
Performance analysis of moving object detection using BGS techniques in visual surveillance

Lavanya Sharma*

Department of Computer Science,
Uttarakhand Technical University,
Dehradun, Uttarakhand, India
Email: shm.lavanya@gmail.com

*Corresponding author

Nirvikar Lohan

Department of Computer Science,
College of Engineering Roorkee,
Roorkee, Uttarakhand, India
Email: Nirvikarlohan@yahoo.com

Abstract: Over the last decennium, the object detection is the pivotal step in any machine vision and image processing application. It is the initial step applied to extract most informative pixel from the video stream. Many algorithms are available in literature for extraction of visual information or foreground object from video sequence. This paper also provides a detailed overview of both conventional and traditional approaches used for detection of object. This paper explores various related methods, major challenges, applications, resources such as datasets, web-sources, etc. This paper presents a detailed overview of a moving object detection using background subtraction techniques in the video surveillance system that provide safety in cities, towns or home when video sequence is captured using IP cameras. The experimental work of this paper is performed over change detection, I2R, and wallflower datasets. The experimental work also depicts a comparative analysis of some of the peer methods. This work demonstrates several performance metrics to check robustness of the compared state-of-the-art methods.

Keywords: digital image processing; BGS techniques; object detection; object tracking; fuzzy logic; artificial intelligence; internet of things; smart cities; spatio-temporal data.

Reference to this paper should be made as follows: Sharma, L. and Lohan, N. (2019) 'Performance analysis of moving object detection using BGS techniques in visual surveillance', *Int. J. Spatio-Temporal Data Science*, Vol. 1, No. 1, pp.22–53.

Biographical notes: Lavanya Sharma received her PhD degree from Uttarakhand Technical University, India in 2018 and MTech in Computer Science and Engineering in 2013 from Maharshi Dayanand University, Haryana, India. She is a recipient of several prestigious awards during her academic career. She also received *TEQIP* scholarship during her PhD. She has contributed in academic research and published several papers in international conferences and international journals including IEEE, Inderscience, Elsevier, Springer and many more. She has authored a text book on object detection with

background subtraction and also contributed as Program Technical Committee member and organised Springer's ICACDS conferences 2016 and 2018 held in India. She is an editorial member/reviewer of various Inderscience and many other reputed journals. Her primary research interests are image processing and computer vision, vehicular adhoc network, mobile adhoc networks, internet of things and cloud computing.

Nirvikar Lohan received his PhD in Computer Science Engineering and presently working as Associate Professor in College of Engineering Roorkee, Uttarakhand, India. He has more than 13 years of experience in research and academics and also conducted several conferences, workshops and FDP programs. He is presently guiding five PhD students and guided more than ten MTech students. His primary research interests are image processing and computer vision, wireless networks, internet of things and big data.

1 Introduction

The motion based object detection is a pivotal and a preliminary step in many visual surveillance system and machine learning applications, as it is the initial step applied to detect independent motion of object within a particular video stream. Since, 1990's this approach has been broadly studied and applied by several authors for motion detection, object detection, tracking of people, indoor-outdoor sequence, logo detection, unusual activity detection, traffic analysis, medical analysis. In computer vision, the effectiveness usage of internet of things, GPS trajectory, and location prediction can be carried out by maintaining a connection between an IP cameras with visual output devices over an internet protocol.

Mining of motion based object data such as trajectory prediction (GPS), and location prediction is rapidly increasing day by day. Trajectory classification can be best suited to utilise the video sequences of the locations that are visited by the moving or motion based object. Location prediction also plays another important key role in predicting the historical location of motion based object, basically it aims to forecast that the object may visit. The upcoming or next location of moving object can be predicted by comparing the current object location with its previous or historical location (Barnich and Van Droogenbroeck, 2011; Shekhar et al., 2011; Zhou et al., 2014; Chen et al., 2011; Shekhar et al., 2003). Several algorithms have been designed to chunk the foreground mask from background scene from video stream and share the same stratagem (Stauffer et al., 1999; Lee, 2005; Haque et al., 2008; Bouwmans et al., 2014; Chiranjeevi and Sengupta, 2014; (Barnich and Van Droogenbroeck, 2011; Shekhar et al., 2011).

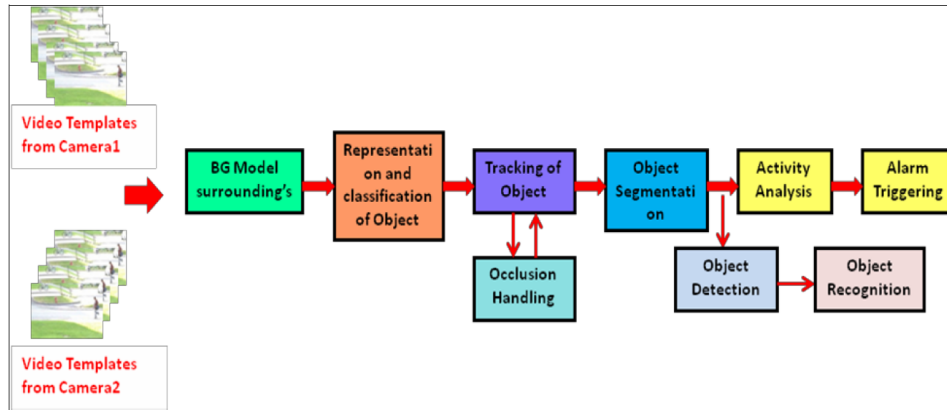
- *Background initialisation*: this scheme aims to create a background model by taking few initial frames into consideration without any moving object (Stauffer and Grimson, 1999; Akula et al., 2013; Yadav and Singh, 2015a). This mode can be designed by several methods such as fuzzy, neural, estimation, wavelet, statistical, etc.)
- *Foreground detection*: after taking few initial frames into consideration a background modelled is designed then by comparing current template from reference template lead to the foreground object of a particular scene (Bouwmans et al., 2014;

Barnich and Van Droogenbroeck, 2011; Shekhar et al., 2011; Akula et al., 2013; Yadav and Singh, 2015a; Karpatne et al., 2016).

- *Background maintenance*: in case of dynamic scenario, background model is updated periodically with respect to the learning rate to tackle various issues and challenges such as shadowing, camera jitter, slow leafy movement or water rippling in the background, camouflage etc. (Stauffer and Grimson, 1999; Lee, 2005; Akula et al., 2013; Yadav and Singh, 2015a).

Even if several algorithms have been developed in literature to track and segment the object of interest but still have some open issues that lead to the domain of digital image processing and computer vision in its infancy due to illumination or intensity change, dynamic or cluttered background, and shadowing, repetitive motion of leaves etc. (Lee, 2005; Haque et al., 2008). Therefore, to get ideal results better understanding of data and background techniques are required.

Figure 1 Basic steps of background subtraction technique (see online version for colours)



Considering all of above problematic, we present a comprehensive review of background subtraction techniques with major challenges, performance, and adaptable parameter required from external user, suitable application and applicability based comparison of all mathematical models.

In this paper, we present a brief discussion on object detection using background model. In Section 2, major contributions of researchers in this area are discussed including both conventional and traditional approaches used for detection of object. Here, we categorised all the approaches in terms of mathematical models. In Section 3, this paper explores the various applications of object detection. In Section 4, we provide a brief description of current issues and challenges. In Section 5, a brief summary of resources, datasets and codes that are available in this research field are explored. Section 6 provides experimental evaluation of various state-of-the-art methods. Finally, we discussed numerous promising directions for future perspectives.

2 State-of-the-art

Many implementations of background subtraction algorithms have been investigated till now. They have been extensively surveyed in the literature. In this article we propose a brief survey of most popular methods for background subtraction. This article helps several researchers and academicians to select the most appropriate method for a particular application.

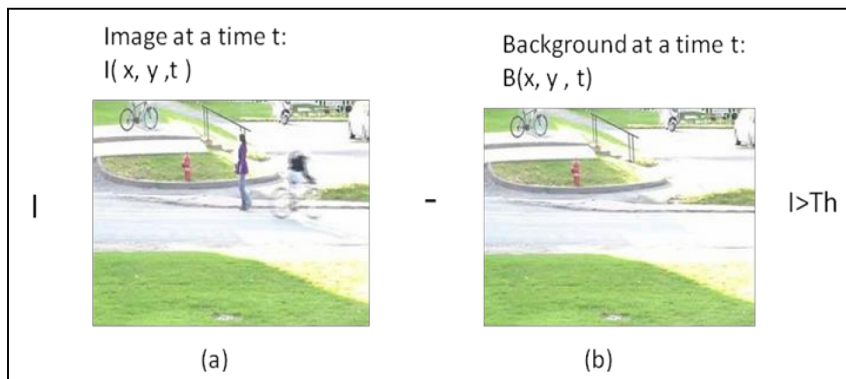
2.1 Basic background models

The basic models include creating a background model by taking few initial frames into consideration without any moving object. For each video sequence an absolute difference is computed between the reference template and current template known as frame difference. In these models background can also be modelled using mean, median, average or histogram analysis (Bouwman et al., 2014) over time. But these models can handle only some specific challenges and best suited for static background. To make them suitable for dynamic background various improvements are done by several authors.

$$BM_N(x, y) = \frac{\sum_{m=1}^N I_m(x, y)}{N} \quad (1)$$

where, $B_{MN}(x, y)$ represents pixel's (x, y) intensity of background model, $I_m(x, y)$ represents pixel's (x, y) intensity of m^{th} frame and N refers to the total number of frames used to build background model.

Figure 2 Frame differencing steps of background subtraction technique (see online version for colours)



2.2 Statistical background models

These models are more robust and realistic to dynamic background and intensity changes. Statistical variables are used to classify the foreground pixels from background pixels.

2.2.1 Gaussian model

The background is modelled using a single Gaussian or mixture of Gaussian (MoG). Pixel intensity values can be modelled either by any of above two Gaussians. Stauffer and Grimson (1999) proposed a model based on Gaussian mixture for motion detection and this work later extended by Lee (2005), Haque et al. (2008), Chiranjeevi and Sengupta (2014), Barnich and Van Droogenbroeck (2011) and Yadav and Singh (2015a) these models which comprises of background modelling and distribution of pixel's value over time. Haque et al. (2008) improved the performance of both research works Kim et al. (2005) and Guo et al. (2011) by introducing another modelling and learning mode. Jung (2009) proposed another robust BGS method and shadow removal using statistical descriptors to detect and discard the shadowing area. Barnich and Van Droogenbroeck (2011) employed particle swarm optimisation that automatically decides GMM algorithms parameters. A combined approach of GMM with region based algorithm based on colour histogram and texture information is proposed. Lee (2005) also proposed a new mechanism for object segmentation based on kernel density estimation to overcome sudden illumination change and shadow problem. Further, Yadav and Singh (2015a) have also improved the outcomes of the GMM (Guo et al., 2011) by introducing another distance estimation method and post-processing technique that outperforms well than previous work. This process results in improvement of detection quality and also performs better in colour video frames. Yadav et al. (2014) also presented an improved BGS method by enhancing the existing model by using some image processing techniques (such as dilation, erosion, imfill) in order to get better detection quality results. Yadav and Singh (2015a) proposed an efficient combined approach using Kullback Lieber divergence (KLD) and BGS for detection of a moving object in thermal video. Yadav and Singh (2015b) proposed a Quasi-Euclidian distance based threshold in background subtraction method provided by Haque et al. (2008), as given in Table 1.

Table 1 Gaussian model

<i>Author</i>	<i>Colour channel</i>	<i>Adaptable parameter requirement</i>	<i>Major challenge</i>	<i>Suitable application</i>	<i>Model representation</i>
Stauffer and Grimson (1999)	RGB	$\mu, \sigma, \alpha, \omega, T$	Sudden variation in background	Outdoor	Gaussian
Lee (2005)	RGB	$W, L, \alpha, C_k, \eta_k$	Mixed based surveillances	Indoor and outdoor	Gaussian mixture learning
Haque et al. (2008)	Gray	$\mu, \sigma, \alpha, \omega, S$	Sudden illumination variation in background	Indoor, outdoor	GMM model
Bouwmans et al. (2014)	Gray	A, T	Sudden change in background intensity	Remote surveillance system	Multimodal (gaussian)
Jung (2009)	Gray	$A, T, k, \alpha, \gamma_\alpha, \beta, S$	Shadow and noise problem, homogenous FG object	Indoor and outdoor	Robust statistical descriptor based model

Table 1 Gaussian model (continued)

<i>Author</i>	<i>Colour channel</i>	<i>Adaptable parameter requirement</i>	<i>Major challenge</i>	<i>Suitable application</i>	<i>Model representation</i>
Barnich and Van Droogenbroeck (2011)	Gray	R, N, ϕ , # K, G, N, α , β ,	Motion detection	Indoor, outdoor	ViBe
Lee (2005)	Hsv	B, w, n, γ	Sudden illumination change and shadow	Indoor and outdoor	Kernel Density Estimation
Akula et al. (2013)	IR	N, μ , σ^2	Complex BG with different env. conditions	Outdoor	Contour based statistical BGS
Yadav and Singh (2015a)	Gray	A	Noise, outlier and BG variations	Indoor and Outdoor	KLD based

Table 1 explored various statistical background model methods of the literature. The first column gives the name of authors, second column deals with the colour channel. The third column indicates different parameters required from user, and fourth column deals with major challenging issues. Fifth column represents suitable applications and the last column of the table indicates the representation of model.

2.2.2 Cluster background model

Recently, Butler et al. (2005) developed another efficient set of rules that represents each pixel as a clusters bunch. These bunches are reordered as possibility that they sculpt the background and are also redesigned to tackle complex background and illumination disparity. In this work each incoming pixels are matched against its corresponding clusters bunch and then categorised matching cluster as a part of background. Individual pixel in the frame is represented as agglomeration of K mean algorithm or code book model and all arriving pixels are matched against corresponding agglomerated group also termed as cluster's group and then classified as whether pixel belongs to background or foreground. Kim et al. (2005) developed a real time background model and used code book model to segment foreground and background pixels. This method handles the illumination changes and moving object. Guo et al. (2011) proposed another hierarchical scheme for foreground detection using codebook model to compress the information in order to get high processing speed. This paper also suggested the colour model to classify input pixel's as shadowing, background or foreground. Li et al. (2013) suggested an algorithm for foreground detection towards illumination variation using non moving cameras. Firstly, an online expectation maximisation algorithm was developed to update the Gaussian mixture and then, combined a spherical K-mean method to get accurate direction to update the model. Other authors, Varma and Sreeraj (2014) also proposed a hybrid method using code book to extract the foreground region of interest (ROI) from background scene. In this work, two code books, block based and pixel based are used for efficient object detection. Block based generates the coarse foreground region, which is then polished by pixel stage.

Table 2 described the cluster background model. In this table, the first column gives the name of authors, second column deals with the colour channel, third column indicates

different parameters required from user, fourth column deals with major challenging issues Fifth column represents suitable applications and the last column of the table indicates the representation of model.

Table 2 Cluster background model

<i>Model</i>	<i>Colour channel</i>	<i>Adaptable parameter required</i>	<i>Major challenge</i>	<i>Suitable application</i>	<i>Model representation</i>
Butler et al. (2005)	Ycbr RGB	T, W, α , L	Background variations and lighting variations.	Indoor and outdoor	Cluster based grouping of pixels
Kim et al. (2005)	Rescaled RGB	I, f, γ , p, q, α , β	Moving backgrounds or illumination variations	Indoor and outdoor	Codebook model
Guo et al. (2011)	RGB	K, α, γ , β , D, θ	Vision based applications	Indoor and outdoor	Hierarchical scheme with codebook model
Li et al. (2013)	RGB, YCRCB, HSV	α , β , T, Th, η , θ , τ	Illumination and noise	Indoor and outdoor	K means clustering algorithm
Varma and Sreeraj (2014)	RGB	T, α	Lesser memory usage and speedy processing	Dynamic environment and outdoor	Hybrid model based using codebook

2.2.3 Neural and neuro-fuzzy based network background model

The background is entitled as weighted sum of network and network is trained on N clean frames and then mean is calculated. On the basis of this process, pixels classified as a background or foreground is done. Culibrk et al. (2006) proposed another efficient neural network based method to modelled background for moving object detection in video sequences. This approach also enabled efficient and highly-parallelised hardware implementation. Bayesian classifier is used for categorisation of pixels as Background and foreground detection. Mario et al. (2015) proposed a BGS model that comprises of neural self dynamic region analysis. Maddalena and Petrosino (2012) adopted a HSV colour space based self-creating neural network for motion patterns learning to model the background. This newly developed algorithm named, self-organising background subtraction (SOBS) which detects moving objects via motion map and non moving patterns. A modernised neural network mapping approach is also used to structure framework in an easier manner. Same author Maddalena and Petrosino (2014) also proposed an improved SOBS by initiating spatial coherence into the background upgrade mechanism. This algorithm termed as SC-SOBS algorithm gives more robustness against false alarm detections. Hao et al. (2013) proposed PCA based RBF network to handle the variations caused by bit rate with limited bandwidth scenario. Rohanifar and Amiri (2011) developed an adapted backgrounds subtraction which considered radial basis function (RBF) neural network to cross between to states of machine. This system is flexible to illumination changes and few background movements' present in the background scene as shown in Table 3.

Table 3 Neural and neuro-fuzzy based network background model

<i>Model</i>	<i>Colour channel</i>	<i>Adaptable parameter required</i>	<i>Major challenge</i>	<i>Suitable application</i>	<i>Model representation</i>
Culibrk et al. (2006)	RGB	β, W, σ, θ	Real time segmentation	Indoor and outdoor	neural network based approach
Mario et al. (2015)	HSV	Th, α	Illumination and slight shadow problem	Traffic monitoring	SOM-DVA SOM-DVAM
Maddalena and Petrosino (2012, 2014)	IR	A	Robustness against false detections	Indoor and outdoor	fuzzy spatial coherence-based approach (SOBS)
Rohanifar and Amiri (2011)	Gray	Th2, Th3	Light changes and few bG object movement	Indoor and outdoor	RBF neural network
Hao et al. (2013)]	RGB	N, ϵ	Variation caused by bit rate with limited bandwidth	Traffic monitoring	PCA based RBF network

Table 3 shows various neural and neuro-fuzzy based network background model. The first column gives the name of authors, second column deals with the colour channel, third column indicates different parameters required from user, fourth column deals with major challenging issues, fifth column deals with suitable applications and the last column of the table indicates the representation of model.

2.2.4 Estimation model

Any pixel of the current image which divergates notably from its forecasted value can be treated as foreground pixel. This strainer can be a Kalman filter or Wiener filter. Tomaya et al. (1999) used Wiener filter for classification of pixel, this filter is based on present day values, pixel that divergates notably termed as foreground. Scott et al. (2009) proposed a two-step algorithm for updation of mean and standard deviation. In this method 1-D Kalman filter was used to each pixel to track the pixel intensity as discussed in Table 4. Cinar and Principe (2011) used an algorithm to extract higher order information from video sequence and proposed adaptive background estimation using theoretic cost information for hidden state estimation in linear system. Cheng and Ruan (2012) developed a background estimation algorithm using Kalman filter namely self-adaptive background model that verifies whether they can be modelled by linear system secondly, they have zero mean Gaussian distribution. There is a major drawback of Kalman algorithm that gives optimal results only when all processes are Gaussian random processes but not suitable for non-Gaussian behaviour. To overcome this challenge, Pokheriya and Pradhan (2014) proposed another efficient technique to detect and track motion of object by combining an adaptive BG subtraction with silhouette based model and Kalman filter for tracking purpose. Bhaskar et al. (2015) proposed a method that combines both dynamic reverse analysis (DRA) approach with an enhanced

Rao-Blackwellized particle filter (ERBPF) for multiple target detection and tracking of object. Liang and Juang (2015) suggested a new method to classify four kinds of moving objects in intelligent transportation system (ITS). To classify moving object, author combines static and spatio-temporal characteristics based on appearance and movement of local parts of segmented object. Álvarez-Meza et al. (2016) presented a new adaptive background model developed within an adaptive learning framework, named as object-based selective updating with correntropy (OSUC) to support video surveillance systems.

Table 4 Estimation model

<i>Model</i>	<i>Colour channel</i>	<i>Algorithm</i>	<i>Adaptable parameter required</i>
Toyama et al. (1999)	RGB	Wiener filter	a, s, p
Scott et al. (2009)	RGB	1-D Kalman filter	μ, κ
Cinar and Principe (2011)	RGB	Correntropy	Σ
Cheng et al. (2012)	Gray	Kalman filter	α_1, α_2
Pokheriya and Pradhan (2014)	Gray	Kalman filter	μ, κ
Bhaskar et al. (2015)	RGB	Rao-Blackwellized particle filter	$\alpha, \dot{\rho}$
Liang and Juang (2015)	RGB	Kalman filter	K, ω, T, Th
Álvarez-Meza et al. (2016)	RGB	Correntropy adaptation-based	α, T_e
	<i>Major challenge</i>	<i>Suitable application</i>	<i>Model representation</i>
Toyama et al. (1999)	Canonical problems	Indoor and outdoor	Wallflower
Scott et al. (2009)	Track background intensity of scene	Outdoor	Kalman filter based with no temporal coherency
Cinar and Principe (2011)	Hidden state estimation	Outdoor	information theoretic cost function
Cheng et al. (2012)	Real-time detection of moving object	Outdoor	Self-adaptive background
Pokheriya and Pradhan (2014)	Traffic surveillance	Outdoor	Silhouette based trained model
Bhaskar et al. (2015)	Moving camera occlusion and illumination	Indoor and outdoor	E-RBPF DRA
Liang and Juang (2015)	Object detection real world application	ITS system	Spatial and temporal entropy value of optical flow
Álvarez-Meza et al. (2016)	Real world scenario object detection	Indoor and outdoor	OSUC

Table 4 displays several estimation models. Here, the first column gives the name of authors, second column deals with the colour channel, third gives the details of

algorithms used by authors in their respective work, fourth column indicates different parameters required from user, fifth column deals with the performance of BG algorithm under dynamic scenario, sixth column represents suitable applications and the last column of the table indicates the representation of model.

2.2.5 Fuzzy logic based model

Some authors (Baf and Bouwmans 2007; Maddalena and Petrosino, 2010; Sivabalakrishnan and Manjula, 2010; Chiranjeevi and Sengupta, 2011, 2012) introduced concept of fuzzy logic for object detection and object tracking to measure uncertainty and imprecision such as moving leaves, camera jitter, illumination variations, etc. Baf and Bouwmans (2007) proposed two algorithms, uncertainty over mean known as T2-FMOG-UM and uncertainty over variance known as T2-FMOG-UV. Both of them proved to be more robust than crisp. Maddalena and Petrosino (2010) devised a fuzzy model that deals with decisive issues that emerges during frangible setting and a new a spatial coherence variant to magnify robustness against false alarm detections. Sivabalakrishnan and Manjula (2010) proposed an adaptive background using fuzzy logic. The author used fuzzy interference to improve the detection of moving object. Chiranjeevi and Sengupta (2011) presented an approach to handle multimodal background distribution without using multiple models per pixels. In this work author proposed a multi channel correlogram using inter pixel relation and colour planes then derived a multichannel kernel fuzzy correlogram by applying the membership transformation. Chiranjeevi and Sengupta (2012) also present a new approach basic Histon roughness index (BHRI) to overcome problems arises due to dynamic background. Kim and Kim (2012) presented a fuzzy colour histogram (FCH) based approach to handle dynamic texture scenes. Panda and Meher (2016) presented a BGS model with some important contributions. Firstly, colour difference histogram (CDH) to reduce false alarm, and secondly, a fuzzy colour difference histogram (FCDH) by using fuzzy C means (FCM). Narayan et al. (2016) proposed an effective moving cast shadow segmenting technique based on local information and fuzzy integration of Markov random field as displayed in Table 5.

Table 5 Fuzzy logic model

<i>Author's name</i>	<i>Colour channel</i>	<i>Adaptable parameter required</i>	<i>Major challenge</i>	<i>Suitable application</i>	<i>Model representation</i>
Baf and Bouwmans (2007)	RGB	μ, d, o, σ, k	Camera jitter, waving leafy movement and rippling of water	Indoor and outdoor	Type-2- fuzzy mixture of Gaussian model
Maddalena and Petrosino (2010)	HSV	p, N, Ω	Crisp moving object detection	Indoor and outdoor	Fuzzy spatial coherence-based approach
Sivabalakrishnan and Manjula (2010)	RGB	n, SAD	Difficulty in handling moved object	Indoor and outdoor	Fuzzy logic-based
Chiranjeevi and Sengupta (2011)	RGB	$\alpha, \beta, \gamma_1, \gamma_2, \rho$	Multimodal distribution	Indoor and outdoor	KFCM

Table 5 Fuzzy logic model (continued)

<i>Author's name</i>	<i>Colour channel</i>	<i>Adaptable parameter required</i>	<i>Major challenge</i>	<i>Suitable application</i>	<i>Model representation</i>
Chiranjeevi and Sengupta (2012)	RGB	α, β	Dynamic background such as swaying vegetation, camera jitter	Indoor and outdoor	Basic Histon roughness index (BHRI)
Kim and Kim (2012)	RGB	A	Dynamic texture scenes	Indoor	fuzzy colour histogram
Panda and Meher (2016)	RGB	C, r, α , E	False alarm reduction	Industrial automation security surveillance	FCDH
Narayan et al. (2016)	RGB	α, K_1, K_2	BG separation and shadow detection	Indoor and outdoor	GMRF

Table 5 represents various fuzzy logic models. The first column gives the name of authors, second column deals with the colour channel third column indicates different parameters required from user, fourth column deals with major challenging issues, fifth column represents suitable applications and the last column of the table indicates the representation of model.

2.2.6 Principal component analysis modelling approaches

In literature, various author proposed new approaches based on robust principal component analysis. Ding et al. (2011) presented a Bayesian model that infers an approximate portrait noise analysis and also infers low-rank and sparse components. Wohlberg et al. (2012) presented a local principal component analysis (LPCA) for nonlinear datasets and also used a decomposition corresponding to model comprises of low-dimensional subspaces for local imbalance. Wang and Yeung (2013) proposed a Bayesian approach to robust matrix factorisation to handle text and noise problem. Model parameters have conjugate priors in case of generative process and in other case noise model takes a Laplace mixture form. Kim et al. (2015) used PCA plane and adaptive Gaussian kernel for target segmentation in non homogenous infrared imaginary. The author first apply PCA approach minimising total least square error and then calculate second derivative of sum of square errors (SDSSE). To solve the problem of multiplex scenarios that encountered in BGS a hierarchical RPCA method is proposed with some particular assumptions regarding background by Gao et al. (2012) as shown in Table 6.

Table 6 provides various principal component analysis (PCA) model. In this table, first column gives the name of authors, second column deals with the colour channel, third column indicates different parameters required from user, fourth column deals with major challenging issues, fifth column represents suitable applications and the last column of the table indicates the representation of model.

Table 6 Principal component analysis model

<i>Author name</i>	<i>Algorithm Bayesian</i>	<i>Adaptable parameter required</i>	<i>Major challenge</i>	<i>Suitable application</i>	<i>Model representation</i>
Ding et al. (2011)	RPCA	$K, \alpha, \beta, N, \sigma$	Non-stationary noise, slow changing foreground	Indoor and outdoor	Robust principal component analysis
Wohlberg et al. (2012)	PCP	λ, S, σ, N	video background removal problem	Outdoor	Local PCP
Gao et al. (2012)	PCA	$L, M, K, \mu, \rho, \gamma$	Complex intensity due to illumination change, BG motion	Outdoor	Block sparse RPCA
Wang and Yeung (2013)	Bayesian	$\mu, \alpha, \rho, W, \nu, \beta, a, b$	Text and noise removal	Indoor and outdoor	Bayesian robust matrix factorisation
Kim et al. (2015)	PCA	$\mu, \sigma, \rho = \tau$	Extracting target region in non homogenous environment	Non-homogenous environment	PCA and adaptive Gaussian kernel

2.2.7 Subspace and low rank modelling approach

Xiong et al. (2011) implemented a robust matrix factorisation (DRMF) with presumption that a little fragment of matrix A which has been corrupted by few random outliers to evaluate true low-rank structure of this matrix and to classify the outliers. Later, Dong and DeSouza (2011) discussed a new adaptive learning algorithm using multiple eigensubspaces to tackle sudden and 21 gradual variations in background. This algorithm provides a robust clustering method for initialisation of the subspaces. Zhou and Tao (2011) developed a ‘Go Decomposition’ (GoDec) that evaluate low-rank part L and sparse part S of a matrix $X = L + S + G$ with noise case G . He et al. (2012) suggested a new algorithm namely Grassmannian robust adaptive subspace tracking algorithm (GRASTA) for tracking subspaces from highly partial information. This work used a cost function (l-norm) to non-stationary and guessimate subspaces tracking in case of streaming data vectors found to be corrupted with outliers. Furthermore, He et al. (2013) presented transformed GRASTA. This algorithm iteratively accomplished incremental gradient descent compel to the Grassmann manifold of subspaces to get simultaneously evaluation of images which are decaying into a low-rank subspace, a sparse fraction of occlusions and foreground pixels, and a transformation such as image spinning. Xu et al. (2013) proposed a Grassmannian online subspace updates with structured-sparsity (GOSUS) based on alternating direction method of multipliers (ADMM) to improve accuracy of online subspace updates listed in Table 7.

Table 7 provides several Sub-spaces and low rank model. The first column gives the name of authors, second column deals with the colour channel, third column indicates different parameters required from user, fourth column deals with major challenging issues, fifth column represents suitable applications and the last column of the table indicates the representation of model.

Table 7 Subspace and low rank model

<i>Author's name</i>	<i>Colour channel</i>	<i>Adaptable parameter required</i>	<i>Major challenge suitable</i>	<i>Suitable application</i>	<i>Model representation</i>
Xiong et al. (2011)	RGB	$K, S, L, \gamma, \sigma, \lambda, n$	Anomaly detection of tasks, robust low-rank factorisations and outliers.	Indoor and outdoor	direct robust matrix factorisation
Dong and DeSouza (2011)	RGB VGA	hreshc,f,	illumination changes	Indoor and outdoor	Eigen subspaces-based algorithm (adaptive learning of multi-subspace)
Zhou and Tao (2011)	Gray	$L, S, G, q, \Omega, c,$	Shadow and light in training images	Indoor	GoDec, bilateral random projections (BRP) based low-rank approximation
He et al. (2012)	Gray	d, L	face images as well as camera jitter	Indoor and outdoor	Grassmannian robust adaptive subspace tracking algorithm (GRASTA)
He et al. (2013)	RGB/ Gray	$\zeta_t, t, \sigma, C, K, s,$	estimates a low-rank model from noisy, corrupted data	Indoor and outdoor	GRASTA algorithm
Xu et al. (2013)	RGB	$\lambda, \rho^1, y^1, \mu_i, A$	online background subtraction and multiple face tracking	Indoor and outdoor	Grassmannian online subspace updates with structured-sparsity

2.2.8 Sparse and domain transform modelling approaches

Porikli and Wren (2005) proposed a robust algorithm referred as Wave-Back that makes a background representation on the basis of pixels history. The discrete cosine transform (DCT) coefficients are calculated for the training and testing frame. Dikmen and Huang (2008) provided a novel structure for BGS method to improved robustness with additional advantage of ease of integration of local discriminative knowledge. Cevher et al. (2008) applied the compressive sensing (CS) theory to recoup object silhouettes when they sparse in spatial domain and also provide some solutions based on convex optimisation. David et al. (2009) obtained a global trained dictionary by using a k-means classifier and matching pursuit scheme a set of coefficients is estimated. By linear combination of dictionary vectors and set of coefficients author computed background estimation for each frame to get foreground and background object. Lu et al. (2013) implemented a new online framework that use ℓ^1 sparse information fitting terms in dictionary learning to the usability and practicality of this approach. Chun et al. (2016) proposed a robust motion saliency detection technique using Temporal Fourier Transform that results in good accuracy and fast computation as compared to other conventional approaches, for fast motion saliency detection as given in Table 8.

Table 8 represents several sparse and domain transform modelling approaches. The first column gives the name of authors, second column deals with the colour channel, third column indicates different parameters required from user, fourth column deals with major challenging issues, fifth column represents suitable applications and the last column of the table indicates the representation of model.

Table 8 Sparse and domain transform model

<i>Model</i>	<i>Colour channel</i>	<i>Adaptable parameter required</i>	<i>Major challenge</i>	<i>Suitable application</i>	<i>Model representation</i>
Porikli and Wren (2005)	IR	α, M, N	Motion of plants by blowing wind, wavy motion on a beach, and the look of rotating objects.	Outdoor	Wave-Back
Dikmen and Huang (2008)	RGB	$\{-1, 0, 1\}$ filters, α	Real applications, sudden and gradual illumination changes	Outdoor	Sparse error estimation
Cevher et al. (2008)	BS images	$L, N, \beta, \sigma, 2t$	Adaptation to illumination changes	Outdoor	Compressive sensing
David et al. (2009)	Gray scale	α ,	Invariant to any illumination, shadows or other environmental changes	Indoor	Based on learned dictionary
Lu et al. (2013)	RGB	P, μ ,	Large scale and dynamic data processing in computer vision	Outdoor	Online robust dictionary learning
Chen et al. (2016)	Independent of parameters and prior knowledge	Detection of dynamic semantic areas	Dynamic semantic areas	Indoor and outdoor	Temporal Fourier transform

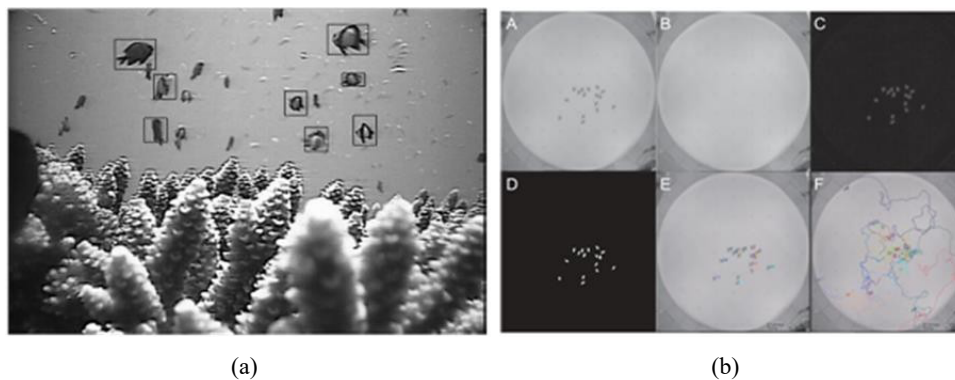
3 Application of object detection

In the domain of computer vision and image processing object detection is the initial step and is an ideal choice for video surveillance system and various computer vision application areas such as (Bouwmans et al., 2014; Baf and Bouwmans, 2017; Bouwmans, 2012):

- *Intelligent visual surveillance*: foreground object detection and background modelling is an ideal choice for this kind of system. The main aim is to detect moving object or salient object for security purpose of particular area or for statistical computation on traffic such as in airport, roads or maritime surveillances. Object of interest is quite different as compared to others. For example, surveillances can be important to judge the shopping behaviour of customer, in banking for safety purpose, battle field etc. (Chen et al., 2016; Bouwmans, 2012).

- *Optical motion capture*: this application deals with full capturing of human being from camera's and then silhouette can be extracted using BGS technique (Bouwmans et al., 2014).
- *Intelligent visual observation of forestry*: surveillance system can also be important to check the activities or behaviour of animals in restricted or protected areas such as zoo's, national park, ocean, rivers etc for ethological purpose as shown in Figure 3 (Spampinato et al., 2008; Campbell et al., 2008; Taycher et al., 2005; Himmelsbach et al., 2005).

Figure 3 Visual observation of forestry in different scenarios: (a) fish detection and tracking under water and (b) tracking of multiple honey bees with complex behaviour (see online version for colours)



Source: Spampinato et al. (2008), Campbell et al. (2008), Taycher et al. (2005) and Himmelsbach et al. (2005)

- *Human machine interaction*: in real time, several applications need interaction between machines and humans via video's captured with static cameras such as games (Kimura et al., 2012).
- *Content based video coding*: to analyse video content firstly video has to be partitioned in No. of frames to obtain the object of interest. Object is first detected and then tracked in video sequence. Then background and object are encoded separately, for this purpose an efficient and effective BGS technique is needed for object detection (Bouwmans, 2012).
- *Medical surveillance*: to provide high quality and affordable healthcare. The medical department sometime assign a person known as patient sitters to keep an eye on patient's health condition minutely at critical level or for those patients who are demented or perturbed (Kimura et al., 2012; Sharma et al., 2016; Sharma, 2016; Ma et al., 2012).
- *Biometric identification*: it means automatic identification or verifications of individuals based on their behavioural or physiological characteristics. In addition to faces and fingerprints of a person, signature, iris, tone of voice, facial thermograms, person's way of using key pad and the famous DNA (Bouwmans et al., 2014; Ma et al., 2012).

- *Aerial surveillance*: digital imaging technologies have strengthened a lot in hardware of this surveillance such as micro-aerial vehicles, forward looking infrared (IR) and high resolution imagery that are capable of identifying the objects at extremely long distances as shown in Figure 4(a) (<https://www.jenoptik.com/us-military-thermal-imager-target-acquisition>).
- *Satellite surveillance*: this system is used for terrain visualisation scenarios in both dimensions (2D or 3D). They can even penetrate the cloud cover, detects objects in buildings or in underground places, and also provide realistic video sequence having higher resolution as shown in Figure 4(b) (<http://www.satimagingcorp.com/gallery/>; <https://en.wikipedia.org/wiki/Surveillance>; <http://pursuitwire.com/2011/12/verilook-surveillance-2-0-sdk-advances-facerecognition-for-large-multi-camera-video-surveillance-systems/>).

Table 9 An overview of categorisation of visual surveillance system with some challenges

<i>Major categories</i>	<i>Sub-categories</i>	<i>Major challenges</i>	<i>Targeted objects</i>
Intelligent visual surveillances of human activities (Chiranjeevi, and Sengupta, 2014 Xu et al., 2013)	Traffic surveillance	DB, SF, IC, SH	Cars, people
	Surveillance of illegal activities	NB, DB, SF, MO	People
	Maritime traffic surveillance	SF, NB	Boats, people
	Surveillance of public event	NB, DB, MO, SF, IB	People
	Border surveillance	DB, NB, SF, IC, MO	People, vehicles
Intelligent visual observation of forestry (Spampinato et al., 2008; Campbell et al., 2008; Taycher et al., 2005; Himmelsbach et al., 2005)	Aerial surveillance	DB, SH	People, vehicles
	Birds surveillance	NB, DB, IC, MO	Birds
	Insect surveillance	NB, MO, SF,	Honeybees
	Fishery surveillance	DB, SH SF, DB	Fish
Intelligent visual observation of natural environment (Bouwman et al., 2014; Goyette et al., 2012)	River surveillances	NB	Woods
	Nuclear accident surveillance	NB, DB, SH, MO	Object, people
	Surveillance of dike	IB, NB	Object
Human machine interaction (HMI) (Bouwman et al., 2014; Kimura et al., 2012)	Coastal surveillance	NB, DB, SF	Object
	Games	DB, SF, NB, IB	People
	Ludo applications	IB, DB	People animals
	Virtual keyboard	SF, SH, MO, IC	Object, people
Content-based video Coding (Rohanifar, and Amiri, 2011; Bouwman 2012)	Gaze detection	NB, SF	People
	Video content	SF, FF, IC, SH	Video objects

Table 9 represents an overview of various kind of visual surveillance systems where first column represents major categorisation of visual surveillance system, second column gives sub categorisation view of several visual surveillance system, third column indicates several challenges of surveillances system, where descriptions are as: non stationary background (NB), dynamic background (DB), slow foreground movement (SF), moved object (MO), fast foreground movement (FO), illumination changes (IC), shadowing (SH), and background to store (IB) and fourth column represents type of object targeted.

Figure 4 Aerial and satellite surveillance system (see online version for colours)



Source: <http://www.satimagingcorp.com/gallery/>,
<https://en.wikipedia.org/wiki/Surveillance>

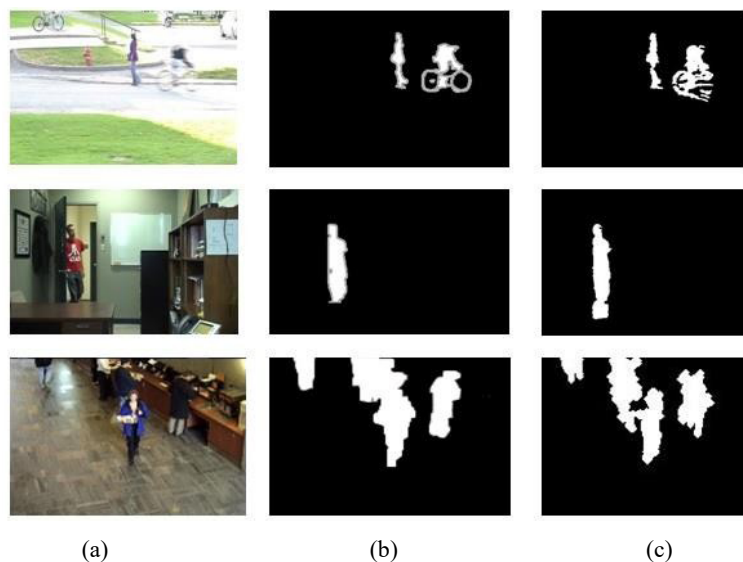
4 Challenges and issues

For efficient object detection and object tracking there are three main conditions; fixed camera, static background and constant illumination. In above ideal condition object detection using BGS gives best results. But while implementing practically, some issues or critical conditions arises and may lead to disturbance of the whole procedure. Various major challenging issues are listed below in case of visual surveillances system (Chiranjeevi and Sengupta, 2012 <https://github.com/andrewssobral/bgslibrary/wiki/Release-notes>; <https://github.com/andrewssobral/bgslibrary/wiki/List-of-available-algorithms>; <http://www.dis.uniroma1.it/~bloisi/software/imbs.html>; I2R Dataset, http://perception.i2r.a-star.edu.sg/bk_model/bkindex.html; CSIR-CSIO Moving Object Thermal Infrared Imagery Dataset (MOTI-ID): CSIR Dataset 09, <http://www.vcipl.okstate.edu/otcbvs/bench/>).

- *Noise image*: noisy image are result of poor grade picture source such as images taken from web cameras or images that are being compressed.
- *Camera jitter*: at sometime, camera is displaced due to motion of wind and causes nominal motion that result in detection of false images.
- *Illumination changes*: there can be sudden changes such as light switch on or off that can lead to illumination changes or intensity changes or in case of outdoor light intensity changes during day scenario. These sudden or gradual variations often lead to false detection of pixel.

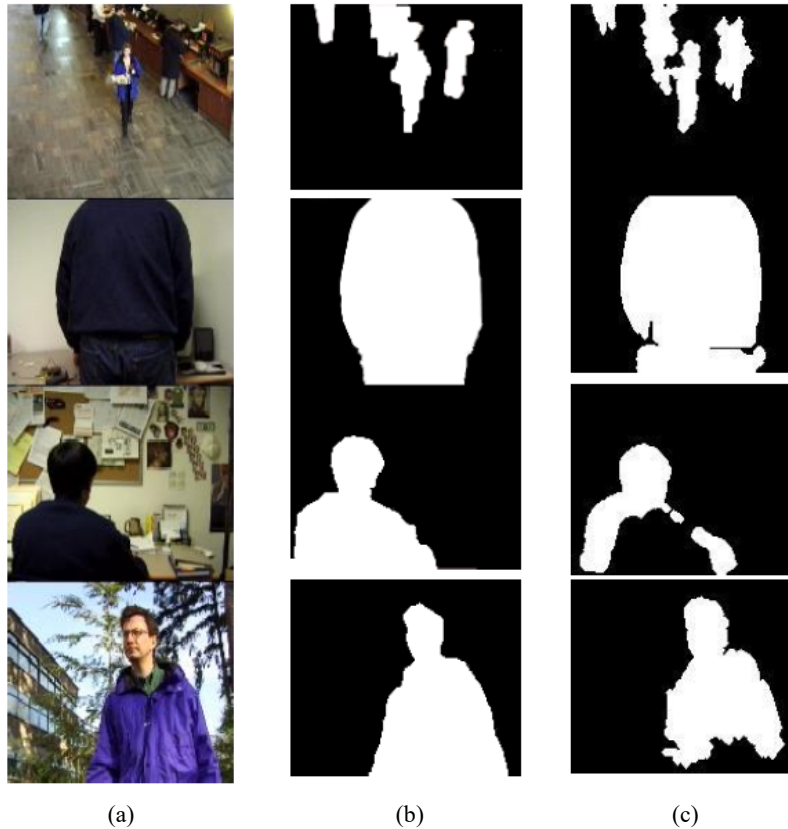
- *Bootstrapping*: during the training period in some environmental conditions background is not present. Then, it becomes very difficult to compute a background.
- *Foreground aperture*: if moved object has uniform coloured region than in that case changes occurs inside these region cannot be detected that result in false negative detection because entire object may not detected as foreground.
- *Moved background object*: sometimes the background object is moved and should not considered as a part of foreground. So, both initial and current object position are detected without efficient background maintenance mechanism.
- *Camouflage*: characteristic of foreground object are assumed on the basis of background model that often lead to miss interpretation of background and foreground pixel.
- *Inserted background object*: in case a new background object is inserted then this inserted object cannot be considered as a foreground object because this object is detected without any robust maintenance mechanism.
- *Dynamic background*: background is sometimes so cluttered it become very difficult to differentiate a pixel as a foreground or background. There are various types of background that can lead to false detection of pixel such as waving
- Leaves, rippling of water, water surface, etc.
- *Sleeping foreground object*: it become very difficult to distinguish non moving objects from background. Then they are considered as background.
- *Shadowing*: shadow detection is an active research area itself and can be detected as foreground comes from moving object or background.

Figure 5 Row wise, (a) original image (b) ground truth (c) foreground mask (see online version for colours)



In Figure 5, the first column shows frame sequence, the second column represents the generated background ground truth (GT) and third column represents the foreground mask. These frames come from Wallflower dataset (Toyama et al., 1999).

Figure 6 Row wise, (a) original image (b) ground truth (c) foreground mask (see online version for colours)



In Figure 6, the first column shows frame sequence, the second column represents the generated background ground truth (GT) and third column represents the foreground mask. These frames come from change detection dataset (Ali et al., 2013).

5 Resources, dataset and codes

- Background subtraction websites

This website contains links to the datasets and list of references. Data is updated regularly and different background models are classified as per the information provided by this website.

- BGS libraries

BGS algorithms are available in several libraries such as opencv but it consists of only algorithms. MOG (Bouwmans et al., 2014, Goyette et al., 2012; <https://www.behance.net/gallery/3943089/BGS-Library-A-Background-Subtraction-Library>) and foreground detection methods developed by Li et al. (2013). An open source multi-platform computer vision framework was developed by Laurence Bender using opencv known as scene (<https://github.com/andrewssobral/bgslibrary/wiki/Release-notes>). Sobral (<http://github.com/andrewssobral/bgslibrary/wiki/List-of-available-algorithms>) also developed a BGS Library that provide c++ framework to perform BGS and contains 29 BGS algorithm for background modelling.

- IMBS libraries

IMBS (independent multimodal background subtraction, Domenico Daniele Bloisi) library is a C++ library specifically designed for real time accurate foreground extraction. IMBS creates a multimodal model of the background in order to handle several challenging issues such as gradual or sudden illumination variations, activities of minute background rudiments, camera jitter, and change of background framework. A statistical analysis of the frames is performed for background modelling (<http://www.dis.uniroma1.it/~bloisi/software/imbs.html>). Bootstrap is also required to build the initial background model. This library exploits open CV functions.

5.1 Datasets

There are various kinds of datasets that are freely available online with their ground truth. In this research work we have several datasets that are freely available online (Rohanifar, and Amiri, 2011; I2R Dataset, 2014; CSIR-CSIO Moving Object Thermal Infrared Imagery Dataset (MOTI-ID): CSIR Dataset 09 <http://www.vcipl.okstate.edu/otcbvs/bench/>; OTCBVS Benchmark Dataset: OSU-Pedestrian, <http://www.vcipl.okstate.edu/otcbvs/bench/>; <http://www.cvg.reading.ac.uk/PETS2016/>; <http://www.openvisor.org/3dpes.asp>; <http://bmc.iut-auvergne.com/>; <https://www.houstonzoo.org/>). Some popular datasets that have been used till now during this work are listed below:

- *Wall flower*: this dataset was provided by Toyama et al. (2009). It consists of full fledge details of seven different video sequences with obstacles that are likely to encounter like illumination or intensity changes, dynamic or cluttered background. This dataset is most widely used dataset with 160×120 pixel size of image but it has major drawback that is only one ground truth available per sequence.
- *I2R8 dataset*: this dataset mainly consist of nine video sequences and was provided by Lin and Huang (I2R Dataset, 2014). Each sequence has illumination effect or dynamic background with 176×144 pixels image size. The important benefit of this dataset provides 20 ground truth images for every single frame sequence. There ground truths are captured only when critical situation occurs.

- *PETS*: this dataset contains different datasets such as PETS2001, PETS2003, PETS 2006 and related to conference performance, evaluation of tracking and surveillances. These are mainly used for object tracking evaluation as compared to BGS (<http://www.cvg.reading.ac.uk/PETS2016/>).
- *Change detection.net (CDNET)*: the dataset namely CDW (Goyette et al., 2012) comprises of realistic, large scale video sequence nearly about 90,000 frames in 31 video sequence that represent six different categories to cover vast orbit of challenging issues in two modularity's (colour and thermal IR). Each frame contains their ground truth foreground, background, and shadow region to allow precise analysis and comparison of BGS algorithm.
- *3DPeS*: 3D people Surveillance dataset is specifically designed for re identification of human beings in case of multi camera systems with non-overlapping field of view. This dataset is also applicable to several tasks such as people detection and tracking, trajectory analysis (<http://www.openvisor.org/3dpes.asp>).
- *OTCBVS dataset*: This dataset is freely or publicly available benchmark for testing and evaluating different computer vision algorithms. This benchmark contains video sequences that recorded in and beyond the visible range of spectrum (OTCBVS Benchmark Dataset: OSU-Pedestrian, <http://www.vcipl.okstate.edu/otcbvs/bench/>).
- *BMC 2012 dataset*: a workshop was organised within ACCV (Asian conference in computer vision) named, BMC mainly focused on comparison of BGS technique with both real and synthetic videos. Initially, there is set of 20 synthetic video sequences with their ground truth images for each video. Out of these 20 sequences, ten sequences are used for learning phase and remaining is used for evaluation (<http://bmc.iut-auvergne.com>).
- *UCSD background subtraction dataset*: this dataset consists of 18 video sequences with their ground truth images in 3D array format starting from frame1 of the sequence and each frame is in the format of JPEG but in some cases ground truth mask is smaller than the No. of frames in the video sequence (http://www.svcl.ucsd.edu/projects/background_subtraction/ucsdbgsub_dataset.htm).
- *CSIR-CSIO (CSIR India) dataset*: this dataset has been provided by Council for Scientific and Industrial Research (CSIR), India. This dataset is used for object detection on thermal video. This dataset contains gray scale video frames of thermal imaging (CSIR-CSIO Moving Object Thermal Infrared Imagery Dataset (MOTI-ID): CSIR Dataset 09 <http://www.vcipl.okstate.edu/otcbvs/bench/>).
- *Houston zoo dataset*: this dataset comprises of both coloured and gray scale video sequences. Main advantage of this dataset is that it comprises of training, testing frames with their ground truth video sequences (<https://www.houstonzoo.org/>). Apart from these datasets there are several datasets which are publicly available.

In this paper, we have tried to show some of the related datasets together as shown in Table 10. These are based on the real-time problematic scenarios. These sequences consist of motion in the background and illumination variation in the background scene or both.

Table 10 Various resources and datasets with major challenging issues

<i>Dataset</i>	<i>Major challenging issues</i>	<i>Authors</i>
Microsoft's Wallflower dataset	Dynamic background, illumination variation, camouflage, foreground aperture, light switch on or off, bootstrapping.	Toyama et al.(1999)
Change detection dataset	Camera jitter, cluttered background, intermittent object Motion, shadow. (http://changedetection.net/)	Goyette et al. (2012)
I2R dataset	Dynamic background or illumination variation http://perception.i2r.a-star.edu.sg/bk_model/bk_index.html)	Lin and Huang (I2R Dataset, 2014), (http://www.cvg.reading.ac.uk/PETS2016)
CSIR-CSIO dataset	Background motion, illumination variation. (http://www.vcipl.okstate.edu/otcbvs/bench/)	Council for scientific and industrial research CSIR-CSIO Moving Object Thermal Infrared Imagery Dataset (MOTI-ID): CSIR Dataset 09 (http://www.vcipl.okstate.edu/otcbvs/bench/)
OTCBVS dataset	Moving person with moving background in thermal imaging (http://www.vcipl.okstate.edu/otcbvs/bench/) (http://www.cvc.uab.es/~ivanhc/ObjDect/huertaDect.html)	Baf and Bouwmans (2007), OTCBVS Benchmark Dataset: OSU-Pedestrian (http://www.vcipl.okstate.edu/otcbvs/bench/)
PETS dataset	Background motion, illumination variation (http://ftp.pets.rdg.ac.uk/pub/)	Bouwmans et al.(2014) (http://www.cvg.reading.ac.uk/PETS2016/)

6 Experimental evaluation

In this section, the statistical measurement for system performance, combination of metrics and benchmark evaluation of various BGS algorithms are shown. Firstly the statistical measurements followed by combination metrics are presented, and then benchmark and performances are evaluated.

6.1 Statistical measurement for system performance

The standard datasets used for experimental evaluation is given below in Table 11. The major challenging issues covered in the data sequences are dynamic background (DB), slow foreground movement (SF), illumination changes (IC), shadowing (SH), non fixed or non stationary background (NB), and background to store (IB).

In Table 11, first column shows the dataset. The second column indicates different data sequence used for experimental evaluation, and third column represents various challenges of BG modelling and last column deals gives detailed overview of data

sequence contents. For qualitative analysis, few frames are demonstrated from several standard datasets with their ground truth and output results in Figure 7. The system performance of top BGS algorithms are evaluated by using several statistical metrics: precision, recall, F1-measure, fp_err, fn_err and Total error as shown in below Table 12.

Table 11 Description of datasets available online

<i>Dataset</i>	<i>Sequence</i>	<i>Challenges</i>	<i>Description</i>
Wallflower (Toyama et al., 1999)	Waving trees	NB, DB	Waving trees occupied by people
	Campus	SF, IC,IB	Moving / non moving person
I2R (Magee, 2004)	Water surface	DB, SF	Nominal motion results in false image detection
Change detection (Goyette et al., 2012)	Fountain	IB, DB	Disguise the presence of object/ person
	Overpass	SF, NB,DB, SH	Person or vehicle moving across on a scene
	Highway	NB, DB, IC	People or vehicle entering or leaving via roads
	Canoe	IC, SF, DB, SH	Object moving, vehicle moving

Table 12 Performance matrices

<i>Method</i>	<i>Description</i>
Precision	$tp / (tp + fp)$
Recall	$tp / (tp + fp)$
F-measure	$2 * (precision * recall) / (precision + recall)$
FP_error	$fp * 100 / rxc$
FN_error	$fn * 100 / rxc$, r = No. of rows and c = no. of columns
Total_Error	FP_error + FN_error
tp = No. of true positive pixels	fn_err = false positive error
fn = No. of false negative pixels	fp_err = true positive error

More, precisely, assessing difficulties of BGS methods originates from lack of statistical evaluation. Clearly, the above metrics alone cannot attain an objective comparison so we also compute several other metrics. To ensure a significant evaluation of accuracy, some other metrics – PBC, FPR, FNR, and Tot_error are derived from precision and recall.

6.2 Performance analysis of considered state-of-the-art of the literature

In literature, over past few decades various background subtraction techniques have been proposed for motion based detection of object. In this research work, we have considered some existing state-of-the-art methods for performance analysis across several quality performance metric and comparison in this paper.

6.2.1 Benchmark evaluation

In this section we compare different models of the literature. In our experiment we considered several standard datasets of real time scenarios for performance evaluation. It

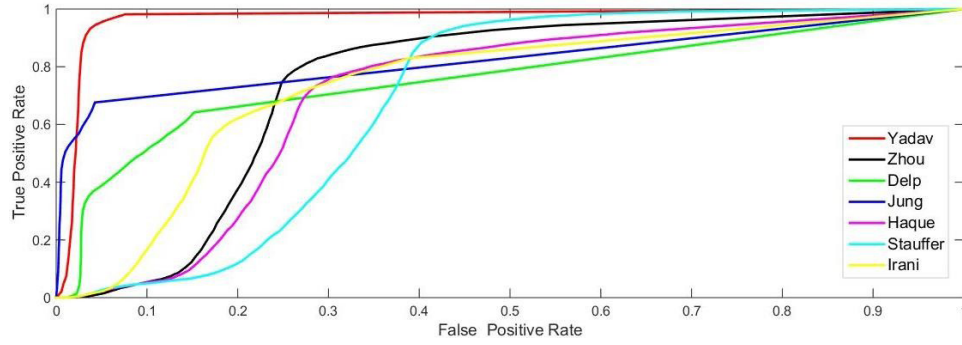
becomes necessary to evaluate system performance in terms of memory consumption, CPU efficiency and execution of time of a BGS algorithm. In this article evaluation is done on windows-XP environment, 8 GB RAM and core-2 duo processor with 2.33 GHz of speed. Experimental results have been carried out on I2R dataset, Wall flower dataset and Change detection using Matlab tool (2011b). For qualitative analysis, few frames are demonstrated from above mentioned standard datasets with their ground truth mask and output results as shown in Figure 7.

Figure 7 Qualitative results of considered method over problematic video sequences (see online version for colours)

	I2R Dataset			Wallflower Dataset	Change Detection Dataset (CD'NET Dataset)			
	Campus 372	Water Surface 499	Fountain 202	Waving Trees 247	fountain02 745	Canoe 963	overpass 2372	Highway
Orig								
GT								
Yadav								
Zhou								
Jung								
KaKi Ng								
Haque								
Stauffer								
Irani								

Rows: [1] Original Frame; [2] Ground Truth [3] Yadav [11], Zhou [88], Jung [89], KaKiNg [90], Haque [3] Stauffer [1], Irani [91].

Figure 7 shows the results of considered state-of-the-art methods with I2R Dataset, Wallflower Dataset and Change Detection Dataset. These are shown in the Figure 7 as original frame, ground truth, Yadav et al. (2014) (PIBBS), Zhou et al. (2013) (DECOLOR), Jung (2009), Ng and Delp (2011), Haque et al. (2008), Stauffer and Grimson (1999) and Irani et al. (1994).

Figure 9 Result analysis using ROC curve (see online version for colours)**Table 13** Total error analysis for Yadav (PIBBS), Zhou (DECOLOR), Ng and Delp, Jung, Haque, Stauffer, Irani

<i>Total error analysis for each model</i>					
<i>Datasets</i>	<i>Sequence</i>	<i>Yadav et al. (2014)</i>	<i>Zhou et al. (2013)</i>	<i>Ng and Delp (2011)</i>	<i>Jung (2009)</i>
I2R	Campus 372	1.1475	6.4111	2.4268	1.2842
	WaterSurface 499	1.3184	7.6123	4.2773	2.3438
	Fountain 202	0.8398	4.5313	2.7295	1.6943
Wallflower	WavingTrees 247	2.7604	10.9844	12.9271	10.7188
Change detection	fountain02 (745)	0.4123	2.6427	1.5312	1.0489
	Canoe 963	2.2461	6.9336	5.4180	5.6172
	Overpass 2398	2.6328	7.3021	14.2930	3.1302
	Highway 1171	0.6510	1.9740	1.7253	7.2188
<i>Total_Error</i>		<i>12.0083</i>	<i>48.3915</i>	<i>45.3282</i>	<i>33.0562</i>
<i>Datasets</i>	<i>Sequence</i>	<i>Haque et al. (2008)</i>	<i>Stauffer and Grimson (1999)</i>	<i>Irani et al. (1994)</i>	
I2R	Campus 372	15.2637	19.6582	7.3486	
	WaterSurface 499	5.7764	7.1973	3.3887	
	Fountain 202	5.9375	7.0117	7.2168	
Wallflower	WavingTrees 247	12.5521	14.5052	11.9479	
Change detection	fountain02 (745)	5.1617	2.1910	7.9234	
	Canoe 963	13.5990	18.7734	6.4180	
	Overpass 2398	7.9779	8.1055	12.9245	
	Highway 1171	3.4141	3.0820	4.4805	
<i>Total_Error</i>		<i>12.0083</i>	<i>69.6824</i>	<i>80.5243</i>	

Table 14 Average analysis for Yadav (PIBBS), Zhou (DECOLOR), Ng and Delp, Jung, Haque, Stauffer, Irani

<i>Model</i>	<i>Precision</i>	<i>Recall</i>	<i>F-measure</i>	<i>TPR</i>
Yadav et al. (2014)	0.8481	0.9208	0.8721	0.9208
Zhou et al. (2013)	0.4281	0.4550	0.3950	0.4550
Ng and Delp (2011)	0.6019	0.1854	0.2584	0.1854
Jung (2009)	0.7244	0.6555	0.6505	0.6555
Haque et al. (2008)	0.3485	0.7967	0.4628	0.7967
Stauffer and Grimson (1999)	0.3411	0.8280	0.4589	0.8280
Irani et al. (1994)	0.3518	0.7217	0.4310	0.7217
	<i>FPR</i>	<i>PBC</i>	<i>Accuracy</i>	<i>Total_Error</i>
Yadav et al. (2014)	0.0155	1.6316	0.9837	1.5010
Zhou et al. (2013)	0.0607	8.4927	0.9150	6.0489
Ng and Delp (2011)	0.0080	6.6162	0.9338	5.6660
Jung (2009)	0.0174	4.6136	0.9538	4.1320
Haque et al. (2008)	0.1370	13.4765	0.8652	8.7103
Stauffer and Grimson (1999)	0.1800	15.4495	0.8454	10.0655
Irani et al. (1994)	0.0892	11.1032	0.8889	7.7060

The overall experimental observations have demonstrated that the PIBBS (Yadav et al., 2014) method is able to handle motion and illumination issues of the background scene. Such kind of work is more suitable for real-time applications.

7 Conclusions and future direction

This paper provides an overview of both conventional and recent approaches used for detection of object. The main focus of this paper is the study of existing literature, application areas, challenges and comparison of performance using some state-of-the-art method. Initially, this paper covers a brief categorisation of the background models based on some mathematical tools. Secondly, this paper focused on the challenging issues, applications, resources and datasets which are publicly available for research and academic activities with ground-truth mask. In the final stage, this work has explored various state-of-the-art methods and compared them. All these methods are based on the background subtraction techniques. All the experimental work is carried out on benchmark datasets such as change detection, I2R, and wallflower. The qualitative and quantitative analysis is experimented over these problematic sequences. For quantitative analysis, this work has computed various metrics, like F-measure, precision, recall, PBC, Total_Err, TPR, FPR, accuracy, FP_Err, FN_Err etc. These methods cover a large period of time in literature. Fig. 7 provides visual results of various methods; Yadav et al. (2014) (PIBBS), Zhou et al. (2013) (DECOLOR), Jung (2009), Ng and Delp (2011), Haque et al. (2008), Stauffer and Grimson (1999) and Irani et al. (1994). This work has demonstrated a comparative study of all considered peer method in terms of error analysis, f-measure, accuracy etc., and presented the best average results. The comparison is also shown in terms of precision-recall and roc-curve. According to the analysis of all methods, the

PIBBS method Yadav et al., 2014 shows better results and demonstrated outranking performance.

In future, we would like to compare BGS techniques with other benchmarks which are publicly available. We would also plan to study and evaluate results on thermal imaging and using IoT things.

References

- Akula, A., Khanna, N., Ghosh, R., Kumar, S. Das, A. and Sardana, H.K. (2013) 'Adaptive contour based statistical background subtraction method for moving target detection in infrared video sequences', *Journal of Infrared Physics and Technology*, Elsevier, December, Vol. 63, pp.103–109.
- Ali, I., Mille, J. and Tougne, L. (2013) 'Adding a rigid motion model to foreground detection: application to moving object detection in rivers', *Pattern Analysis. Applications*, August, Vol. 17, No. 3, pp.567–585.
- Álvarez-Meza, A.M., Molina-Giraldo, S. and Castellanos-Dominguez, G. (2016) 'Background modeling using object-based selective updating and correntropy adaptation', *Image and Vision Computing*, Elsevier, January, Vol. 45, pp.22–36.
- Baf, F.E. and Bouwmans, T. (2007) 'Comparison of background subtraction methods for a multimedia learning space', *International Conference on Signal Processing and Multimedia, SIGMAP*, July.
- Barnich, O. and Van Droogenbroeck, M. (2011) 'ViBe: a universal background subtraction algorithm for video sequences', *IEEE Transactions on Image Processing*, June, Vol. 20, No. 6, pp.1709–1724.
- Bhaskar, H., Dwivedi, K., Dogra, D., Mualla, A. and Mihaylova, L. (2015) 'Autonomous detection and tracking under illumination changes, occlusions and moving camera', *Journal of Signal Processing*, Elsevier, December, Vol. 117, pp.343–354.
- Bouwmans, T. (2012) 'Background subtraction for visual surveillance: a fuzzy approach', in *Handbook on Soft Computing for Video Surveillance*, January Vol. 5, Taylor and Francis Group.
- Bouwmans, T., Porikli, F., Horferlin, B. and Vacavant, A. (2014) 'Background modeling and foreground detection for video surveillance: traditional and recent approaches, implementations, benchmarking and evaluation', *Handbook*, June, Taylor and Francis Group, CRC Press.
- Butler, D., Bove, V. and Shridharan, S. (2005) 'Real time adaptive foreground/background segmentation', *EURASIP*, pp.2292–2304.
- Campbell, J., Mummert, L. and Sukthankar, R. (2008) 'Video monitoring of honey bee colonies at the hive entrance', *ICPR Workshop on Visual Observation and Analysis of Animal and Insect Behavior, VAIB*, December.
- Cevher, V., Reddy, D., Duarte, M., Sankaranarayanan, A., Chellappa, R. and Baraniuk, R. (2008) 'Background subtraction for compressed sensing camera', *European Conference on Computer Vision, ECCV*, October.
- Chen, J., Wang, R., Liu, L. and Song, J. (2011) 'Clustering of trajectories based on Hausdorff distance', in *2011 International Conference on Electronics, Communications and Control (ICECC) (IEEE, 2011)*, pp.1940–1944.
- Chen, Z., Wang, X., Sun, Z. and Wang, Z. (2016) 'Motion saliency detection using a temporal fourier transform', *Journal of Optics and Laser Technology*, Elsevier, June, Vol. 80, pp.1–15.
- Cheng, F.C. and Ruan, S.J. (2012) 'Accurate motion detection using a self-adaptive background matching framework', *IEEE Transactions on Intelligent Transportation Systems*, Vol. 13, No. 2, pp.671–679, June.

- Chiranjeevi, P. and Sengupta, S. (2011) 'Detection of moving objects using fuzzy correlogram based background subtraction', *ICSIPA*.
- Chiranjeevi, P. and Sengupta, S. (2012) 'Robust detection of moving objects in video sequences through rough set theory framework', *Image and Vision Computing, IVC 2012*.
- Chiranjeevi, P. and Sengupta, S. (2014) 'Neighborhood supported model level fuzzy aggregation for moving object segmentation', *IEEE Trans. Image Process.*, February, Vol. 23, No. 2, pp.645–657.
- Cinar, G. and Principe, J. (2011) 'Adaptive background estimation using an information theoretic cost for hidden state estimation', *International Joint Conference on Neural Networks, IJCNN 2011*, August.
- CSIR-CSIO Moving Object Thermal Infrared Imagery Dataset (MOTI-ID): CSIR Dataset 09 [online] <http://www.vcipl.okstate.edu/otcbvs/bench/> (accessed 2 January 2015).
- Culibrk, D., Marques, O., Socek, D., Kalva, H. and Furht, B. (2006) 'A neural network approach to Bayesian background modeling for video object segmentation', in *International Conference on Computer Vision Theory and Applications, VISAPP 2006*, February.
- David, C., Gui, V. and Alexa, F. (2009) 'Foreground/background segmentation with learned dictionary', *International Conference on Circuits, Systems and Signals, CSS*, pp.197–201.
- Dikmen, M. and Huang, T. (2008) 'Robust estimation of foreground in surveillance videos by sparse error estimation', *International Conference on Pattern Recognition, ICPR*, December.
- Ding, X., He, L. and Carin, L. (2011) 'Bayesian robust principal component analysis', *IEEE Transaction on Image Processing*, May, Vol. 20, No. 12, pp.3419–3430.
- Dong, Y. and DeSouza, G. (2011) 'Adaptive learning of multi-subspace for foreground detection under illumination changes', *Computer Vision and Image Understand*, January, Vol. 115, No. 1, pp.31–49.
- Gao, Z., Cheong, L.F. and Shan, M. (2012) 'Block-Sparse RPCA for consistent foreground detection', *Computer Vision – ECCV, of the Series Lecture Notes in Computer Science*, Vol. 7576, pp.690–703.
- Goyette, N., Jodoin, P.M., Porikli, F. and Ishwar, P. (2012) 'changedetection.net: a new change detection benchmark database', *Workshop on Change Detection on CVPR, Proceedings of IEEE*, June, pp.1–8.
- Guo, J.M., Liu, Y.F. and Shih, M.H. (2011) 'Hierarchical method for foreground detection using codebook model', *IEEE Transactions on Circuits and Systems for Video Technology*, June, Vol. 21, No. 6.
- Hao, J.Y., Li, C., Kim, Z. and Xiong, Z. (2013) 'Spatio-temporal traffic scene modeling for object motion detection', *IEEE Transactions on Intelligent Transportation Systems*, Vol. 14, No. 1, pp.295–302, March.
- Haque, M., Murshed, M. and Paul, M. (2008) 'On stable dynamic background generation technique using Gaussian mixture models for robust object detection', *5th International Conference on Advanced Video and Signal Based Surveillance*, IEEE, September, pp.41–48.
- He, J., Balzano, L. and Szlam, A. (2012) 'Incremental gradient on the Grassmannian for online foreground and background separation in subsampled video', in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June.
- He, J., Zhang, D., Balzano, L. and Tao, T. (2013) 'Iterative Grassmannian optimization for robust image alignment', *Image Vis. Comput.*, October, Vol. 32, No. 10, pp.800–813.
- Himmelsbach, M., Knauer, U., Winkler, F., Zautke, F., Bienefeld, K. and Meffert, B. (2005) 'Application of an adaptive background model for monitoring honeybees', *VIIIP*.
- I2R Dataset (2014) May [online] http://perception.i2r.a-star.edu.sg/bk_model/bkindex.html (accessed 4 January 2014).
- Irani, M., Rousso, B. and Peleg, S. (1994) 'Computing occluding and transparent motions', *International Journal of Computer Vision*, Springer, February, Vol. 12, No. 1, pp.5–16.

- Jung, C.R. (2009) 'Efficient background subtraction and shadow removal for monochromatic video sequences', *IEEE Transaction on Multimedia*, February, Vol. 11, No. 3, pp.571–577.
- Karpatne, Z., Jiang, R., Vatsavai, R., Shekhar, S., and Kumar, V. (2016) 'Monitoring land-cover changes: a machine learning perspective', *IEEE Geoscience Remote Sensing Magazine*, Vol. 4, No. 2, pp.8–21.
- Kim, K., Chalidabhongse, T., Harwood, D. and Davis, L. (2005) 'Real time foreground background segmentation using codebook model', *Real Time Imaging*, Vol. 11, No. 3, pp.167–256.
- Kim, W. and Kim, C. (2012) 'Background subtraction for dynamic texture scenes using fuzzy color histograms', *IEEE Signal Processing Letters*, Vol. 3, No. 19, pp.127–130, March.
- Kim, Y.M., Park, K.T. and Moon, Y.S. (2015) 'Target segmentation in non-homogeneous infrared images using a PCA plane and an adaptive Gaussian kernel', *KSII Transactions on Internet and Information Systems*, June, Vol. 9, No. 6, pp.2302–2316.
- Kimura, T., Ohashi, M., Crailsheim, K., Schmickl, T., Odaka, R. and Ikeno, H. (2012) 'Tracking of multiple honey bees on a flat surface', in *International Conference on Emerging Trends in Engineering and Technology, ICETET 2012*, pp.36–39, November.
- Lee, D.S. (2005) 'Effective Gaussian mixture learning for video background subtraction', *IEEE Trans. Pattern Analysis Machine Intelligence*, May, Vol. 27, No. 5, pp.827–832.
- Li, D., Xu, L. and Goodman, E.D. (2013) 'Illumination-robust foreground detection in a video surveillance system', *IEEE Transactions on Circuits and Systems for Video Technology*, October, Vol. 23, No. 10, pp.1637–1650.
- Liang, C.W. and Juang, C.F. (2015) 'Moving object classification using a combination of static appearance features and spatial and temporal entropy values of optical flows', *IEEE Transactions On Intelligent Transportation Systems*, December, Vol. 16, No. 6, pp.3453–3464.
- Lu, C., Shi, J. and Jia, J. (2013) 'Online robust dictionary learning', *IEEE Computer Vision Foundation*, pp.415–422.
- Ma, K., Sun, R. and Abraham, A. (2012) 'Toward a lightweight framework for monitoring public clouds', *4th International Conference on Computational Aspects of Social Networks*, IEEE, pp.361–365, November.
- Maddalena, L. and Petrosino, A. (2010) 'A fuzzy spatial coherence-based approach to background foreground separation for moving object detection', *Neural Comput. and Application*, Vol. 19, No. 2, pp.179–186.
- Maddalena, L. and Petrosino, A. (2012) 'The SOBS algorithm: what are the limits?', *IEEE Workshop on Change Detection, CVPR 2012*, June.
- Maddalena, L. and Petrosino, A. (2014) 'The 3dSOBS+ algorithm for moving object detection', *Computer Vision and Image Understanding*, Vol. 122, pp.65–73.
- Magee, D. (2004) 'Tracking multiple vehicles using foreground, background and motion models', *Image and Vision Computing*, Vol. 22, pp.143–155.
- Mario, I., Murguia, C. and Alonso, G.R. (2015) 'Fuzzy-neural self-adapting background modeling with automatic motion analysis for dynamic object detection', *Journal of Applied Soft Computing*, November, Vol. 36, No. C, pp.570–577.
- Narayan, B., Ghosh, S., Cho, S. and Ghosh, A. (2016) 'Integration of fuzzy Markov random field and local information for separation of moving objects and shadows', *Journal of Information Sciences*, Elsevier, February, Vol. 331, No. C, pp.15–31.
- Ng, K. and Delp, E.J. (2011) 'Background subtraction using a pixel-wise adaptive learning rate for object tracking initialization', *Visual Information Processing and Communication II, Proceedings of SPIE*, January, Vol. 7882, pp.1–9.
- OTCBVS Benchmark Dataset: OSU-Pedestrian [online] <http://www.vcipl.okstate.edu/otcbvs/bench/> (accessed 4 July 2013).

- Panda, D. and Meher, S. (2016) 'Detection of moving objects using fuzzy color difference histogram based background subtraction', *IEEE Signal Processing Letters*, January, Vol. 23, No. 1. pp.45–49.
- Pokheriya, M. and Pradhan, D. (2014) 'Object detection and tracking based on silhouette based trained shape model with Kalman filter', *IEEE International Conference on Recent Advances and Innovations in Engineering (ICRAIE-2014)*, 9–11 May, Jaipur, India.
- Porikli, F. and Wren, C. (2005) 'Change detection by frequency decomposition: wave-back', in *International Workshop on Image Analysis for Multimedia Interactive Services, WIAMIS 2005*, April.
- Rohanifar, M. and Amiri, A. (2011) 'The background modeling for the video sequences using Radial Basis Function neural network', *6th International Conference on Computer Sciences and Convergence Information Technology (ICCIT)*, pp.371–374.
- Scott, J., Michael, A. and Cornish, D. (2009) 'Kalman filter-based video background estimation', *IEEE Applied Imagery Pattern Recognition Workshop (AIPRW)*, pp.1–4, 14–16 October.
- Sharma, S. (2016) 'Expanded cloud plumes hiding big data ecosystem', *Future Generation Computer Systems*, June, Vol. 59, pp.63–92.
- Sharma, S., Chang, V., Tim, U.S., Wong, J. and Gadia, S. (2016) 'Cloud based emerging services systems', *International Journal of Information Management*, Elsevier, pp.1–12, March.
- Shekhar, S., Evans, M.R., Kang, J.M. and Mohan, P. (2011) 'Identifying patterns in spatial information: a survey of method', *Wiley Interdis. Rev. Data Min. Knowl. Disc.*, Vol. 1, No. 3, pp.193–214.
- Shekhar, S., Lu, C. and Zhang, P. (2003) 'A unified approach to detecting spatial outliers', *GeoInformatica*, Vol. 7, No. 2, pp.139–166.
- Sivabalakrishnan, M. and Manjula, D. (2010) 'Adaptive background subtraction in dynamic environments using fuzzy logic', *International Journal of Image Process.*, Vol. 4, No. 1, pp.270–273.
- Spampinato, C., Burger, Y., Nadarajan, G. and Fisher, R. (2008) 'Detecting, tracking and counting fish in low quality unconstrained underwater videos', *VISAPP*, pp.514–519.
- Stauffer, C. and Grimson, W.E.L. (1999) 'Adaptive background mixture models for real-time tracking', *IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, June, pp.246–252.
- Taycher, L., Fisher, J. and Darrell, T. (2005) 'Incorporating object tracking feedback into background maintenance framework', *IEEE Workshop on Motion and Video Computing*, Vol. 2, pp.120–125.
- Toyama, K., Krumm, J., Brumitt, B. and Meyers, B. (1999) 'Wallflower: principles and practice of background maintenance', in *International Conference on Computer Vision*, pp.255–261, September.
- Varma, S. and Sreeraj, M. (2014) 'Hybrid background subtraction in video using bi-level codebook model', *5th International Conference on the Applications of Digital Information and Web Technologies (ICADIWT)*, IEEE, May, pp.124–130.
- Wang, N. and Yeung, D.Y. (2013) 'Bayesian robust matrix factorization for image and video processing', *IEEE International Conference on Computer Vision (ICCV)*, pp.785–792.
- Wohlberg, B., Chartrand, R. and Theiler, J. (2012) 'Local principal component pursuit for nonlinear datasets', *International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2012*, March.
- Xiong, L., Chen, X. and Schneider, J. (2011) 'Direct robust matrix factorization for anomaly detection', *International Conference on Data Mining, ICDM 2011*.
- Xu, J., Ithapu, V., Mukherjee, L., Rehg, J. and Singhy, V. (2013) 'GOSUS: grassmannian online subspace updates with structured sparsity', *International Conference on Computer Vision, ICCV 2013*, December.

- Yadav, D.K. and Singh, K. (2015a) 'A combined approach of Kullback-Leibler divergence method and background subtraction for moving object detection in thermal video', *Infrared Physics and Technology*, Elsevier, Vol. 76, pp.21–31, 23, December.
- Yadav, D.K. and Singh, K. (2015b) 'Moving object detection for visual surveillance using Quasi-Euclidian distance', *IC3T-2015, LNCS, Advances in Intelligent Systems and Computing Series*, Springer, September, Vol. 381, pp 225–233.
- Yadav, D.K., Sharma, L. and Bharti, S.K. (2014) 'Moving object detection in real-time visual surveillance using background subtraction technique', *14th International conference on Hybrid Intelligent Systems (HIS2014)*, IEEE, December, pp.79–84.
- Zhou, T. and Tao, D. (2011) 'GoDec: randomized low-rank and sparse matrix decomposition in noisy case', *International Conference on Machine Learning, ICML 2011*.
- Zhou, X., Shekhar, S. and Ali, R.Y. (2014) 'Spatiotemporal change footprint pattern discovery: an interdisciplinary survey', *Wiley Interdis. Rev. Data Min. Knowl. Disc.*, Vol. 4, No. 1, pp.1–23.
- Zhou, X., Yang, C. and Yu, W. (2013) 'Moving object detection by detecting contiguous outliers in the low-rank representation', *IEEE Transactions on Pattern Analysis Machine Intelligence*, March, Vol. 35, No. 3, pp.597–610.

Websites

- <http://bmc.iut-auvergne.com/>.
- <http://pursuitwire.com/2011/12/verilook-surveillance-2-0-sdk-advances-facerecognition-for-large-multi-camera-video-surveillance-systems/>.
- <http://www.cvg.reading.ac.uk/PETS2016/>.
- <http://www.dis.uniroma1.it/~bloisi/software/imbs.html>.
- <http://www.openvisor.org/3dpes.asp>.
- <http://www.satimagingcorp.com/gallery/>.
- http://www.svcl.ucsd.edu/projects/background_subtraction/ucsdbgsub_dataset.htm.
- <https://en.wikipedia.org/wiki/Surveillance>.
- <https://github.com/andrewssobral/bgslibrary/wiki/List-of-available-algorithms>.
- <https://github.com/andrewssobral/bgslibrary/wiki/Release-notes>.
- <https://www.behance.net/gallery/3943089/BGS-Library-A-Background-Subtraction- Library>.
- <https://www.houstonzoo.org/>.
- <https://www.jenoptik.com/us-military-thermal-imager-target-acquisition> (accessed 13 August 2016).