Comparative Analysis of Background Subtraction Models Applied on a Local Dataset Using a New Approach for Ground-truth Generation

https://doi.org/10.3991/ijes.v10i03.34317

Maryam A. Yasir[1(✉)], Yossra Hussain Ali[2]
[1]University of Baghdad, Baghdad, Iraq
[2]University of Technology, Baghdad, Iraq
maryam.a@sc.uobaghdad.edu.iq

Abstract—Background subtraction is the dominant approach in the domain of moving object detection. Lots of research have been done to design or improve background subtraction models. However, there is a few well known and state of the art models which applied as a benchmark. Generally, these models are applied on different dataset benchmarks. Most of the time choosing appropriate dataset is challenging due to the lack of datasets availability and the tedious process of creating the ground-truth frames for the sake of quantitative evaluation. Therefore, in this article we collected local video scenes for street and river taken by stationary camera focusing on dynamic background challenge. We presented a new technique for creating ground-truth frames using modelling, composing, tracking, and rendering each frame. Eventually we applied nine promising benchmark algorithms used in this domain on our local dataset. Results obtained by quantitative evaluations exposed the effectiveness of our new technique for generating the ground-truth scenes to be benchmarked with the original scenes using number of statistical metrics. Furthermore, results show the outperformance of SuBSENSE model against other tested models.

Keywords—video surveillance, background subtraction, moving objects detection, ground-truth, evaluation metrics

1 Introduction

Detection of moving object is the initial and essential workflow step of the video surveillance system where further steps could be taken for classifying the moving objects. Background subtraction techniques have been used to detect the moving objects along with optical flow and temporal differencing techniques, although background subtraction is the most well-known and still the dominant technique among others due to the easiness and high performance when implemented in a vast scope of video surveillance environments [1]. Normally, Background subtraction techniques deal with different challenges according to the various environments. For example, challenges can be, dynamic background, illumination changes, camera jitter, low frame rates or other challenges [2].

Through decades, many background subtraction models have been developed and introduced to tackle the background subtraction challenges, these models are classified into different categories by several surveys and review articles, [3][4]. ***Basic models*** are the simplest models they depend on a threshold to decide whether it's a foreground or background pixel and it is more convenient to models with a single background distribution [5], mean model [6], median model [7], and histogram analysis model [8] are examples of the basic models. ***Filter models*** expect the background depending on the intensity or orientation of the previous pixels [9], Wiener filter [10], Tchebychev filter [11], Correntropy filter [12], optical flow [13] and Kalman filter [14], are examples of the single processing models (filter models). ***Mathematical models*** consist of two classes, statistical parametrical models and statistical non-parametrical models. Gaussian Mixture Model (GMM) [15] is example of parametrical statistical models [16]. On the other hand, Visual Background extractor (ViBe) [17], Substance Sensitivity Segmenter (SuBSENSE) algorithm, kernel density estimation (KDE) [18] and fuzzy models [19] are examples of non-parametrical statistical models [3].

Clustering models depend on the color intensity of the pixel to recognize whether a pixel belongs to the background or the foreground clusters, K-means [20], Codebook [21] and background reconstruction [22] are examples of clustering models.

Furthermore, ***Machine learning models*** are the state-of-the-art models which encompass various techniques like, support vector machines (SVM) [23], robust subspace tracking [24], reconstructive and discriminative subspace learning techniques [25][26], deep learning neural networks [27][28] and convolutional neural network (CCN) [29] which have been broadly embraced due to the massive development of hardware processing power [30].

Fusion of several models and strategies is another approach that has been adopted in many articles for a better performance, real time semantic background subtraction (RT-BSB) [31] is an example. Background subtraction technique is mainly consisting of pipelined stages, four main stages and two secondary stages. The main stages are: ***background initialization***, ***background modelling***, ***background maintenance*** and ***background detection*** where the secondary stages are: ***The pre-processing*** and ***post-processing*** stages.

Generally, background subtraction models are applied on a benchmark dataset, choosing a suitable dataset is quite challenging due to the lack of available datasets that providing a specific challenge to be tested or providing the ground-truth frames for quantitative evaluations. As a result only few datasets can be considered as a benchmark, CDnet 2012 and CDnet 2014 [2] are the most well-known benchmark datasets used in this domain for providing a variety of challenges along with supporting the ground-truth frames. Therefore, in this article we shed the light on how to collect a local dataset with dynamic background challenges using stationary camera and proposing a new technique for creating the ground-truth frames for quantitative evaluation. We applied the most well-known and benchmarked background subtraction algorithms on our local dataset for qualitative and quantitative assessment.

2 Local video scenes

In addition to the global benchmark dataset used by researchers around the world, in this article, we attempt to collect local scenes to be tested by benchmark background subtraction models. The utilization of a locally gathered dataset along with the global benchmark datasets is of importance as it adds the value of testing algorithms on local scenarios captured using locally widespread technology, as the use of Closed-Circuit Television (CCTV) cameras increased significantly with the rise of security demands. For this sake we captured three videos with different dynamic background challenges and distinct moving objects to be detected. The video scenes were taken by a common stationary CCTV camera with a frame rate of 30 fps. Furthermore, we followed a new technique to generate the ground-truth frames for each input frame. Table 1 contains details about the local dataset videos.

2.1 Original video scenes

In this section we present the details with sample frames from each original video captured as follows:

Street scenes. In the street scenes, two different moving vehicles are captured in both right and left directions. the dynamic background challenge is the moving palm frond in the upper right corner of the scene frame. The following Figure 1 depicts frame samples of street videos, the scenes (a) and (c) are before the emergence of the target moving vehicle, where (b) and (d) are the scenes with the target moving vehicle.

(a) **(b)** **(c)** **(d)**

Fig. 1. Frame samples before and after the emergence of the target moving vehicle

Tigress river scene. Tigress river scene is recorded from the riverbank of Tigress capturing the river and the moving ferry target. In this scenario, the dynamic background challenge is the moving water surface. Figure 2 depicts the frame samples where (a) is the original scene before the emergence of the moving ferry and (b) is the frame containing the target moving ferry.

Fig. 2. Frame samples before and after the emergence of the target moving ferry

Table 1. Dynamic background videos from local dataset

Videos	Dynamic Challenge Description	Total Frames	Region of Interest Frames
Van	Moving palm frond in the upper right corner	346	188–313
Car	Moving palm frond in the upper right corner	491	254–383
Ferry	Moving water surface	3811	235–3412

2.2 Ground-truth scenes

In this section we present the steps of generating the correspondent ground-truth frames to each original frame in our video scenes.

Modelling. Generally, 3D modeling is creating edges, polygons and vertices for an object through representing mathematical coordinates of the object's shell in a virtual three dimensions space using specific applications [32]. In addition, there are different approaches of generating a 3D model, the manual approach where the designer is responsible for creating the model from scratch, the procedural approach is when designer follows an algorithmic steps to build the 3D model, and scanning is another approach where designer scan the real object in order to create the 3D model [33].

In this work three moving objects were modelled manually using Autodesk 3ds Max application, the following Figure 3 depicts the modelling perspectives for the three moving objects (Van, car, and ferry).

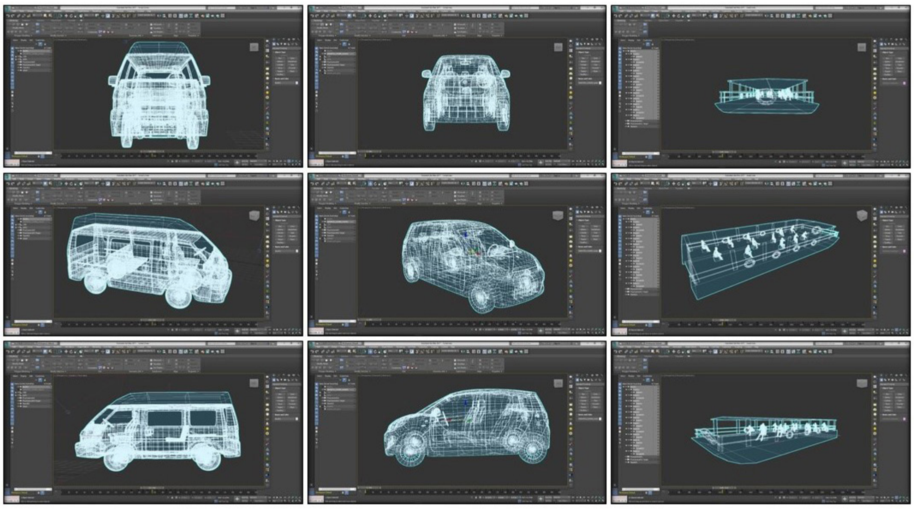

Fig. 3. Modelling perspectives

Composing. In general, this step involves choosing the suitable camera lenses, controlling the camera angle, posing and blocking, if the object is static also, it could involve some specific composing techniques [34].

In this work we used, the real dimensions for the objects and the real camera distance with lenses similar to the CCTV camera lens. The following Figure 4 depicts the composition screenshots.

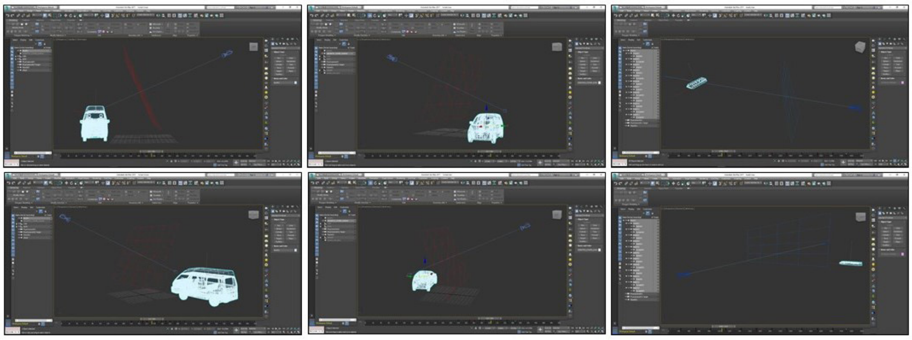

Fig. 4. Composing the camera to the objects

Tracking. Animating the moving object is the approach we followed to imitate and track the movement of our objects in the original scene. The following Figure 5 showing the screenshots of tracking the moving objects.

(a) **(b)** **(c)**

Fig. 5. Tracking (a) van (b) car (c) ferry

Rendering. Rendering is the final step where; the designer generates a moving object based on the 3D scene. In this work the negative mode is used to create the scene. The following Figure 6 showing the rendering process.

(a) (b) (c)

Fig. 6. Rendering on (a) van (b) car and (c) ferry

3 Experiments and analysis

In this article we compare the foreground detection using different nine benchmarked background subtraction algorithms. The models are tested using our local dataset considering the dynamic background challenge. In these experiments we are comparing the background subtraction image with the correspondent ground-truth image to evaluate the performance of each method in respect to quantitative evaluation metrics at the pixel level, and the background subtraction method classifies the pixels into background or foreground. Seven metrics are used for the performance quantitative evaluation as illustrated in the following equations:

$$Accuracy = \frac{TP + TN}{TP + FN + TN + FN} \tag{1}$$

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

$$Recall(sensitivity) = \frac{TP}{TP + FN} \tag{3}$$

$$F - Measures(F1) = \frac{2 * Precision * Recall}{Precision + Recall} \tag{4}$$

$$FPR = \frac{FP}{FP + TN} \tag{5}$$

$$FNR = \frac{FN}{TP + FN} \tag{6}$$

$$Error\,rate(PWC) = \frac{FP + FN}{TP + FN + TN + FP} * 100 \tag{7}$$

Where, *TP* is the number of foreground pixels correctly classified, *TN* is the number of background pixels correctly classified, *FP* is the number of background pixels incorrectly classified as foreground pixels, and *FN* is the number of foreground pixels incorrectly classified as background pixels.

Accuracy in Eq. (1) indicates the correct classification for a pixel whether it is a foreground or a background pixel, *Precision* in Eq. (2) indicates the proportion of truly detected foreground pixels to the number of all pixels classified as foreground pixels, *recall* in Eq. (3) indicates the number of pixels that are correctly classified as a foreground of all the foreground pixels and the *F-measure* in Eq. (4) is the harmonic mean of recall and precision. On the other hand, we have the metrics: (False Positive Rate) *FPR* in Eq. (5) is the number of background pixels that are misclassified as foreground pixels, (False Negative Rate) *FNR* in Eq. (6) is the number of foreground pixels that are misclassified as background pixels, and **P**ercentage of **W**rong **C**lassifications (*PWC*) in Eq. (7) indicates the error rate which is the percentage of misclassified pixels to the original pixels [3].

Normally, we measure the relevance by recall and precision. A low recall is an indication of over segmentation of the foreground objects, where a low precision is an indication of under segmentation of the foreground objects. High F-measures is an indication of a robust background subtraction algorithm and the lower the FPR, FNR and PWC is an indication of a better performance.

In the following Tables 2–10 we illustrate the analytical metrics results of applying the benchmark background subtraction models SuBSENSE [16], ViBe [35], LOBSTER [36], GMM [15], KNN [37], KDE [18], Fuzzy Choquet Integral [38], Fuzzy Sugeno Integral [38], and Codebook [39] respectively on our local dataset videos, providing detailed results, the highest F1 is highlighted in bold. While Table 11 illustrates the average performance metrics of the forementioned models on the local dataset, the best result in each metric is highlighted in bold.

Table 2. Performance metrics results of applying **SuBSENSE** model

Video	Accuracy	Precession	Recall	F1	FPR	FNR	PWC
Car	0.9970	0.7957	0.6741	**0.766**	0.0013	0.3259	0.3034
Van	0.9896	0.8370	0.5338	0.6665	0.0024	0.4662	1.0412
Ferry	0.9948	0.4037	0.811	0.5398	0.0045	0.189	0.5211

Table 3. Performance metrics results of applying **ViBe** model

Video	Accuracy	Precession	Recall	F1	FPR	FNR	PWC
Car	0.9935	0.2806	0.3962	0.5182	0.0015	0.6038	0.6508
Van	0.9853	0.2795	0.4164	**0.5383**	0.0033	0.5836	1.4717
Ferry	0.9958	0.4102	0.3346	0.376	0.0016	0.6654	0.4202

Table 4. Performance metrics results of applying **LOBSTER** model

Video	Accuracy	Precession	Recall	F1	FPR	FNR	PWC
Car	0.9946	0.7757	0.4589	**0.6317**	0.0014	0.5411	0.5428
Van	0.9862	0.8545	0.4011	0.5727	0.0031	0.5989	1.3808
Ferry	0.9965	0.5231	0.7039	0.5863	0.0023	0.2961	0.3511

Table 5. Performance metrics results of applying **GMM** model

Video	Accuracy	Precession	Recall	F1	FPR	FNR	PWC
Car	0.9575	0.1494	0.7303	0.6309	0.0414	0.2697	4.252
Van	0.985	0.2482	0.6351	**0.6445**	0.0092	0.3649	1.4952
Ferry	0.9942	0.2299	0.3861	0.3088	0.0039	0.6139	0.584

Table 6. Performance metrics results of applying **KNN** model

Video	Accuracy	Precession	Recall	F1	FPR	FNR	PWC
Car	0.9830	0.1941	0.6436	**0.6555**	0.0152	0.3564	1.7045
Van	0.9774	0.2619	0.5862	0.6294	0.0157	0.4138	2.2598
Ferry	0.9937	0.3241	0.8654	0.5255	0.006	0.1346	0.6330

Table 7. Performance metrics results of applying **KDE** model

Video	Accuracy	Precession	Recall	F1	FPR	FNR	PWC
Car	0.9748	0.0948	0.5965	**0.4382**	0.0227	0.4035	2.5232
Van	0.9242	0.1066	0.7473	0.3891	0.0735	0.2527	7.5783
Ferry	0.9727	0.089	0.7213	0.1834	0.0261	0.2787	2.7350

Table 8. Performance metrics results of applying **Fuzzy Choquet Integral** model

Video	Accuracy	Precession	Recall	F1	FPR	FNR	PWC
Car	0.9926	0.5117	0.2523	0.402	0.0008	0.7477	0.7405
Van	0.9852	0.6292	0.3405	**0.4895**	0.0019	0.6595	1.4774
Ferry	0.9942	0.1366	0.1129	0.1198	0.0022	0.8871	0.5800

Table 9. Performance metrics results of applying **Fuzzy Sugeno Integral** model

Video	Accuracy	Precession	Recall	F1	FPR	FNR	PWC
Car	0.9922	0.5343	0.2097	0.3487	0.0007	0.7903	0.7818
Van	0.9848	0.6413	0.3136	**0.4624**	0.0017	0.6864	1.5211
Ferry	0.9946	0.1363	0.0866	0.1039	0.0017	0.9134	0.5414

Table 10. Performance metrics results of applying **Codebook** model

Video	Accuracy	Precession	Recall	F1	FPR	FNR	PWC
Car	0.9499	0.0626	0.7211	**0.3415**	0.0489	0.2789	5.0112
Van	0.8212	0.0536	0.7684	0.2331	0.1825	0.2316	17.8830
Ferry	0.9569	0.067	0.8708	0.1469	0.0427	0.1292	4.3053

Table 11. Average performance metrics results of applying background subtraction models

Model	Accuracy	Precession	Recall	F1	FPR	FNR	PWC
SuBSENSE	**0.994**	0.679	0.673	**0.657**	0.003	0.327	**0.622**
ViBe [35]	0.992	0.323	0.382	0.478	0.002	0.618	0.848
LOBSTER	0.992	**0.718**	0.521	0.597	0.002	0.479	0.758
GMM [40]	0.979	0.209	0.584	0.528	0.018	0.416	2.111
KNN [8]	0.985	0.260	0.698	0.604	0.012	0.302	1.532
KDE	0.957	0.097	0.688	0.337	0.041	0.312	4.279
Fuzzy Choquet Integral	0.991	0.426	0.235	0.337	0.002	0.765	0.933
Fuzzy Sugeno Integral	0.991	0.437	0.203	0.305	**0.001**	0.797	0.948
Codebook	0.909	0.061	**0.787**	0.241	0.091	**0.213**	9.067

Overall, the forementioned comparison results of each model on each video from the local datasets, the observations have been made specifically on the F-measure metric results, this is due to the fact that it is the harmonic mean of recall and precision, and F-measure is the most important metric to be considered for evaluating the overall robustness of the background subtraction model. The overall results of all models shows that the best performance in F-measure is achieved more frequently on the car and van videos.

This is generally indicating that models behave similarly to the environmental video challenges even though each model produces its own performance results. What can be noticed from Table 11 is that SuBSENSE on average outperforms all other models in terms of accuracy, F1 measure and PWC metrics when applied on the local dataset.

Moreover, Figures 7–9 depict the visual results comparison of the obtained foreground masks from applying the models on the local dataset. Looking at figures where qualitative evaluation is illustrated; one can see that the results generated by SuBSENSE is visually the closest to the ground-truth frame.

(a) (b) (c) (d) (e)

Fig. 7. Comparison of foreground detection results (a) original scenes (b) ground-truths (c) SuBSENSE subtraction (d) ViBe subtraction (e) LOBSTER subtraction

(a) (b) (c) (d) (e)

Fig. 8. Comparison of foreground detection results (a) original scenes (b) ground-truths (c) GMM subtraction (d) KNN subtraction (e) KDE subtraction

(a) (b) (c) (d) (e)

Fig. 9. Comparison of foreground detection results (a) original scenes (b) ground-truths (c) Fuzzy Choquet Integral subtraction (d) Fuzzy Sugeno Integral subtraction (e) Codebook subtraction

4 Conclusion

In this work, we recorded three local videos with dynamic background challenges in an attempt to prepare a local dataset. We proposed a new technique of creating a ground-truth for our local dataset by utilizing the concept of 3D modelling. Ground-truth created for each correspondent original frame to be employed for quantitative evaluation. Benchmark algorithms ***(SUBSENSE, ViBe, LOBSTER, GMM, KNN, KDE, Fuzzy Choquet Integral, Fuzzy Sugeno Integral and Codebook)*** subtraction were applied on the local dataset and both quantitative and qualitative assessments are presented. Qualitative evaluations illustrated in the screenshots figure depicting the visual assessment of the subtracted mask resulted from each algorithm. On the other hand, different evaluation metrics have been employed for the quantitative evaluation. Our results showed the efficiency of the proposed ground-truth generation technique in creating suitable input for benchmark algorithms, thus allowing the developers of practical computer vision software targeting the local environment to test their solutions on local scenes.

5 Acknowledgment

We thank Haidder Yassir (3D Producer) for his help and guidance in modeling the ground-truth scenes.

6 References

[1] D. Zeng, M. Zhu, and A. Kuijper, "Combining background subtraction algorithms with convolutional neural network," *J. Electron. Imaging*, vol. 28, no. 1, p. 13011, 2019. https://doi.org/10.1117/1.JEI.28.1.013011

[2] Y. Wang, P.-M. Jodoin, F. Porikli, J. Konrad, Y. Benezeth, and P. Ishwar, "CDnet 2014: An expanded change detection benchmark dataset," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2014, pp. 387–394. https://doi.org/10.1109/CVPRW.2014.126

[3] M. Yasir and Y. Ali, "Review on real time background extraction: models, applications, environments, challenges and evaluation approaches," *Int. J. Onl. Eng*., vol. 17, no. 2, pp. 37–68, 2021. https://doi.org/10.3991/ijoe.v17i02.18013

[4] T. Bouwmans and B. Garcia-Garcia, "Background subtraction in real applications: Challenges, current models and future directions," *arXiv Prepr. arXiv1901.03577*, 2020. https://doi.org/10.1016/j.cosrev.2019.100204

[5] S. Lee and D. Kim, "Background subtraction using the factored 3-way restricted Boltzmann machines," *arXiv Prepr. arXiv1802.01522*, 2018.

[6] P. W. Power and J. A. Schoonees, "Understanding background mixture models for foreground segmentation," in *Proceedings Image and Vision Computing New Zealand*, 2002, vol. 2002, pp. 10–11.

[7] M. Hadiuzzaman, N. Haque, F. Rahman, S. Hossain, M. R. K. Siam, and T. Z. Qiu, "Pixel-based heterogeneous traffic measurement considering shadow and illumination variation," *Signal, Image Video Process.*, vol. 11, no. 7, pp. 1245–1252, 2017. https://doi.org/10.1007/s11760-017-1081-z

[8] Z. Zivkovic and F. Van Der Heijden, "Efficient adaptive density estimation per image pixel for the task of background subtraction," *Pattern Recognit. Lett.*, vol. 27, no. 7, pp. 773–780, 2006. https://doi.org/10.1016/j.patrec.2005.11.005

[9] K. Goyal and J. Singhai, "Review of background subtraction methods using Gaussian mixture model for video surveillance systems," *Artif. Intell. Rev.*, vol. 50, no. 2, pp. 241–259, 2018. https://doi.org/10.1007/s10462-017-9542-x

[10] K. Toyama, J. Krumm, B. Brumitt, and B. Meyers, "Wallflower: Principles and practice of background maintenance," in *Proceedings of the Seventh IEEE International Conference on Computer Vision*, 1999, vol. 1, pp. 255–261. https://doi.org/10.1109/ICCV.1999.791228

[11] R. Chang, T. Gandhi, and M. M. Trivedi, "Vision modules for a multi-sensory bridge monitoring approach," in *Proceedings. The 7th International IEEE Conference on Intelligent Transportation Systems (IEEE Cat. No. 04TH8749)*, 2004, pp. 971–976.

[12] G. T. Cinar and J. C. Príncipe, "Adaptive background estimation using an information theoretic cost for hidden state estimation," in *The 2011 International Joint Conference on Neural Networks*, 2011, pp. 489–494. https://doi.org/10.1109/IJCNN.2011.6033261

[13] X. Yu, X. Chen, and H. Zhang, "Accurate motion detection in dynamic scenes based on ego-motion estimation and optical flow segmentation combined method," in *2011 Symposium on Photonics and Optoelectronics (SOPO)*, 2011, pp. 1–4. https://doi.org/10.1109/SOPO.2011.5780637

[14] F. Lei and X. Zhao, "Adaptive background estimation of underwater using Kalman-Filtering," in *2010 3rd International Congress on Image and Signal Processing*, 2010, vol. 1, pp. 64–67. https://doi.org/10.1109/CISP.2010.5647080

[15] C. Stauffer and W. E. L. Grimson, "Adaptive background mixture models for real-time tracking," in *Proceedings. 1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat. No PR00149)*, 1999, vol. 2, pp. 246–252.

[16] P.-L. St-Charles, G.-A. Bilodeau, and R. Bergevin, "Subsense: A universal change detection method with local adaptive sensitivity," *IEEE Trans. Image Process.*, vol. 24, no. 1, pp. 359–373, 2014. https://doi.org/10.1109/TIP.2014.2378053

[17] O. Barnich and M. Van Droogenbroeck, "ViBe: A powerful random technique to estimate the background in video sequences," in *2009 IEEE International Conference on Acoustics, Speech and Signal Processing*, 2009, pp. 945–948. https://doi.org/10.1109/ICASSP.2009.4959741

[18] M. Hua, Y. Li, and Y. Luo, "Robust background modeling with kernel density estimation," *Int. J. Online Eng.*, vol. 11, no. 8, 2015. https://doi.org/10.3991/ijoe.v11i8.4880

[19] F. El Baf, T. Bouwmans, and B. Vachon, "Type-2 fuzzy mixture of Gaussians model: Application to background modeling," in *International Symposium on Visual Computing*, 2008, pp. 772–781. https://doi.org/10.1007/978-3-540-89639-5_74

[20] D. Butler, S. Sridharan, and V. M. J. Bove, "Real-time adaptive background segmentation," in *2003 IEEE International Conference on Acoustics, Speech, and Signal Processing, 2003. Proceedings.(ICASSP'03)*, 2003, vol. 3, pp. III–349. https://doi.org/10.1109/ICME.2003.1221318

[21] L. Wang and C. Pan, "Effective multi-resolution background subtraction," in *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2011, pp. 909–912. https://doi.org/10.1109/ICASSP.2011.5946552

[22] M. Xiao, C. Han, and X. Kang, "A background reconstruction for dynamic scenes," in *2006 9th International Conference on Information Fusion*, 2006, pp. 1–7. https://doi.org/10.1109/ICIF.2006.301727

[23] I. Junejo, A. Bhutta, and H. Foroosh, "Dynamic scene modeling for object detection using single-class SVM," in *Proc. of IEEE International Conference on Image Processing (ICIP)*, 2010, vol. 1, pp. 1541–1544.

[24] N. Vaswani, T. Bouwmans, S. Javed, and P. Narayanamurthy, "Robust subspace learning: Robust PCA, robust subspace tracking, and robust subspace recovery," *IEEE Signal Process. Mag.*, vol. 35, no. 4, pp. 32–55, 2018. https://doi.org/10.1109/MSP.2018.2826566

[25] E. J. Candès, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis?" *J. ACM*, vol. 58, no. 3, pp. 1–37, 2011. https://doi.org/10.1145/1970392.1970395

[26] D. Farcas, C. Marghes, and T. Bouwmans, "Background subtraction via incremental maximum margin criterion: A discriminative subspace approach," *Mach. Vis. Appl.*, vol. 23, no. 6, pp. 1083–1101, 2012. https://doi.org/10.1007/s00138-012-0421-9
[27] T. Minematsu, A. Shimada, H. Uchiyama, and R. Taniguchi, "Analytics of deep neural network-based background subtraction," *J. Imaging*, vol. 4, no. 6, p. 78, 2018. https://doi.org/10.3390/jimaging4060078
[28] L. Maddalena and A. Petrosino, "Self-organizing background subtraction using color and depth data," *Multimed. Tools Appl.*, vol. 78, no. 9, pp. 11927–11948, 2019. https://doi.org/10.1007/s11042-018-6741-7
[29] M. Braham and M. Van Droogenbroeck, "Deep background subtraction with scene-specific convolutional neural networks," in *2016 International Conference on Systems, Signals and Image Processing (IWSSIP)*, 2016, pp. 1–4. https://doi.org/10.1109/IWSSIP.2016.7502717
[30] T. Bouwmans, S. Javed, M. Sultana, and S. K. Jung, "Deep neural network concepts for background subtraction: A systematic review and comparative evaluation," *Neural Networks*, 2019. https://doi.org/10.1016/j.neunet.2019.04.024
[31] A. Cioppa, M. Van Droogenbroeck, and M. Braham, "Real-time semantic background subtraction," in *2020 IEEE International Conference on Image Processing (ICIP)*, 2020, pp. 3214–3218. https://doi.org/10.1109/ICIP40778.2020.9190838
[32] "3D modeling," *Siemens Digital Industries Software*.
[33] Carmel, "How to 3D scan with a phone: Here are our best tips," *Sculpteo*. Sculpteo, 2022.
[34] "3D photography," *3D Modeling Resources*. 2019.
[35] M. Van Droogenbroeck and O. Paquot, "Background subtraction: Experiments and improvements for ViBe," in *2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, 2012, pp. 32–37. https://doi.org/10.1109/CVPRW.2012.6238924
[36] P.-L. St-Charles and G.-A. Bilodeau, "Improving background subtraction using local binary similarity patterns," in *IEEE Winter Conference on Applications of Computer Vision*, 2014, pp. 509–515. https://doi.org/10.1109/WACV.2014.6836059
[37] M. A. Yasir and Y. H. Ali, "Comparative analysis of GMM, KNN, and ViBe background subtraction algorithms applied in dynamic background scenes of video surveillance system," *Eng. Technol. J.*, vol. 40, no. 4, pp. 617–626, 2022. https://doi.org/10.30684/etj.v40i4.2154
[38] T. Bouwmans, "Background subtraction for visual surveillance: A fuzzy approach," *Handb. soft Comput. video Surveill.*, vol. 5, pp. 103–138, 2012.
[39] A. Ilyas, M. Scuturici, and S. Miguet, "Real time foreground-background segmentation using a modified codebook model," in *2009 Sixth IEEE International Conference on Advanced Video and Signal Based Surveillance*, 2009, pp. 454–459. https://doi.org/10.1109/AVSS.2009.85
[40] Z. Zivkovic, "Improved adaptive Gaussian mixture model for background subtraction," in *Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004.*, 2004, vol. 2, pp. 28–31. https://doi.org/10.1109/ICPR.2004.1333992

7 Authors

Maryam A. Yasir received her bachelor's degree in computer science from the University of Baghdad (UOB) –Iraq 2004. Since 2008, she is working as a lecturer at the computer science department, college of science, University of Baghdad up till now. In 2011 she received a certificate for a seven months course in IT Administration from the Technische Universität Berlin (TU)Berlin-Germany. In 2014 she received her master's degree in computer science from University Putra Malaysia (UPM)-Malaysia.

Meanwhile, she is a PhD candidate at the University of Technology (UOT) – Iraq. Maryam has participated in many scholar courses and activities, local and abroad including Fulbright visiting scholar at the University of Central Oklahoma (UCO)-US 2015.

Assistant Professor Dr. Yossra Hussain Ali. She received her B.Sc, M.Sc and PhD degrees in 1996, 2002 and 2006 respectively from Iraq, University of technology, department of Computer Sciences. She Joined the University of Technology, Iraq in 1997. During her postgraduate studies, she worked on Computer Network, Information systems, Agent Programming and Image Processing as well as some experience in Artificial Intelligent and Computer Data Security. She is a reviewer at many conferences and journals and she supervised a number of undergraduates and postgraduates (PhD. and MSc.) dissertations in Computer sciences. Yossra has many professional certificates and she has published in well regarded journals (e-mail: yossra.h.ali@uotechnology.edu.iq).

Article submitted 2022-07-29. Resubmitted 2022-09-23. Final acceptance 2022-09-26. Final version published as submitted by the authors.