaussians for foreground detection - A survey, Recent Pat. Comput. Sci. RPCS 2008 1 (3) (2008) 219–237.

[7]. T. Bouwmans, Subspace Learning for Background Modeling: A Survey, Recent Patents on Computer Science, RPCS 2009 2 (3) (2009) 223–234.

[8]. T. Bouwmans, Recent advanced statistical background modeling for fore- ground detection: A systematic survey, Recent Pat. Comput. Sci. RPCS 2011 4 (3) (2011) 147–176.

[9]. T. Bouwmans, Traditional and recent approaches in background modeling for foreground detection: An overview, Comput. Sci. Rev. 11 (2014) 31–66.

[10]. T. Bouwmans, E. Zahzah, Robust PCA via principal component pursuit: A review for a comparative evaluation in video surveillance, Comput. Vis. Image Underst. 122 (2014) 22–34 Special Isssue on Background Models Challenge.

[11]. T. Bouwmans, F. El Baf, B. Vachon, Statistical Background Modeling for Foreground Detection: A Survey, in: Part 2, Chapter 3, Handbook of Pattern Recognition and Computer Vision, vol. 4, World Scientific Publishing, Prof C.H. Chen, 2010, pp. 181–199.

[12]. T. Bouwmans, Background subtraction for visual surveillance: A fuzzy ap- proach, in: S.K. Pal, A. Petrosino, L. Maddalena (Eds.), Handbook on Soft Computing for Video Surveillance, Taylor and Francis Group, 2012, pp. 103– 139 (Chapter 5).

[13]. C. Guyon, T. Bouwmans, E. Zahzah, Robust principal component analysis for background subtraction: Systematic evaluation and comparative analysis, in: Principal Component Analysis, INTECH, 2012, pp. 223–238 (Book 1, Chap- ter 12).

[14]. T. Bouwmans, F. Porikli, B. Hoferlin, A. Vacavant, Handbook on Background Modeling and Foreground Detection for Video Surveillance, Chapman and

Hall/CRC, 2014.
[15]. T. Bouwmans, J. Gonzalez, C. Shan, M. Piccardi, L. Davis, Special issue on back- ground modeling for foreground detection in real-world dynamic scenes, Mach. Vis. Appl. (2014) (special issue).
[16]. Y. Zhang, X. Li, Z. Zhang, F. Wu, L. Zhao, Deep learning driven blockwise mov- ing object detection with binary scene modeling, Neurocomputing (2015).
[17]. M. Braham, M. Van Droogenbroeck, Deep background subtraction with scene-specific convolutional neural networks, in: International Conference on Systems, Signals and Image Processing, IWSSIP 2016, 2016.
[18]. Y. Wang, Z. Luo, P. Jodoin, Interactive deep learning method for segmenting moving objects, Pattern Recognit. Lett. (2016) Special Issue on Scene Back- ground Modeling and Initialization.
[19]. P. Viola, M. Jones, Rapid object detection using a boosted cascade of simple features, Comput. Vis. Pattern Recognit. (2001).
[20]. C. Stauffer, E. Grimson, Adaptive background mixture models for real-time tracking, in: IEEE Conference on Computer Vision and Pattern Recognition, CVPR 1999, 1999, pp. 246–252.
[21]. X. Fang, W. Xiong, B. Hu, L. Wang, A moving object detection algorithm based on color information, in: International Symposium on Instrumentation Science and Technology, vol. 48, 2006, pp. 384–387.
[22]. H. Bhaskar, L. Mihaylova, A. Achim, Video foreground detection based on symmetric alpha-stable mixture models, IEEE Trans. Circuits Syst. Video Technol. (2010).
[23]. Y. Lin, Y. Tong, Y. Cao, Y. Zhou, S. Wang, Visual-attention based background modeling for detecting infrequently moving objects, IEEE Trans. Circuits Syst. Video Technol. (2016).
[24]. N. Jacobs, R. Pless, Shape background modeling : The shape of things that came, in: IEEE Workshop on Motion and Video Computing, WMVC 2007, 2007, pp. 1–7.
[25]. P. Jodoin, V. Saligrama, J. Konrad, Behavior subtraction, IEEE Trans. Image Process. (2011).
[26]. S. Javed, S. Oh, A. Sobral, T. Bouwmans, S. Jung, Background subtraction via superpixel-based online matrix decomposition with structured foreground constraints, in: Workshop on Robust Subspace Learning and Computer Vi- sion, ICCV 2015, 2015.
[27]. Zheng, M. Xu, H. Shi, B. Luo, C. Li, CLASS: Collaborative low-rank and sparse separation for moving object detection, Cogn. Comput. (2017) 1–14.
[28]. S. Varadarajan, P. Miller, H. Zhou, Region-based mixture of gaussians mod- elling for foreground detection in dynamic scenes, Pattern Recognit. 48 (11) (2015) 3488–3503.
[29]. S. Varadarajan, P. Miller, H. Zhou, Spatial mixture of Gaussians for dynamic background modelling, in: IEEE International Conference on Advanced Video and Signal Based Surveillance, AVSS 2013, 2013, pp. 63–68.
[30]. M. Heikkila, M. Pietikainen, A texture-based method for modeling the back- ground and detecting moving objects, IEEE Trans. Pattern Anal. Mach. Intell. 28 (4) (2006) 657–662.
