APPLICATION OF HYBRID MODELS FOR BACKGROUND SUBTRACTION IN VIDEOS

A DISSERTATION

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE AWARD OF THE DEGREE
OF

MASTER OF TECHNOLOGY
IN
SIGNAL PROCESSING AND DIGITAL DESIGN

Submitted by:

SAUMYA MAURYA

2K19/SPD/15

Under the supervision of

Dr. M.S. CHOUDHRY

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Bawana Road, Delhi-110042

JULY, 2021

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Bawana Road, Delhi-110042

CANDIDATE'S DECLARATION

I, SAUMYA MAURYA, Roll No. 2K19/SPD/15 student of MTech. (Signal Processing and Digital Design), hereby declare that the project titled **"Application of Hybrid Models for Background Subtraction in videos"** which is submitted by me to the Department of Electronics and Communication Engineering, Delhi Technological University, Delhi in partial fulfillment of the requirement for the award of the degree of Master of Technology, is original and not copied from any source without proper citation. This work has not previously formed the basis for the award of any Degree, Diploma Associateship, Fellowship or other similar title or recognition.

Place: Delhi

Date: 28/07/2021

SAUMYA MAURYA

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Bawana Road, Delhi-110042

CERTIFICATE

I hereby certify that the Project Dissertation titled "**Application of Hybrid Models for Background Subtraction in videos**" which is submitted by SAUMYA MAURYA, Roll No. 2K19/SPD/15, Electronics and Communication Department, Delhi Technological University, Delhi in partial fulfillment of the requirement for the award of the degree of Master of Technology, is a record of the work carried out by the students under my supervision. To the best of my knowledge this work has not been submitted in part or full for any Degree or Diploma to this University or elsewhere.

Place: Delhi

Dr. M.S. CHOUDHRY

Date: 28/07/2021

SUPERVISOR

ACKNOWLEDGEMENT

I would like to articulate my deep gratitude to my supervisor **DR. M.S. CHOUDHRY** (Professor, ECE), for continuous support, patience, moral support, motivation and immense knowledge during my MTech research project. Furthermore, I would like to thank him for the continuous support on the way for the completion of this project report under the guidance of professional like him. I could not have imagined having a better guide and advisor for my MTech study.

I am also grateful to **Prof. N.S. RAGHAVA**, HOD, Department of Electronics and Communication Engineering, DTU for providing the right academic resources & environment for this work to be carried out. I would also like to thanks my parents and brother for their love and support. Last but not the least, my whole-heartedly thanks to all my friends who have patiently extended all sorts of help for accomplishing this undertaking.

SAUMYA MAURYA

ABSTRACT

In recent years, Background Subtraction (BGS) has become a major topic of study in the field of Computer Vision for performing moving object detection. BGS is typically utilized in video surveillance cameras to monitor moving objects and is a widely used technique in the field of computer vision for non-stationary object identification and tracking. Hybrid models are one of the many types of approaches that can be found in the BGS literature as a result of extensive ongoing studies. Hybrid models are created by combining two or more models, allowing them to benefit from each other's strengths while overcoming the weaknesses of the original models. In this report, firstly some of the recently developed hybrid models like Hierarchical Modeling and Alternating Optimization (HMAO), randomized dynamic mode decomposition (rDMD), Adaptive Motion Estimation and Sequential Outline Separation (AME+SOS) etc. are descriptively analyzed using a tabular form of review for a clear and easy understanding in addition with the comparative analysis which is performed based on F-m values of the models for a video sequence from the very popular CDnet dataset.

Furthermore, in this project a new method based on Hybrid modelling technique that combines two models, namely Robust Principal Component Analysis (RPCA) using Principal Component Pursuit (PCP) and randomized Singular Value Decomposition (rSVD) is proposed. Video sequence is decomposed into frames and a data matrix is created with frames as a column vector and the above-mentioned methods are then applied on the data matrix to achieve the backdrop and foreground of the video.

Experiments are run on two well-known datasets, CDnet (2014) and BMC, as well as a random YouTube video sequence. In comparison to the classic PCP model, experimental results reveal that the proposed model has lower computing complexity and produces results in less time.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 Background

Object tracking and its detection are the two most significant tasks in Computer Vision finding its applications in a number of activities like Robotics, surveillance of videos, facial recognition, behavioural recognition, scene and image analysis etc. Detection of objects is performed to identify the varied contour from the background of the scene. To achieve the mentioned task a number of methods like filtering using spatio temporal features, Optical flow and Background Subtraction (BGS) can be employed. Among these methods the most common approach followed by researchers is the BGS method. To separate the background from foreground a simple three step process involving background initialization, moving object detection using intra frame difference and background maintenance is the basic idea followed by researchers incorporating different methods.

Background extraction often referred as foreground detection or background subtraction is a standard technique used in the area of Computer Vision, whose process involves separation of the background of a frame for further processing. An easy way to achieve the mentioned task involves acquiring a background image devoid of any moving object as shown in fig. 2. In a brief note steps of BGS process can be summarized in three steps- 1) Initialization of the background using frames of video sequence to construct first background image, 2) Detecting the moving objects either using foreground or background through current frame and background image contrast and 3) Regular updating of background image through background maintenance. In ideal scenario without any challenges, it is an easy task to perform BGS but problem arises in the real world when sometimes background is unavailable or constantly changing in the presence of illumination change, new object introduction etc. Thus, a superior BGS model should be more robust to such challenges over time. As mentioned, [1-2] in the literature two allied issues in BGS are change detection and salient motion detection which aims to identify intraframe variations and semantic regions to percolate unimportant ones. The fig. 1 below shows a simple block diagram of BGS process. Researchers are putting their best efforts to develop models that can provide an accurate BGS by overcoming the challenges of real-world scenarios that has led to the emergence of numerous new methodologies in the recent years.

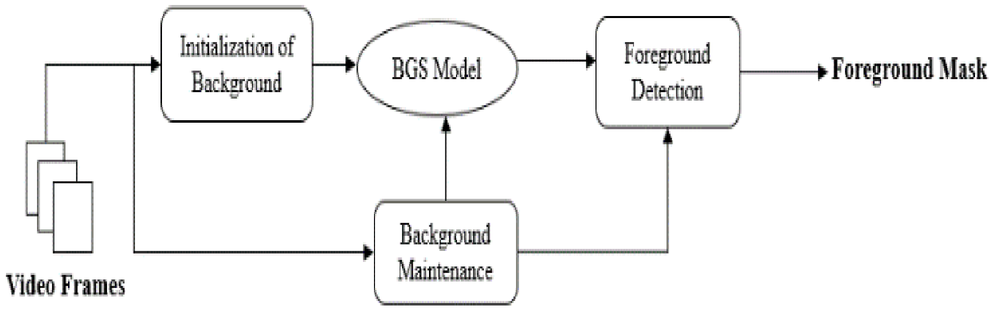

Fig. 1 Block Diagram of BGS process

In perfect conditions BGS is rather an easy task to perform but in real world scenario the presence of various challenges like intensity variation, moving background, absence of background in some frames makes the task a little challenging. There are many challenges that have been identified [9], some being simple like Dynamic Background (DB), noise, big moving object etc. and some being complex ones like camouflage, complex dynamic background, moving camera etc. Thus, a supercilious BGS model must be more immune to all these challenges. The next section provides an insight on various challenges of BGS present in today's real-world scenarios.

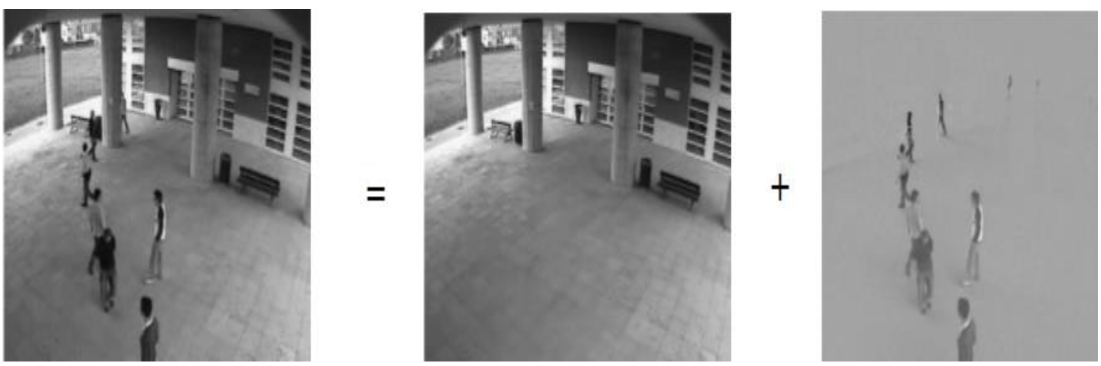

Fig. 2 Decomposition of image into Background and Foreground

1.2 Challenges and Issues in BGS

The ideal conditions needed to achieve excellent results for background subtraction includes fixed camera position, constant illumination in the surroundings and static background (no objects are removed or added in the background during camera recording), but in real world scenario the process is not so simple as certain conditions may arise time to time to cause some hinderance in the process. Bouwmans [9] nicely categorized them into 13 categories in his recent work. The extended list of these challenges including some new ones is as follows-

A. **Dynamic Background-** Background may not always contain static objects but can also be comprised of some movement leading to false detection in the BGS process. Examples of some BGs are fountains, water surfaces with ripples, waving trees due to presence of wind etc.

B. **Camera Jitter**- Movement of camera due to some external factors like heavy winds lead to the production of shifts between video frames without any motion blur. This causes the imperfect formulation of foreground masks.

C. **Camouflage-** Presence of similar type of objects in foreground and background give birth to the issue of camouflage. Similarity in pixels makes foreground and background quite impossible to distinguish.

D. **Illumination Variations-** They can be two types sudden like switching lights ON/OFF and gradual like setting of sun. Since illumination variation has effect on overall pixels it leads to a large number of false detections.

E. **Noise**- Many a times recording sources introduces some sort of noise in the recorded images or videos due to their poor quality. Presence of noise makes the task of BGS a difficult one.

F. **Ghost Effect-** Initial movement of background object leads to the detection of the background and new revealed parts of background called as ‘ghosts’ simultaneously, leading to the ghosting effect.

G. **Sleeping Foreground Object-** In some cases objects remain motionless for a long period of time leading to the misclassification of these as background.

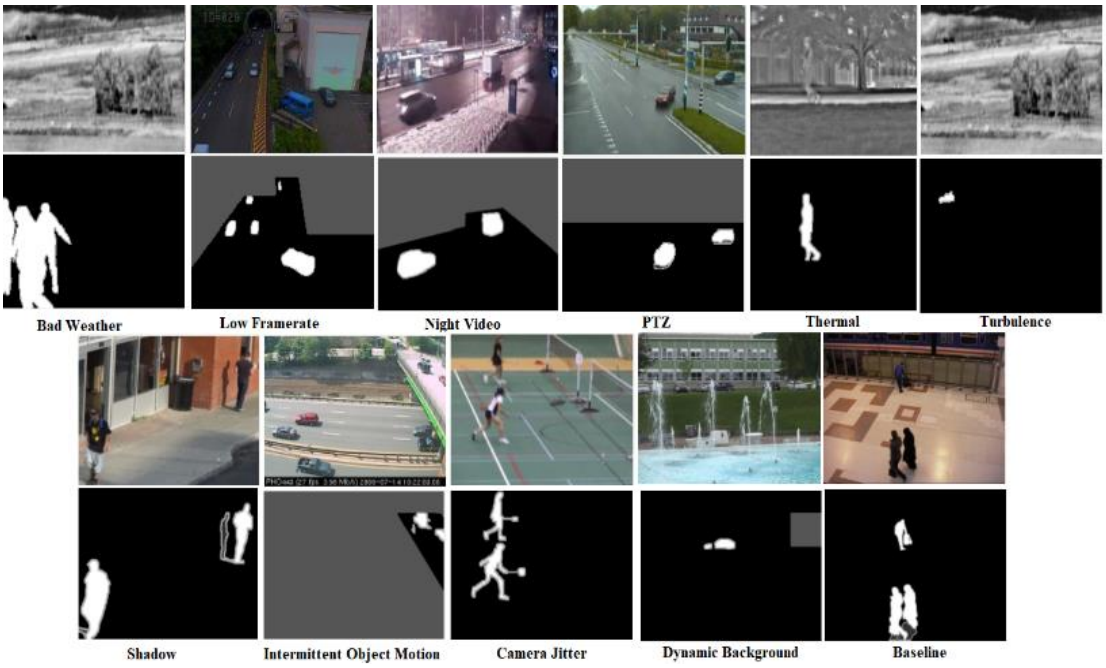

Fig. 3 Image frames and their ground truths of some real-world video sequences representing different challenges from CDnet (2014) dataset

H. **Inserted Background Object**- Sometimes a new background is inserted in the frame at some point of time, but care should be taken not to include that in foreground. E.g. - Introduction of sun in the frame at dusk when camera position is towards west.

I. **Moved Background Object -** In certain cases, the background objects can also move which leads to false detection of foreground.

J. **Foreground Aperture -** Presence of uniform-coloured regions in the moved objects results in non- detection of these regions as foreground, leading to hole effect in the foreground mask.

K. **Bootstrapping -** This challenge arises in the absence of background in some environments that makes the task to compute a representative background image impossible.

L. **Shadows -** Existence of shadows in the foreground leads to the false detections of background making the task of BGS more challenging.

M. **Automatic Camera Adjustment -** With the advancement in the technology, new smart cameras are being invented that can adjust their focus, gain control, white balance and brightness. These properties change the inter- frame colour level dynamics leading to inaccurate BGS. E.g.- PTZ camera.

N. **Big Moving Objects in Foreground -** Often big objects in the foreground moving very slowly leads to false detections in the foreground mask, mistaken these objects as background.

O. **Clutter Environment -** Cluttering refers to the presence of a lot of unorganized stuff in the environment making it a difficult task to distinguish all things accurately from the background.

1.3 Motivation

Hybrid Modelling based circuits can be easily found in the literature of background separation models where a lot of models are inspired from this modelling technique to present a robust BGS model. These models help the researchers in overcoming the obstacles and limitations of the traditional models and thus tackling the various challenges like dynamic background, moving camera, camouflage, complex background etc. as mentioned in above section. Various advantages of Hybrid Models that makes them so efficient are as follows:

1) These are developed as a result of the combination of various traditional models, thus works more efficiently by overcoming the limitations of the models used to form these.

2) They can be useful while dealing with some complex challenges in the field of BGS by combining methods that can support the limitations of each other, as a result giving us a robust method.

3) A lot of time is saved while working on these models, as they are more effective and less complex to work on, only requirement being the exhaustive research for the combination of techniques.

1.4 Objective

The main objective of this report is to describe the proposed Hybrid Modelling based technique for background subtraction. The proposed model combines two well known traditional models namely Principal Component Pursuit (PCP) and randomised SVD (rSVD) to carry out the task of foreground separation from the given video sequence. The proposed technique effectively carries out the mentioned task in less time and comparatively low computational complexity as compared to the traditional PCP model using RPCA. Resolution of the data matrix formed can be easily controlled thus providing a factor to control execution time of the model as time is directly proportional to the resolution of the data matrix.

1.5 Organisation of Report

This report has been divided into five chapters. In Chapter 1, the topic of the report is introduced followed by the current challenges looming in the BGS field. This chapter also gives insight on the motivation and objective behind the report. Chapter 2 brings out the comprehensive literature review of background separation models including full length surveys and recent models falling under various categories. Chapter 3 deals with the various hybrid models emerged in recent years with their mathematical and theoretical description supported by their comparative analysis. In Chapter 4, the proposed model is explained using the developed algorithm in terms of background modelling and foreground separation. Lastly, Chapter 5 presents the experimental analysis of the model by performing the experiments on some well-known datasets, followed by the conclusion and future scope of the proposed work.

CHAPTER 2

LITERATURE REVIEW

2.1 Full Length Surveys

Time to time some researchers contribute their time in bringing out comparative reviews and surveys in different fields. In the field of BGS subtraction also there are quite a few surveys done in past years some being compact and others being detailed. The survey done by Ivor [3] can be said to be first in the field comparison been done on 9 techniques. Piccardi [4] compared some methods based on speed, accuracy and memory requirements in 2004. First categorized survey was done by Cheung and Kamath [5] dividing models into recursive and non-recursive category, thus providing more organized data. Cristani et al. [6] end up classifying the techniques singular monocular sensor or multi sensors but with limitation of ambiguity in the categories of some methods. First detailed survey was conducted by Bouwmans et al. [7] in 2010 and extended version of it came into light in 2013, but again the survey dealt only with statistical category of models. RPCA based [8] survey was conducted in 2013 in which all RPCA models in the field were surveyed. In 2014, a comprehensive survey was again provided by Bouwmans [9] which can be said to be the most detailed survey done including all models, challenges, datasets available till 2014. Low rank-based models [10] were surveyed in 2016 and Neural network-based models [11] in 2019 which thoroughly surveyed all models of low rank and neural network category and presented us a compact survey but again limiting the category of models. Latest in the literature is a detailed survey of background subtraction in real applications [12] which contains a detailed description of BGS models in real applications with their scope and limitations as well including the current challenges in real time applications.

A detailed survey conducted on models developed in recent years (2016-2021) including the latest ones as well is provided in the section below.

2.2 Recent Models in the Field of BGS

There are a number of models developed in the recent years that can be categorized in various categories such as Robust subspace, Statistical, fuzzy, Neural Networks etc. These category models from recent years (2016-2021) are reviewed as below-

2.2.1 Robust Subspace Models

The models belonging to this category separate the background and foreground using a robust subspace model constructed on low rank and sparse decomposition. To achieve this decomposition different techniques namely Robust Principal Component Analysis, Robust Non-Negative Matrix Factorization or Robust Orthonormal Subspace Learning can be used. LRSD based models have attracted a great deal of interest in the recent years specially the RPCA based models. A quick review of the recent LRSD based models can be done as below-

A. RPCA Models- Boumans [9] presented a detailed study on the usage of RPCA in various emerging fields. It is stated that RPCA via decomposition into low rank plus sparse matrices proffers a sturdy outline for a wide variety of applications including image processing, video processing, computer vision etc. Introduction of statistical PCA goes back to 1901, courtesy of Karl Pearson, and from then it has found its usage in quite a few applications, but disadvantages like outlier sensitivity, high computational cost and memory requirements made it infamous. Candes et al. [13] proposed the RPCA model by decomposing the data matrix as a low rank and sparse component. The matrix under observation can be modelled as- M=L+S, where L represents the low rank subspace and S the sparse subspace. Recently several techniques based on the RPCA model have been introduced in the literature of foreground – background separation based on different approaches integrated with RPCA.

Feng *et al.* [14] proposed an Online RPCA (ORPCA) model to increase the computation time and storage efficiency of RPCA by processing one sample per time instance. Guang *et al.* [15] modified this ORPCA algorithm by adjusting it to the diversification of sparsity of foreground objects among frames through introduction of a prior information based adaptive weighing parameter in the algorithm, as a result, increasing the accuracy of foreground detection. The weighing parameter controls the level of care to be given to the sparsity of the estimated frames and it can be adjusted in accordance with the foreground pixels. The model provides promising results in the videos suffering from the problem of camera jitter. The drawback of the model includes the slow processing and less accurate results for large video frames. To solve the problem of background separation in videos with complex background Xue *et al.* [16] proposed the method in which more refined prior representation for static and dynamic component of video sequence is employed. The model describes the well-defined low rank property of background layer by focusing on the exploitation of rank-1 constraint, and spatial temporal smoothness of foreground layer is calculated using 3-D total variation measure and L1 norm. The inclusion of rank-1, smooth and sparse properties in RPCA framework gives better results with increased accuracy in complex backgrounds.

To deal with the problem of high computational time and cost in methods involving Singular Value Decomposition (SVD) in large matrices Wang *et al.* [17] proposed a method that enhances the robustness and flexibility of the traditional nuclear norm by embracing two non-convex low rank approaches. These two non-convex low rank approaches give an upper bound to the true rank in the method. Model is solved using Augmented Lagrangian Multiplier (ALM). In the model proposed in [18], a challenging problem of moving camera with possible dense and sparse corruptions is addressed. The video under consideration is patterned as an integration of low rank component that encapsulates the background, a smooth component encapsulating the dynamic foreground and a sparse component that segregates possible outliers and other sparse corruptions in the video. OptShrink technique is used to separate out the low rank portion. TV regularization technique is used to separate the dynamic background from sparse corruptions and this proposed model outsources other models working on same technique by reconstructing full field of view of the scene by using frame registration and masking strategy.

In another approach [19], it is assumed that image segmentation consists of low rank background matrix and a dynamic tree structured sparse matrix. To obtrude the spatial coherence on the foreground regions, the method dynamically calculates the aid of the foreground regions via super pixel generation step. Background modelling is achieved using Column Subset Selection problem to tackle the curse of dimensionality. Model provides accurate results handling the problem of foreground aperture, large dynamic background and bootstrapping challenge. Cao *et al.* [20] proposed a model for Background Subtraction from Compressive Measurements (BSCM) in which background subtraction is achieved by breaking the video frames into backgrounds and foregrounds using its Spatio-temporal correlation features and temporal continuity features in a tensor framework, respectively. They enhanced both features used for subtraction by using 3D Total variation for foreground features and Tucker Decomposition for Background features. Since the proposed technique need not to gather information about all video voxels, it significantly reduces the computational and storage costs.

Graph regularized RPCA model known as a motion aware regularization of graphs on low rank component for background modelling is proposed by Javed *et al.* [21]. Authors calculated the optical flow and took advantage of it to create a motion matrix for object detection. The model provided promising results in the field of camera jitter and sudden lighting variation. Chen *et al.* [22] used motion saliency feature to come up with a two-stage model. Here the sparse foreground blocks are calculated based on the spectral residuals in first stage and the spatial correlation of foreground objects in second stage while RPCA is used to solve low rank and sparse matrix decomposition in the first stage. Srivasta *et al.* [23] proposed to use optical flow method with RPCA to estimate motions between frames to obtain more accurate foreground masks.

B. Other LRSD Based Models - The advancement in LRSD based models helped in overcoming many challenges of background foreground segmentation in recent years. Falling in this category Zheng *et al.* [26] came up with the idea to detect moving objects as sparse outliers in contrast with low rank structure background to tackle the large moving objects in the foreground. As authors retain global appearance consistency for both foreground and background accuracy of the model against large moving objects came to be pretty high. SOFT-INPUT algorithm and Markov Random field with Graph cut algorithm were used to optimize background matrix and foreground mask respectively. Athisha *et al.* [27] improved the traditional Dynamic Mode Decomposition (DMD) algorithm by decreasing its computational complexity. Authors converted the video stream into frames, broke each frame into segments and then applied DMD on the segment where moving object or foreground is obtained. The improved method drastically reduced the computational complexity of the DMD based background-foreground separation. Presence of complex scenes is a very common but challenging task in video segmentation and Javed *et al.* [28] showed that it can be handled easily by introducing spatio-temporal information in the low rank component of the frames.

Table 1 Overview of Robust Subspace models

Category	Models	Algorithm	Issues addressed	Dataset Used	Author (Year)
RPCA Models	-Modified Online RPCA	ORPCA	Camera Jitter	CDnet	Guang *et al.* (2017) [15]
	-Total Variation and Rank-1 constraint RPCA	TR1-RPCA	Big and slowly moving objects, multi objects	SABS, CDnet, I2R, UCSD	Xue *et al.* (2018) [16]
	-Robust PCA Using Matrix Factorization	OurLaplace, OurGeman	Computational Time, Deadlines	Self-made	Wang e*t al.* (2018) [17]
	-Panaromic RPCA	PRPCA, OptShrink	Gaussian noise, pepper outliers	I2R	Moore e*t al.* (2019) [18]
	-Tree Structured Sparse RPCA	ARPCA	Bootstrapping challenge	Wallflower, I2R, CDnet (2012), SABS	Ebadi *et al.* (2018) [19]
	-Total Variation Regularised Tensor PCA	H-TenRPCA PG-TenRPCA	Dynamic Background	SABS, CAVIAR, I2R, UCSD, CDnet	Cao *et al.* (2016) [20]
	-N fast principal	nFast Principal	Dynamic BG	Wallflower, I2R	Chen e*t al.* (2019) [22]
	-Sparse RPCA	Multi component	Complex Dynamic Background	Wallflower	Wu *et al.* (2018) [25]
	-Graph Regularised RPCA	SRPCA MAG-RPCA	Intermittent motion of foreground objects	SBMnet	Javed e*t al.* (2016), Javed *et al*. (2017) [21]
	-Robust Online Matrix Factorization	OMoGMF, t-OMoGMF	Camera Jitter, Dynamic background	I2R, Self-made dataset	Yong *et al.* (2018) [24]
	-RPCA via Optical Flow	CORPCA-OF	Complex Foreground	CDnet, SBMnet	Srivasta *at el.* (2018)
Other LRSD Models	LRS separation	CLASS, F-CLASS	Large Moving Object	CDnet (2014), Wallflower	Zheng *et al.* (2017) [26]
	$L_{1/2}$ Norm and Spatial Continuity	SCLR-$l_{1/2}$	Dynamic Background	CDnet (2014), UCSD	Zhu *et al.* (2018) [31]
	Robust Matrix Completion	fRMC	Sudden Illumination change	BMC (2012), SABS	Behnaz *et al.* (2018) [30]
	Generalised Nuclear norm and Laplacian Mixture	GNNLSM	Baseline, Shadow Effect	CDnet (2014)	Yang *et al.* (2020) [34]
	GNN and Structured Sparse norm	GNNSSN, RGNNSSN	Noise	I2R, CDnet (2012)	Yang *et al.* (2020) [35]
	Multi Scale structured	MSSLRSD	Baseline, Intermittent Object Detection	CDnet, GTFD	Zheng *et al.* (2019) [33]
	LR and structured sparse	LRSSD	Illumination change noise	CDnet, BMC, PETS2001	Ma *et al.* (2018) [29]
	Spatiotemporal Low rank modelling	SLMC, SRPCA	Complex Dynamic Background, Bootstrapping	CDnet (2014), I2R, Wallflower, SBMI	Javed *et al.* (2018) [28]
	Structured l_p regularised	ULLR, SLLR	Slow Moving Foreground Objects	I2R, CDnet (2012)	Chen *et al.* (2019) [32]
	l_1 minimization algorithm	l_1 min algorithm	Dynamic Background, Noise	CDnet	Xiao *et al.* (2016) [36]
	Dynamic Mode Decomposition	DMD	Cluster of foreground objects	SBI	Athisha *et al.* (2018) [27]

Assumption that images resting in the background are linearly correlated was taken into account. Also, to remove the redundant data, the help of motion segmented binary matrix generated using Optical Flow information was taken. Authors created two graphs between frames for encoding of local structure for uninterrupted promotion of intrinsic behaviour of low rank model in opposition to the outliers. In the same year Ma *et al.* [29] came up with a low rank and Scale Invariant Local Ternary pattern (SILTP) based structured sparse decomposition method to give robust results against illumination change and noise in video sequences. Accurate results were provided due to the fact that model considers both structural properties of images and spatial distribution of outliers during BGS. SILTP enhanced the structured presentation of foreground region and refined foreground is presented jointly by the fusion of sparse components and attention mapping. Behnaz *et al.* [30] proposed an unsupervised approach for moving objects with non-salient features detection. For increasing efficiency of the model, spatial information is also incorporated in the model and proposed algorithm was used to solve the formulated background model.

Most of the LRSD methods use L1 norm in low rank modelling and this approach has several limitations while dealing with statistically variable underlying foreground pixels in background, so to handle this weakness Zhu *et al.* [31] proposed a unified regularization framework. Flexibility in choosing the reasonable sparse domain is provided by adjusting a shrinkage parameter q and model is solved using Augmented Lagrange multiplier (ALM) method. Introduction of $l_{1/2}$ constraint and spatial continuity regularisation term helped in enhancing the accuracy of the model. In the same field Chen *et al.* [32] incorporated flexible l_p regularised schemes to handle the variation scenario. Local group structure along with homogeneous group structure were used to encode the underlying objects. The value of p between 0 and 1 guaranteed the method to achieve robustness against outliers. In continuation of the category Zheng *et al.* [33] explored the structured smoothness using both appearance consistency and spatial compactness in LRSD framework. Authors fused structured information at various scales into formulation of robustness and computational efficiency is increased by incorporating a low rank design scheme in the proposed model. Although model performed well video sequences with noise, the accuracy was not up to the mark with tiny moving objects in foreground during dark environment.

Furthermore, Yang *et al.* [34] recently proposed a Generalized Nuclear Norm infusion with Laplacian Scale Mix to approximate low rank and sparse matrix and the proposed problem statement is solved using Alternating Direction Method of Multipliers (ADMM). The model performs well in presence of baseline effect. Latest in the literature is the Generalized Nuclear Norm and structured Sparse norm (GNNSNN) method by Yang *et al.* [35] that overcomes the limitation of low accuracy of the surrogate functions of rank and sparsity in LRSD methods. Here, rank function is approximated by GNN and l_o norm by SNN.

2.2.2 Fuzzy Models

Background subtraction models often suffers from critical issues like imprecisions and uncertainties and to tackle these issues researchers in recent years incorporated the Fuzzy concepts in the background subtraction models, thus increasing the accuracy and efficiency of background subtraction. The classification of these models can be done as follows:

A. Fuzzy Background Modelling- To improve the precision of dynamic background modelling Zheng *et al.* [37] introduced us with a background subtraction model using histograms drew on strong uniform fuzzy partitions falling under the category of pixel-based techniques. These histograms evince excellent performance in terms of accuracy and can converge to real PDF with fewer samples, thus providing accurate results outperforming other state of art methods. Authors used a saturating linear function to provide misclassification of pixels and threshold for background segmentation is set adaptively based on the shape of histogram. Moving forward, to detect dynamic background accurately irrespective of the size of foreground object, Yu *et al.* [38] presented us with a model based on Fuzzy C means clustering and Fuzzy Nearness Degree. Background model is developed using Fuzzy Histogram, accumulating pixel values in each histogram interval with different probabilities and pixels are categorized into background and foreground based on fuzzy nearness of them. Efficiency of the algorithm is increased due to the fact that parallel computing is used by the authors in the method.

B. Fuzzy Foreground Modelling – A method that uses adjustable Gaussian Mixture Model to model the state of each pixel as an intermediate step with fuzzy classifiers to predict the pixel class of each frame is put forward by Ali *et al.* [39]. Here both selection and classification are purely fuzzy based and method falls into the category of uncertainty methods using intervals based on Type-2 Fuzzy sets.

Table 2 Overview of Fuzzy models

Models	**Algorithm used**	**Issues addressed**	**Dataset Used**	**Authors (Year)**
Fuzzy Background Modelling	Histograms based on Fuzzy Partitions	Dynamic Background and Camouflaged foreground.	CDnet, BMC2012, Fish4Knowledge, MAR, Wallflower	Zheng *et al.* (2017) [37]
	FCFN	Inaccuracy and uncertainty in Dynamic Background	CDnet, SABS	Yu *et al.* (2019) [38]
Fuzzy Foreground Detection		Dynamic Background	CDnet	Ali *et al.* (2018) [39]

2.2.3 Cluster Models

In cluster models each pixel in the frame of the video sequence can be temporally constituted by clusters. In the traditional codebook method proposed by Kim *et al.* [40], formulation of a codebook based on the codewords was done and based on that samples at each pixel are clustered into a set of codeword depending upon the colour distortion metric and finally background encoding is done using pixel by pixel basis. To make the method more robust to challenges like illumination change, dynamic background and shadow effect, Krungkaew and Kusakunniran [41] came up with an advanced version of traditional codebook method by integrating a technique of making a dynamic codeword in the model. To increase the robustness of the model to the above mention challenges authors also used a lab colour space feature, whose key advantage is that the illumination component is completely cut off from the colour components. The model provides promising results by outperforming previous codebook-based methods.

Furthermore, to cope up with more advanced challenges like bootstrapping and dynamic background Qin *et al.* [42] formulated a regression-based foreground detection method including a novel background basis selection process. Authors included basis matrix construction and basis matrix update processes to improve the performance of the method. Global foreground and background are disjoint using linear regression-based foreground detection method along with locally constructed basis matrix. The results are more accurate as compared with other methods as shown in the paper.

Table 3 Overview of Cluster Models

Models	Algorithm used	Issues addressed	Dataset Used	Authors (Year)
Novel Dynamic Codebook	DCB	Illumination change, Clutter environment	CDnet	Krungkaew *et al.* (2016) [41]
Background Basis Selection	Regression based Foreground Detection & Background basis selection process	Frequent appearing foreground, Bootstrapping	CDnet, I2R	Qin *et al.* (2016) [42]

Table 4 Overview of Neural Network Models

Models	Issues addressed	Dataset Used	Authors (Year)
Robust Multi Independent Layered Self - Balanced SOBS(RMIL-SB-SOBS)	Illumination change, Camera motion, Jitter effect	CDnet (2012, 2014)	Gemignani and Rosa (2016) [43]
Dual Multi scale 3D fully convolutional network (DMFC3D)	Night video, Thermal, Turbulence	CDnet (2014)	Wang *et al.* (2018) [44]
Deep Probabilistic Background Model (DeepPBM)	Complex Backgrounds	BMC (2012)	Farnoosh *et al.* (2019) [45]
Convolutional Neural Network Features	Dynamic Background	CDnet	Yu *et al.* (2019) [46]
Deep Context Prediction	Blurry objects, Intermittent motion detection, Illumination challenge, Weather	ImageNet (Training), SBM.net (FG estimation, CDnet (BG estimation)	Sultana *et al.* (2019) [49]
Foreground Segmentation Technique using Deep Neural Networks	Illumination change, Camera Jitter, Shadows	CDnet (2014)	Vijayan and Mohan (2020) [47]

2.2.4 Neural Network Models

The models coming under this category classifies each pixel of the image frame as either moving object or background and here Foreground Segmentation is considered to be a pixel level classification problem. In recent years many Neural Network based methods have been presented for solving the background subtraction problem and they have been proven pretty successful in solving the issue in a better way in terms of computational complexity and handling various challenges of video sequences termed earlier in the paper.

A multi layered artificial NN was presented by Gemignani and Rosa [43] which being capable of building a multi modal colour distribution for each pixel provided efficient results handling the issues of camera motion, camera jitter and illumination change. Individual layer of the network represents an independent scene of background scene and computation of each layer can be done independently. Authors used two resolution levels- pixel and image to perform motion classification over time thus increasing the accuracy of motion segmentation in the model. Wang *et al.* [44] came up with the idea of dual multi scale 3D fully CNN for foreground detection hinderance. The model takes help of a decoder -encoder structure to obtain pixel wise classification results by formulating a mapping from image sequence. Encoder and Decoder are separately trained using two stage training procedure to achieve better results. For foreground segmentation, spatial temporal features from image sequence is taken into account. Although model performs well with night video and thermal images, it fails to provide robust results for issues like camera jitter and intermittent object motion.

To improve accuracy of NN models for complex backgrounds Farnoosh *et al.* [45] used Deep NN in learning compressed representation of video sequences and reformulate them back to original domain. The important property of the proposed unsupervised and generalized model is its generative modelling ability for background that can help in building synthetic background of specific scene. The model performs the task of background subtraction by thresholding the difference between a generative low dimensional model of the background and the original input frame. For more effective handling of dynamic background Yu *et al.* [46] proposed a refined Background subtraction method based on features learned from a Deep CNN, the nature of which being unsupervised and concise. In the model, general features of a video are taken out from the lower convolutional layers and since only few channels containing crucial information of the video are applied to background subtraction, the computational complexity of the method reduces drastically. The run time is also less as model need not to be trained unlike other supervised learning models.

Another NN based model is proposed by Vijayan and Mohan [47] based on handcrafted features incorporating both high- and low-level features effectively. Authors used both motion and appearance information for more accurate moving object detection. Background Pixel library is created using PAWCS [48] algorithm and model is trained with the help of optical flow information, spatial information and background model information. The method provides promising results for challenges like illumination change, camera jitter etc. outperforming several art of state methods. An amalgamation of General Adversarial network and image impainting is presented by Sultana *et al.* [49] to achieve better accuracy while predicting homogeneous or blurry objects. The method is a context prediction based and presents a Estimation of missing pixel values for blurry objects is done by training CNN via impainting method.

2.2.5 Statistical Models

A. Gaussian Models- The Gaussian Mixture Model (GMM) is a sturdy method for background modelling where dynamic backgrounds are under questioning. Although GMM works well with background variations like multi- modal and moderate illumination changes, it often fails to provide accurate results with instantaneous illumination change and irregular background motions. In recent years a variety of methods based on GMM have been developed to cope up with the above-mentioned limitations. Shi *et al.* [50] proposed a model to handle the challenge of well-characterizing the varying sparsity of foreground components, by modelling the sparse components with Gaussian scale mixture. Authors represented individual pixel as a multiplication of standard Gaussian random variable and a positive scaling variable. A stochastic optimization strategy for low rank approximation and super-pixel-based segmentation method is incorporated to bring down the computational complexity and enhance the foreground extraction accuracy of the model.

The issue of setting an adaptive threshold in multi modal Gaussian based BG-FG separation is handled by Akilan *et al.* [51] by proposing a novel foreground enhancement strategy accustoming the colour and illumination measures. Histogram developed based on the fused features of both the measures namely colour and illumination is used to find the optimal threshold for foreground extraction and primary FG feature is enhanced by this feature fusion approach. The proposed model outperforms various GMM based models and provides promising results by dealing with various issues of BG modelling effectively but fails to tackle the issues of shadow regions and sudden moving objects. Zhang *et al.* [52] used a fixed length vector known as Mino Vector (MV), that keeps purified pixel values for each pixel in background scene to define the Dynamic nature (DN) of the pixels. The dynamicity of the area in the scene is defined by the Dynamic Rank (DR) of the pixels such as large values of DN represents pixels belonging to dynamic area and so on. The threshold of the pixel is decided based on the DR and Kernel Density Estimator (KDE) is implemented in further steps to improve the estimation accuracy. At the final stage authors introduced Tetris Update Scheme for background subtraction. Method performs well with scenes having issue of camera jitter but needs improvement for scene with appearance changes.

B. Subspace Learning Models - A PCA based approach to determine the low rank subspace of background scene straight away in compressed sensed (CS) domain is used by Liu and Pados [53] to decrease the computational complexity of traditional PCA models. Principal components are computed under L1 norm maximization criteria and background scene is obtained by taking projection of CS measurement vector into L1 principal components followed by Total Variation minimization for image recovery. To further increase the accuracy of PCA and decrease its computational cost Huang *et al.* [54] came up with the idea of Regression PCA, where regression portion of images is assumed to be the background information and foreground detection can be done by feeding the residual portion of images to the RPCA model.

Although model provides promising results with big moving objects but have limitations when background changes a lot in a short time span. Toroghi *et al.* [55] casted the concept of Hilbert-Schmidt Independence Criteria (HSIC) integrated with Independent Component Analysis (ICA) to deal with the challenge of sudden illumination change. Morphological operator helped in enhancing the efficiency of the model by removing holes and noisy part of foreground mask. A method that disintegrates video tensor spatiotemporally into background and foreground is proposed by Kajo *et al.* [56] to accurately detect the static Foreground Objects (SFOs). SFOs can be detected at both pixel and object level using SVD. Although the model is capable of determining the exact time when the foreground object changes its behaviour from being in motion to static, it fails to deliver accurate results when objects with sharp pixel intensities are detected. To further increase the accuracy of foreground detection Pei *et al.* [57] proposed a dictionary learning based method using dynamic dictionary updating modelling and active contour-based modelling for background and foreground respectively. Spatiotemporal Clustering helps in estimating the detailed background model by increasing the temporal quantisation.

C. Support Vector Models - SV based models become advantageous when the discussion of computational complexity comes into picture. Guo *et al.* [58] proposed an adaptive background subtraction model utilizing a novel local SVD Binary Pattern (LSBP) feature that can define potential structure of local regions in an image, in return enhancing the robustness of the model to illumination variations, noise and shadows. Model provides better results than many states of art methods because both LSBP binary string and colour intensities are used to describe a single background pixel although it couldn't outperform SuBSENSE model [59]. Another model based on SV thresholding is proposed by Yang *et al.* [60] that minimises all singular values and make usage of a novel non- convex surrogate function to approximate the low rank matrix. Authors introduced a generalised SVT operator (GSVT) operator which is solved using Alternating Direction Method of Multipliers (ADMM) to formulate a non-convex low rank and sparse decomposition of the scene. The model effectively handles challenges like bootstrapping and noise. An interesting approach of dealing videos as a tensor and decomposing each horizontal/lateral slice into low rank matrices is proposed by Kajo *et al.* [61]. A depth analysis is done to compute left/right singular values and singular vectors for extracting SVD components and rank1 of these first left/ right singular values and vectors forms an effective model of background scene. Model outperforms many known matrix / tensor-based methods in terms of performance and sensitivity to the variation in video contents.

D. Non- Parametric Models - Models falling under this category works on the non-parametric background modelling paradigm. He *et al.* [62] modified the famous ViBe algorithm by merging colour features in texture features with it for a more effective background subtraction technique. For extracting local binary gray scale difference information, a novel local compact binary count (LCBC) is introduced and fusion of this LCBC and colour features in ViBE increased the accuracy of the model significantly. Total Variation norm regularisation deals with enhancing the spatial smoothness of foreground objects. Proposed method performs well with dynamic background but its accuracy got reduced on introduction of noise in image frame. Another approach in this category was developed by Zhong *et al.* [63] to increase the discriminative ability of background model by representing each pixel using fusion of its gray values and gradient. The method relies on dual target updating strategy to update the classification status of the pixel in both foreground and background targets. Authors exploited similar and random schemes for updating background to enhance the accuracy of updated target. Furthermore, for diverse scenarios to dynamically adjust the controlling threshold a controlling threshold adaption process is developed.

E. Hybrid Models - To overcome the problem of inflated computational time to process the high-resolution videos, Ericson and Donovan [64] modeled an integrated data driven method combining PCA with time series analysis (Fourier Transform in time). The proposed technique handled the dynamic background issue in video processing using a randomized matrix algorithm for fast computing the low rank dynamic mode decomposition. DMD is an interesting topic of research for applications where focus is mainly on fast processing rather than precision, since it permits the matrix decomposition representing both space and time. The model proposed by Li *et al.* [65] provide a solution to this precision problem in dynamic background issue of background subtraction by putting forward a joint esteem problem called as Hierarchical Modeling and Alternating optimization (HMAO). The formulated problem is solved using Alternating Directions Multiplier Method (ADMM). The model disintegrates the background temporally into low and high frequency components for the enhance characterization of the class of video with dynamic background and facilitate the improved noise resistant at higher resolutions for foreground by constructing a prior Markov Random Field (MRF) at a spatially low resolution. To make the model more robust, the authors treat the principal component of residuals as the detail background not as outliers. Although the proposed model has minor issue with micro foreground regions, it provides some promising results in the field of illumination change incorporated in dynamic background.

Wang *et al.* [66] dealt with same problem of skepticizing dynamic background as a moving object by fusion of super-pixel segmentation technique into RPCA model. Authors obtained the super-pixel grouping matrix using Linear Spectral Clustering (LSC) algorithm as it produces super pixels with high boundary dependence and then incorporated the improved RPCA model on that to procure the sparse matrix. Finally, motion mask based on the sparse matrix is created with the help of Otsu algorithm which in turn along with grouping matrix accurately locates the motion area from the dynamic background. The accuracy of the proposed method is high for large moving objects but it fails to correctly classify the small moving objects. Apart from the challenge of dynamic background, illumination change during video recording is another hinderance for background subtraction models. To address this issue in outdoor video surveillance in [67] a model was proposed that benefits from two popular approaches namely local adaptive, which tends to fit a distribution pixelwise and global linear, which rebuilds the background based on the low rank version of the scene. The model can predict the whole scene using a small subset of pixel values exploiting the pixel correlations, which are independent of any assumptions about the type of illumination sources, thus making the model effective against illumination change challenge. Experimental results showed that the background can be constructed as early as after 5th frame resulting in a much less computational time. In [68], authors used the mixture of two approaches namely Adaptive Motion Estimation (AME) and Sequential Outline Separation (SOS) to propose an effective object detection technique which efficiently handles issues like closer view gap, sudden light change, and ghosting. Movement region of images is extracted using difference of image sequence between consecutive frames.

An additive model combining Neighbourhood Chain Prediction (NCP) and Differential boundary Pattern (DBP) is proposed by Mohanpriya and Mahesh [69] to handle the issue of shadow region in visual tracking. Clustering of the lumpy background is done using NCP followed by extraction of the texture of the video frame using DBP to vanquish the shadow pixels in existing in the image frame. Machine Learning Classification (MLC) is used for binary labelling of tracking region to separate the background and foreground. Authors exploited the lower intensity estimation and prediction of the area around it to enhance the pixels for shadow removal resulting in an adequate amount of accuracy of the proposed model. Another model dealing with the same issue of moving cast shadows [70] is proposed by Badri *et al.* The median of pixel values at corresponding temporal direction pixel locations is used to construct the background model. The binary image separating the background and the moving object with shadows is obtained using Maximum o’posterior probability (MAP) of the Markov Random Field (MRF) constrained clustering. Authors used the analysis of rg colour chrominance property of shadow to accurately differentiate the shadows from the moving objects. An interesting fusion model dealing with common issues like bootstrapping and illumination change along with complex challenges such as zoom in and zoom out can be found in [71].

The model is a fusion of neural network and clustering algorithms that is based on neural network mapping, relating neurons with each pixel of the scene and representing them with weight vectors. Foreground and background area is recognized and established by spatio temporal tracking of moving key points, followed by neural background subtraction localized in these areas to accomplish foreground detection. Latest in the literature the hybrid model proposed by Umer *et al.* [72] is an unsupervised

Table 5 Overview of Statistical Models

Category	Methods		Issues addressed	Dataset Used	Authors (Year)
Gaussian Models	MVGMM		Camera Jitter, Dynamic background, Illumination change	CDnet	Akilan *et al.* (2018) [51]
	TUS+MV		Camera Jitter	CDnet (2012)	Zhang *et al.* (2018) [52]
	SGSM-BS		Dynamic Background, Irregular moving Objects	PITS, CDnet	Shi *et al.* (2018) [50]
Subspace Learning Models	**MATRICES**	CS-L1-PCA	Illumination change, Bootstrapping, Dynamic Background	CDnet, Airport	Liu & Pados (2016) [53]
		HSICPSO-ICA	Sudden luminance variation,	Wallflower	Toroghi *et al.* (2020) [54]
		RegPCA	Dynamic background, Big Moving Objects	Airport and game video	Huang *et al.* (2020) [55]
	TENSORS	Online SFO	Camera Jitter, Intermittent Object detection, Dynamic background, SFO	CDnet	Kajo *et al.* (2019) [56]
		THHS	Large local noise interference, Dynamic background	CDnet (2014), I2R	Pei *et al.* (2020) [57]
Support Vector Models	LSBP		Illumination change, Shadows	CDnet (2012)	Guo *et al.* (2016) [58]
	Non-LSRD		Baseline effect, Noise in frames	CDnet, I2R	Yang *et al.* (2019) [60]
	SS-SVD		Cluttering, Moving Camera	SBI, SBMC	Kajo *et al.* (2018) [61]
Non-Parametric Models	LcbcBN		Dynamic Background, Noise	I2R, CDnet	He *et al.* (2019) [62]
	DTNBM		Dynamic & Static foreground object	SABS, CDnet	Zhong *et al.* (2019) [63]
Hybrid Models	GMRF+MAP+localGLCM		Shadows	CDnet, BMC	Badri *et al.* (2016) [70]
	rDMD		Illumination change	CDnet, BMC	Erichson *et al.* (2016) [64]
	NCP-DBP		Shadow Region, Illumination change, Dynamic Background	CAVIAR	Mohanpriya *et al.* (2017) [69]
	HMAO		Dynamic and Complex backgrounds with illumination change.	CDnet, I2R	Li *et al.* (2019) [65]
	KC-SONN-ABM		Bootstrapping, Illumination change, Zoom-in and zoom-out	FBMS, Airport MotionSeg, SBMnet	Avola *et al.* (2020) [71]
	DSPB		Repetitive interference with illumination change	LIMU, I2R, Wallflower	Tocker *et al.* (2020) [67]
	AME+SOS		Closer view gap, Ghosting, Sudden lighting change	CDnet	Thenmozhi *et al.* (2020) [68]
	LSCNC-RPCA		Dynamic background	CDnet (2014), I2R	Wang *et al.* (2020) [65]
	FPPVOS		Computational Cost Large moving objects	YouTube - Objects	Umer *et al.* (2021) [72]
	cDMD (FT+SVD)		Illumination change	CDnet, BMC	Erichson *et al.* (2019) [73]

learning-based model fusing Probability Weighted Moments (PWM) with Principal Component analysis (PCA) to decrease the computational cost issue of unsupervised learning methods. Algorithm pays attention mainly to the features having intra class variability, assuming variation between foreground object of interest and its background in terms of size, colour, shape etc. PWM is used to separate the foreground object and background while PCA is used to reduce the dimensionality of data by neglecting irrelevant data.

2.2.6 OTHER MODELS

Some other categories of models have also emerged in the recent years to deal with the various challenges by effectively performing the application of foreground and background separation. The models falling under those categories are as below-

A. Low Rank Minimization Model - In the year 2017 Zhao and Tan [74] proposed unified and systematic framework to carry out the online non-negative matrix factorization in existence of outliers. Two solvers, established on projected gradient descendent and alternation direction method of multipliers, were put forward by the authors in the method. The algorithm was set to learn the basics matrix in an online manner while remaining unaffected by outliers. The model performed fairly well on large scale datasets containing shadow problems.

Zhu *et al.* [75] came up with a model integrating structural and spatiotemporal information into a combined sparse low rank matrix and factorization framework. Model achieved the robustness to noise due to the inclusion of structured sparsity constraint and Total Variation Regularization. In addition to that continuous trajectory of long duration motionless objects in spatiotemporal space made the model immune towards statistical variability in the underlying foreground pixels. The limitation of the model included the wrong absorption of partial shadow region of moving object into the foreground. Latest in the literature is the model with moderate computational complexity scalable to temporal and spatial resolution of videos [76]. Low Rank Matrix Factorization was used in place of batch mode nuclear norm for low rank approximation and moving objects were separated from background with the assistance of information extracted by optical flow. The model is enhanced with an affine transformation to improve the accuracy against presence of camera jitter in video sequence.

B. Unsupervised and Supervised Learning Models - In [77] an unsupervised learning-based method using a visual attention mechanism to model a full background based on subsets of frames and propagating it further to other frames for accurate background subtraction was proposed. To stand against frequent local motions in the scenes a SIFT matching based stabilization algorithm was created. Authors identified an object as a part of background or foreground based on Region of Difference (RoD) identified between frames. Further in the category Chelly *eta l.* [78] came up with a 2D unsupervised approach to handle issue of moving camera. During test time wrapping of previously unseen frame based on prediction was performed and foreground/ background separation is done based on the projection of these wrappings.

C. Spatio-Temporal Feature Based Models - The models taking both space and time features into consideration falls into this category. Ray and Chakraborty [80] suggested that moving objects can be detected as clusters of spatiotemporal blobs created as a result of image sequence analysis by 3D Gabor filter. Assumption that rate with which the foreground changes as compared with the background would always be greater, led to the idea of the proposed model to overcome moving camera limitation in BGS applications. The main advantage is that the model need not to be trained or initialized. In addition to this model Mohanty and Rup [81] proposed the latest model that can overcome quite a few challenges of video sequence in BGS, making it an advanced model. A spatio temporal local Binary Pattern (STLBP) technique is used to take out the duet of spatial texture and temporal motion features from a video sequence. Mean of surrounding pixels is taken into account to find these STLBP features. Model works well with many issues like illumination variation, dynamic background etc. but needs advancement for camouflaged and complex scenes.

D. Super-Pixel Based Approaches - Super pixels are those pixels that share some common characteristics and models based on these pixels falls under this category. Silva *et al.* [82] proposed a novel online one class ensemble-based method to select suitable features of each region of a certain scene to separate the foreground from the background. Incorporation of super pixel approach increases the accuracy of the proposed model in terms of computational time and efficiency. Furthermore, the model presented in [83] calculated super pixels of initial frame using linear iterative clustering method and then employed K means clustering to divide them into k smaller units. To initialize BG model the divided sub super pixels are represented as a multi-dimensional feature vector.

Refinement and simple representation of super pixel method gives the proposed model an edge in accuracy and computational time.

E. Sparse Subspace Models - A model that merges spatial and temporal sparse subspace clustering into RPCA is presented by Javed *et al.* [84]. The information captured from the formed spatial and temporal graphs was used to estimate the proximity matrices based on normalized Euclidean and geodesic distances. These matrices bounded the BG model to be spatially and temporally consistent. Linearized Alternating Direction Method with adaptive Penalty Optimization scheme (LAD MAP) provided the solution to the proposed model. Another model in the category based on Dictionary Learning for Dynamic Mode Decomposition was presented in [85]. Here the moving objects were extracted by separating sequence of video frames into BG and FG information with the help of dictionary learned patches. The model was successful in outperforming the traditional DMD model by providing promising results with complex background scenes.

F. Region Based Models - Shahida and Maheshwari used an interesting approach [86] by performing foreground detection only in a certain region of interest (ROI) which is calculated by the proposed Reducing and Raising by Half (RRH) algorithm. Change in pixel intensity between frames gives the foreground object and the adjustments for ROI is monitored by taken into account the factors like the movement of objects in the background, instable region based on periodic changes and similarity between FG and BG under consideration. The drawbacks of the model include inefficiency in removing shadow and baseline effects.

G. Misc Models - Some other models were also developed in these past years such as Estimation model like FSBE [87] that detects the illumination change in video sequence using value channel from HSV colour space and detect the BG by selection of frames using refinement algorithm, Domain Transfer Model like FWFC [88] that uses wavelet bands to highlight small variations in image domain and thus formulating BG/FG models from these bands, Z-score model [89] where background is modelled based upon the sample differences of temporal sequence obtained using z-score labelling , Twin Background model [91] presenting dual long and short models to enhance foreground exposure rate incorporating statistical method and Motion Histogram Threshold Algorithm [92] analysing smooth histogram peaks and valleys of the inter frame motion, thus presenting the threshold value at run time for accurate background subtraction.

Table 6 Overview of Other Category Models

Category	Methods		Issues addressed	Dataset Used	Authors (Year)
Low Rank Minimization Models	OPGD, OADMM		Large Scale Dataset, shadow Detection	I2R	Zhao & Tan (2017) [74]
	MSMF		Dynamic Background, Irregular moving Objects in foreground	CDnet, I2R, BMC	Zhu *et al.* (2020) [75]
	SSMR		Complex moving objects, Camera Jitter, Large scale dataset	CDnet (2014), SBI	Yang *et al.* (2020) [76]
Learning Models	**Unsupervised**	Visual attention +SIFT	Foreground objects with frequent and salient motion	DARPA, CDnet	Lin *et al.* (2017) [77]
		JA-POLS	Moving Camera	CDnet (2014)	Chelly *et al.* (2021) [78]
	Supervised	CM+DM	Camouflaged Background	CDnet (2014), Some sequences from internet	Xiang *et al.* (2020) [79]
Spatio-Temporal Models	ODTU		Moving Camera	TB50, TB100	Ray *et al.* (2019) [80]
	Proposed STLBP2		Illumination variation, dynamic background, low contrast, noise	CDnet, I2R, SBMnet, Wallflower, OTCBVS	Mohanty & Rup (2021) [81]
Super Pixel Based	IWOC-SVM		Intermittent Object Detection	MSVS, RGB-D, CDnet (2014)	Silva *et al.* (2017) [82]
	SBS		Tiny foreground objects, Ghost artifacts	CDnet (2014), SBI, CMTIS	Chen *et al.* (2020) [83]
Sparse Subspace Models	MSCL, MSCL-FL		Photometric Variations, Camera Jitter, Shadows, Large occlusions	CDnet (2014), BMC, SBMnet, I2R, Wallflower	Javed *et al.* (2017) [84]
	dl-DMD		Background with complex underlying dynamics.	SBMnet, BMC	Huq *et al.* (2020) [85]
Region Based Models	RRH		Intermittent Object Detection, Illumination change	CDnet (2014)	Shahida *et al.* (2020) [86]
Other models	Estimation Model – FSBE		Illumination change, Cluttering	SBMnet (2016)	Djerida *et al.* (2019) [87]
	Domain Transfer Model – FWFC		Camouflage	SBM-RGBD, Self-made (CAMO-UOW)	Li *et al.* (2018) [88]
	Z- Score Model – CBGM		Shadow Effect, Illumination change	Wallflower, I2R	Suman *et al.* (2017) [89]
	Twin Background Model – TBGM		Dynamic Background	CDnet (2014)	Jeeva *et al.* (2019) [90]
	Motion Histogram Model		Baseline	CDnet	Fakhri *et al.* (2020) [91]

CHAPTER 3

BGS USING HYBRID MODELLING BASED TECHNIQUES

3.1 Integration of fuzzy Markov Random Field and local information (GMRF+MAP+localGLCM)

In the year of 2016 Badri et al. [70] proposed a praiseworthy technique that detaches the in-motion objects from their casted shadows on the background. The proposed method has two main steps- subtraction of background and detection of shadows. To perform first step the background model is built with the help of median calculated using the temporal direction pixel values at a desired pixel position. Model can be mathematically described as follows-

$$Z_{(t-1)}(x,y) = median\left\{y_k(x,y), \qquad k = 0,1,2 \ldots, (t-1)\right\} \tag{3.1}$$

Authors took the advantage of RGB features including ten more local features with it at individual pixel position in both modeled reference frame and target frame to diminish the effects of colour frequency changes and frequent illumination variation. Once done with the features difference image can be obtained between reference and target frames as follows-

$$d_j^f(x,y) = |p_t^f(x,y) - Z_{(t-1)}^f(x,y)| \tag{3.2}$$

To further find the non-stationary region in the scene MRF based fuzzy clustering is performed followed by maximum a'posterior probability (MAP) to divide the in-motion object with shows and backgrounds in different groups. Usage of shadow's rg colour chrominance property further separation of cast shadows is done from the moving objects. The mathematical equation for shadow detection is given as-

$$r_{(t+1)}(x,y) = \begin{cases} 1; & if\Delta(x,y) \geq k_2 * \mu_\Delta \\ 0; & otherwise \end{cases} \tag{3.3}$$

Where $r_{(t+1)}(x, y)$ denotes feature-based shadow processing, average value of M*N sized Δ matrix results in μ_{Δ} and k_2 ranging from 0.5 to 1.5 is a constant. The main drawback of the method is that value of features can be changed easily even with a slight change in significant bits' binary level.

3.2 Randomised Low Rank Dynamic Mode Decomposition (rDMD)

To overcome the problem of inflated computational time to process the high-resolution videos, Ericson and Donovan [64] proposed a model based on data driven method DMD fusing PCA and Fourier Transform. The given technique handled the dynamic background issue in video processing with the help of a randomized matrix algorithm for fast computing the low rank Dynamic Mode Decomposition (DMD). Usage of probabilistic SVD algorithm utilising swiftly decomposing singular values of given data increases the computational savings of the presented method in comparison to traditional SVD algorithms significantly by 10 to 30 times as mentioned by authors. Moreover, randomised DMD is 2 to 3 times faster than the DMD infused with deterministic SVD. SV decomposition is given by the equation:

$$Z = X\Sigma Y^* \tag{3.4}$$

where X, Y are orthogonal matrices and Σ is a diagonal matrix. For improved computational time randomised SVD is used by the authors which is nothing but a linear algebra problem made up of projections and random sampling.

For background modeling main focus is given to the low rank features of the given video stream. Low rank DMD can be expressed as- $A \approx \phi R V_{and}$, ϕ denotes the dynamic modes of the given data matrix, R is an amplitude diagonal matrix described as –

$$R = \begin{pmatrix} r_1 & & & \\ & r_i & & \\ & & \ddots & \\ & & & r_k \end{pmatrix} \tag{3.5}$$

$$V_{and} = \begin{pmatrix} 1 & \lambda_1 & \dots & \lambda_1^{n-1} \\ 1 & \lambda_2 & \dots & \lambda_2^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \lambda_k & \dots & \lambda_k^{n-1} \end{pmatrix} \tag{3.6}$$

The difference between original video data and that of background one gives the identification of foreground objects. The main advantage is that dynamic modes can vary according to the requirement of video data.

3.3 <u>Hierarchical Modeling and Alternating Optimization(HMAO)</u>

The method put forward by Li et al. [65] joins hierarchical modeling with alternating optimisation to separate the foreground and background of video sequence. Background is hierarchically represented by breaking the sequence into high and low frequency counterparts ($z = z^l + z^h$) to deal with dynamic background problems in a better way. Principal components belonging to the residuals are taken into account for detail background instead of wasting them by outlier treatment for finer details. For better robustness towards noise, first the foreground is detected at the low resolution and then move on to further resolutions. Furthermore, graph cut method ensures the joint enforcement of l_1 norm foreground constraint and background's rank-1 constraint.

Final joint estimation problem formed as per the defined background and foreground models above is given in (7) which is further solved using ADMM.

$$\min_{X,F,\Omega}\{\sum_{G} f(a)\|\rho_{i,j,n} - \rho_{x,y,z}\| + \gamma D(\rho) + \beta\|\rho\|_0 + \sum_{i,j} ||X_{i,j,:,:} - X^{l*}_{i,j,:,*4} J_{i,j,*4} O - X^{h}_{i,j,:,:}||\} \quad s.t. rank(X^{l*}_{vec}) = 1, \quad ||O||_2 = 1, \quad D = P_{\overline{\rho}} X + Z \tag{3.7}$$

O denotes the changing tendency; X is the background and $D(\rho)$ represents the prior knowledge of foreground. Model can successfully differentiate regularly appearing background objects from noise.

3.4 <u>Generalised Singular Value Thresholding (GSVT) Operator Based non-Convex LRSD (SVT+ADMM)</u>

Since, the nuclear norm minimises the singular values in one go, this limitation of it for approximation of rank function is solved by the researchers in [60] by exploitation of a surrogate function to solve the low rank matrix and they came up with the idea of performing non-Convex LRSD using generalised SVT operator and solving the formulated problem using multipliers' alternating direction technique known as ADMM. The GSVT operator can be defined as follows-

$$Prox_g^{\sigma}(A) = \underset{L}{argmin} \sum_{j=1}^{x_1} g(\sigma_i(L)) + \frac{1}{2}||L - A||_h^2 \tag{3.8}$$

Where g is a non-convex surrogate function, L is low rank matrix and σ denotes L's singular values. ADMM function in mathematical terms is given as-

$$\mathcal{L}(L, B, C, \mu) = \sum_{j=1}^{x_1} g(\sigma_i(L)) + \lambda||S||_1 - \{C, L + B - M\} + \frac{\mu}{2}||L + B - M||_h^2 \tag{3.9}$$

μ controls the penalty, C is Lagrangian multiplier and {.} is inner product of matrix. Logarithmic penalty is used here by the researchers. The model fails to perform efficiently in presence of complex scenarios.

3.5 Adaptive Motion Estimation + Sequential Outline Separation (AME+SOS)

Thenmozhi and Kalpana [68] used Adaptive Motion Estimation (AME) approach with the along with Sequential Outline Separation (SOS) to put forward an effective object detection technique which efficiently handles issues like sudden light change, closer view gap and ghosting. The method proposed by authors performs foreground detection and BGS to detect in motion objects followed by separation of territory of interest from the constructed background. The condition taken into account is that the background is completely stationary devoid of any moving object. The whole process is divided into four stages comprising of initialization of background and its modeling, subtraction of background and AME followed by SOS. For modeling of background BGS strategy based on dual motion position is exploited to point out the differences between current image and background image. AME segmentation helps in identifying frame consisting of a moving object. In SOS process, the refinement of the collection with the help of vectors in motion is studied. Conversion of video frames into gray ones is carried out as given in (1) to reduce the cost of the process.

$$Z_t = 0.125A_t(a, b) + 0.7512B_t(a, b) + 0.072C_t(a, b) \tag{3.10}$$

A, B and C denotes the colour combinations of each frame. Background modeling is performed according to the given equation below in (2).

$$X_1(a, b) = \begin{cases} X_r(a, b), if X_r(a, b) \le P \\ X_{r-1}(a, b), if X_r(a, b) > P \end{cases} \tag{3.11}$$

X_{r-1} represents previous image background, X_r is current background frame and P is the current frame sequence. The threshold deciding a pixel as a current object or background is defined as:

$$Y(x,y) = q(x,y) + k\sigma(x,y) \tag{3.12}$$

q is the local mean of the pixel, σ being its standard deviation controlled by variable k. Presence of pixel energy factor within the threshold of group pixel energy factors ensures the assignment of that pixel to the cluster. Non-stationary region of images is calculated using difference of image sequence between consecutive frames.

3.6 Non-Convex Rank approximation RPCA and Super-pixel Motion Detection (LSCNC+RPCA)

In this approach [66] traditional RPCA technique is infused with motion detection using super-pixels to perform background separation, and the model is further solved with the help of Lagrange's augmented alternating direction technique. Segmentation method performed using super-pixels uses a particular area of image having some visual importance formed by a group of nearby pixels holding some similar features like texture, colour etc. Authors performed segmentation in initial step to extract the grouping matrix consisting of super pixels. The pixels in the possession of the j^{th} super pixel are marked as j (j=1,2…., k) and a grouping matrix $H \in R^{m*n}$ is obtained by arranging each image frames' information as a column vector, which is later used for the extraction of sparse foreground. Moving forward, improved version of traditional RPCA (1) is incorporated for obtaining sparse matrix.

$$A^{k+1} = \underset{A}{argmin}\|A\|_Y + \frac{\mu^k}{2}\left\|A - (C - D^k - E^K + \frac{F^t}{\mu^t}\right\|_f^2 \tag{3.13}$$

A being low rank matrix, C the sequence of visual frequency, μ being the parameter of penalty, D & E being sparse foreground and dynamic backgrounds and F is Lagrange multiplier. In the final steps results of above two steps are combined, forming a motion mask in return to identify the moving objects in the video. The threshold value t deciding the mask result is calculated as follows-

$$t = \underset{g}{argmin}(\alpha_0(a)\sigma_0^2(a) + \alpha_1(a)\sigma_1^2(a)) \tag{3.14}$$

$\alpha_0(a)$ and $\alpha_1(a)$ denotes moving pixels probability having pixel value less than t, and σ represents the variance between classes. The mentioned method has an upper hand in terms of complex computation and storage due to the inclusion of Linear Spectral Clustering in it during segmentation.

3.7 <u>Comparative Analysis of the Techniques</u>

The most common approach used in BGS techniques to comment on efficiency is by taking the F-measure value calculated on various video sequences taken into account. F- measure is defined in terms of precision and recall defined as - correctly classified positive samples in comparison with total samples and true positive specimens in comparison with total positive classified cluster respectively. Mathematical expressions of these terms are given below –

$$Precision = \frac{T_p}{T_P + F_P} \; ; \qquad (3.15)$$

$$Recall = \frac{T_p}{T_P + F_n} \; ; \qquad (3.16)$$

$$F - measure = \frac{2 * precision * recall}{precision + recall} \qquad (3.17)$$

The F-measure of different models calculated on video sequence of Highway from baseline category of CDnet (2014) dataset is shown in the fig 4. The plot shows that rDMD model has the highest Fm among all the reviewed models. Table 7 presents a tabular representation of important aspects of the above discussed models.

Fig. 4 Highway video sequence with its ground truth

Table 7 Overview of above discussed models in terms of various parameters

S. No.	Algorithm	Issues addressed	Remarks	Limitations	Dataset Used	Authors (Year)
1.	GMRF+ MAP+ Local GLCM	Cast Shadows on moving objects	-Shadows are detected using GLCM dependent local features	-Noise sensitivity -Failure in handling sudden illumination change and data recorded by in motion camera	CDnet, BMC	Badri et al. (2016) [70]
2.	rDMD	Illumination change, Bootstrapping, Camouflage, Dynamic Background	-High computational speed with a frame rate of 180 frames per second.	-Cannot perform well with sleeping foreground objects	CDnet, BMC	Erichson et al. (2016) [64]
3.	HMAO	Dynamic and Complex backgrounds with illumination change, Noise	-Residual principal components are not treated as outliers but are used to obtain detail background.	-Limited performance while dealing with night videos, camera jitter, clutter and irregular dynamic motion in background	CDnet, I2R	Li et al. (2019) [65]
4.	SVT+ ADMM (Non-LRSD)	Baseline effect, Noise in frames	-High accuracy and processing speed with 156 f/s frame rate	-Low efficiency with complex scenes and camouflaged videos.	CDnet, I2R	Yang et al. (2019) [60]
5.	AME+ SOS	Closer view gap, Ghosting, Sudden lighting change	-Classification method based on frame difference is used to detect object motion. -High accuracy of 97.45%	-Model struggles in performance when encountered with complex dynamic background scenes.	CDnet	Thenmozhi et al. (2020) [68]
6.	LSCNC-RPCA	Complex Dynamic background	-Super-pixel blocks can be adjusted according the moving objects' size.	-Fails to extract tiny moving objects present in foreground.	CDnet, I2R	Wang et al. (2020) [66]

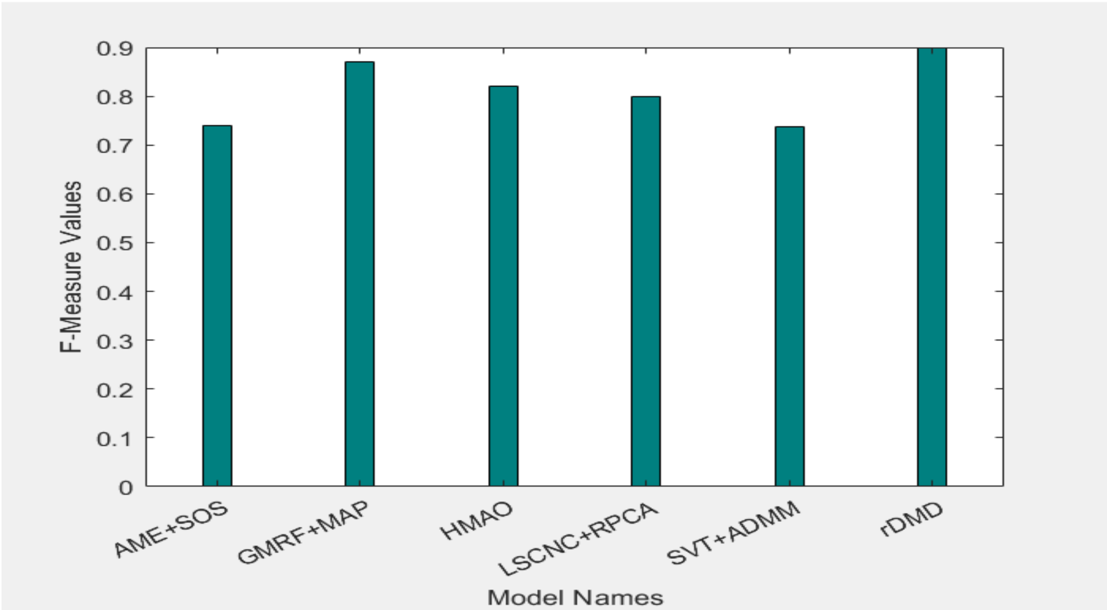

Fig. 5 Plot of Fm values of discussed models for baseline category (Highway) of CDnet dataset

CHAPTER 4

A NEW HYBRID MODELLING BASED TECHNIQUE FOR BACKGROUND SUBTRACTION

4.1 PROPOSED MODEL

We present a new BGS model fusing PCP with rSVD. We are inspired by Candes *et al.* [95], who were first researchers to introduce RPCA model using PCP to represent as a matrix as a decomposition of low rank and sparse components. Here, we have incorporated both the techniques to achieve BGS, by modelling the background using RPCA thus formulating a convex PCP problem and solving it using Augmented Lagrange Multiplier (ALM) method and detecting the foreground by incorporating rSVD to decrease the computational complexity of the method. In addition to this, the computational complexity of the proposed model is further reduced by introducing the scaling factor in the data matrix using which the resolution of the matrix can be controlled and thus having a control over time parameter, as resolution plays a vital role in deciding the time of operation. RPCA model works fine for background modelling but has certain limitations for foreground detection such as, zero prior knowledge of sparse outliers, random uniform distribution of foreground objects etc. which makes it less suitable for foreground detection. Hence, usage of rSVD not only makes the model more robust but also increases the processing speed of the model. Experimental results show that this model is fast as compared to the traditional RPCA model based on PCP.

As described in the introduction, the model proposed in this paper makes use of Principal Component Pursuit (PCP) and randomized Singular value Decomposition (rSVD) to carry out the task of Background Subtraction. Block diagram of the proposed model is shown below in figure 1.

The task of BGS is carried out in three simple steps- a) decomposition of video sequence into frames, b) Background Modelling using PCP and c) Foreground Detection using rSVD. The steps are described as follows-

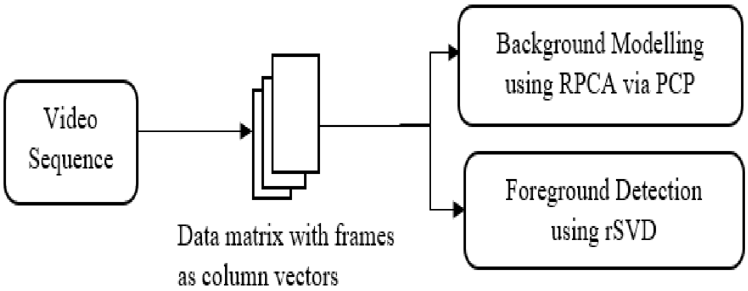

Fig. 6 Block Diagram of the proposed BGS model

A. Video Sequence Decomposition

Video is nothing but a collection of different images known as frames played at a high speed back-to-back. To perform BGS the available video sequence is decomposed into these frames and a data matrix is created by stacking these over one another so that it can be further processed to obtain background and foreground. For a video sequence with say n frames can be described by a data matrix Z, and its one frame can be given as $Z_j \; where \; (1 \leq j \leq n)$.

B. Background Modelling

As proposed by Candes *et al.* [95], we can represent a data matrix easily as a superposition of sparse and low rank components. We incorporated this idea here to acquire the BG model. To obtain the mentioned components PCP is incorporated and LR component gives out the background of the video. For the data matrix Z, the decomposition can be stated as-

$$Z = X + Y \tag{4.1}$$

Where, X denotes the LR component and Y being the sparse one carrying the foreground information. The convex PCP problem of the matrix can be formulated as given in (2).

$$\begin{aligned} &minimize \; ||X||_* + \lambda ||Y||_1 \\ &subject\ to \quad X + Y = Z \end{aligned} \tag{4.2}$$

$||X||_*$ is the nuclear norm of X described as $\Sigma_i \sigma_i(X)$ and $||Y||_1$ is the $l_1$norm of Y, expanded as $\Sigma_{ij}(Y_{ij})$. σ_i represents the matrix's singular values and λ is the control parameter taken generally as (1\n), (where n= no. of frames) for optimized results.

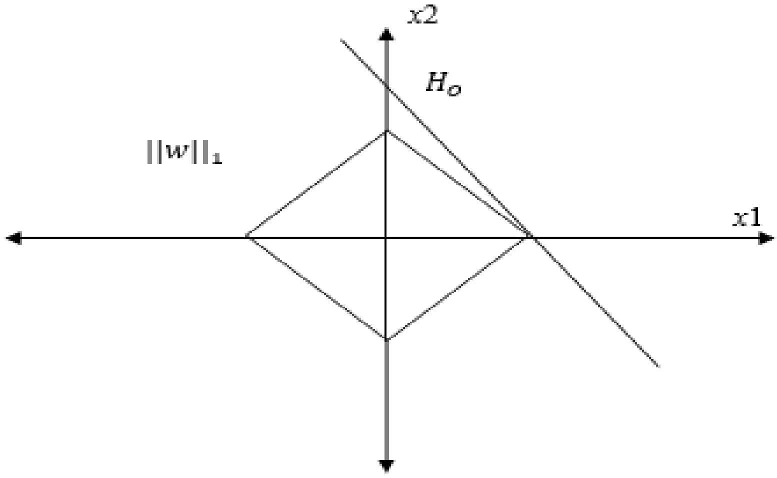

Fig. 7 L1 Regularisation

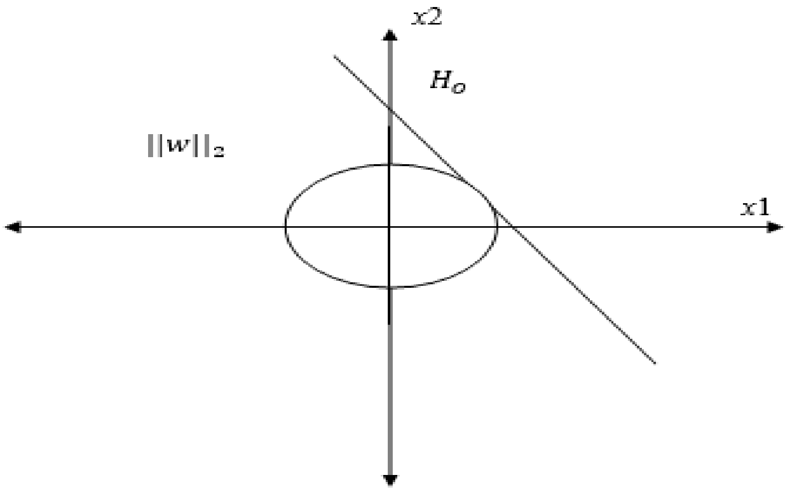

Fig. 8 L2 Regularisation

The nuclear norm is an extension of the trace associated cost, which is often used in the control community for the rank condemnation of positive semidefinite matrices. L_1regularization is similar to MAP approximation, which utilizes Laplacian priors, and it can both inspire sparse solutions and accurately recover a sparse signal. Finally for obtaining the background, the equation (2) is solved by taking help of Augmented Lagrange Multiplier (ALM) method. The mathematical problem statement of ALM to solve (2) is shown below-

$$l(X,Y,M) = \left|\left|X\right|\right|_* + \lambda\left|\left|Y\right|\right|_1 + \{M, Z - X - Y\} + \frac{\mu}{2}\,||Z - X - \,Y||_r^2 \qquad (4.3)$$

μ controls the threshold thus known as thresholding parameter and M defines the ALM matrix. PCP problem to obtain the background can be solved easily by minimizing l with respect to X keeping Y constant and as a result updating A on the basis of residual $(Z - X - Y)$. This equation 3 is solved by continuously updating $(X_k, Y_k) = argmin_{X,Y} l(X, Y, M_k)$ and finally the Lagrange multiplier matrix is updated based on these results as shown in the algorithm below where k represents the step of iteration.

C. Foreground Detection

To achieve this task a method called as rSVD is used. Approximation of matrix factorization can easily be achieved using randomized algorithms reducing the processing complexity and time as discussed in [96]. The listing of traditional decomposing of matrix uses SVD which results in the range of matrix and the prune versions of these are generally incorporated to represent given matrix as a low rank matrix such that

$$E_{m*n} \approx Q_{m*k} \, R_{k*n} \tag{4.4}$$

R is the weak upper triangular matrix and Q being the orthonormal matrix. This factorization scheme can also be exploited to solve computational problems, in this case the problem being sparse matrix Y and E in the above equation can be replaced by Y matrix. Hence rSVD can achieve the task of solving the sparse component Y of data matrix Z, thus leaving us with the foreground of video sequence.

The task at hand to compute Y is shattered into two steps involving the construction of a subspace with low dimensions for capturing matrix action (basis matrix construction) and restricting the subspace matrix to compute standard factorization (rSVD) of the restricted matrix. This process can be explained in detail as follows-

Stage-1: For the computation of estimated basis for matrix Y's range, a matrix with orthonormal columns is needed such that-

$$Y \approx QQ^{*}Y \tag{4.5}$$

Stage-2: Calculation of rSVD of Y. For this let's say we have Q satisfying the equation (5), hence the problem statement of rSVD is given as-

$$Y = U\Sigma V^{*} \tag{4.6}$$

U and V being the m*k and n*k orthonormal matrices of Y and Σ is the k*k non-negative diagonal matrix.

$$\Sigma = \begin{pmatrix} \sigma_1 & & & \\ & \sigma_i & & \\ & & \ddots & \\ & & & \sigma_k \end{pmatrix} \tag{4.7}$$

σ_i denotes matrix Y's singular values, being arranged in the order such as $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_k \geq 0$. Thus, calculation of these components results in the desired matrix Y contributing the foreground of the video sequence. The further iteration steps are added in the algorithm 1 mentioned below. Here, the advantage of using rSVD is the low computational complexity which is achieved by performing only (2q+1) matrix passes, q being the power iteration parameter.

Moreover, when talking about computational complexity, O(m*n(log(k))) floating point operations are required in randomized algorithms in contrast with classic algorithms requiring O(m*n*k) operations. In addition to this randomized approach is less affected by noise and can be identified to incorporate the multi-processor architecture.

Algorithm 1 is the algorithm developed in the proposed model and is defined as shown below.

Algorithm 1

1. **Initialize: M, Y= 0,** $\boldsymbol{\mu = 0}$
2. **while** (convergence≠0) **do**
3. Calculate $\boldsymbol{X_{k+1} = D_{\frac{1}{\mu}}(Z - Y_k + \mu^{-1}M)}$;
4. Calculate$\boldsymbol{M_{k+1} = Y_k + \mu(Z - X_{k+1} + Y_{k+1})}$;
5. **end while**
6. **Output: X** (Background)
7. Form B=$\boldsymbol{Q}^*\boldsymbol{Y}$
8. Calculate B= $\boldsymbol{U\Sigma V}^*$
9. **Output: Y=QB** (Foreground)

4.2 DATASETS AND RESOURCES

In recent years with along with the advancement in the background subtraction techniques, need for the presence of better and enhanced datasets including numerous challenge categories gave rise to many datasets some being small scale datasets and others large scale. Datasets like Wallflower, I2R, CAVIAR etc. are some of the examples of early developed datasets including few challenge categories with video sequences and their ground truth images. Since these were developed early, they were failed to satisfy the need of advanced models being developed. This led to the development of some large-scale datasets like CDnet (2012&2014), GTFD, SBMnet etc. containing realistic video sequences with accurate ground truth covering a fair range of real-world challenges. All of these datasets are publicly available and can be accessed by their links provided in the Background Subtraction Website in the category *Available Datasets*[1][1].Information of some commonly and frequently used datasets by the researchers in the recent years is provided in the table below.

The handiest resource is the *Background subtraction website*[2] that provides a great deal of help being a hub of a huge amount of information available in the field of BGS. It comprises of list of references, Background subtraction steps, issues in performing BGS, links to the surveys, links of available datasets in the field, available implementations, journals and conferences. The information on the site is regularly updated to include new findings. Other resources include a Background subtraction Library developed by Andrews Sobral in 2012 to provide an easy-to-use C++ framework for foreground background separation in videos using OpenCV. The BGS library is compatible with OpenCV 2.4.x, 3.x and 4.x, and compiles under Windows, Linux, and Mac OS X. Currently the library contains **43** algorithms. The source code is available under the MIT license, the library is available free of charge to all users, academic and commercial.

---

[1] http://sites.google.com/site/backgroundsubtraction/testsequences

[2] http://sites.google.com/site/backgroundsubtraction/

Table 8 Overview of some most commonly used datasets

Dataset	**Video Sequences (Year)**	**Categories Included**
Wallflower	7(1999)	Moved Background, Gradual Illumination change, Sudden Illumination change, Dynamic Background, Camouflage, Bootstrapping, Foreground aperture
I2R	9(2004)	Gradual Illumination change, Sudden Illumination change, Dynamic Background, Camouflage, Bootstrapping, Challenging Weather, Video Noise
OTCBVS	10(2004)	Thermal, Challenging Weather and other Dataset categories
CAVIAR(Context Aware Vision Using Image-based Active Recognition)	54(2002-2005)	Different Human Activities (Walking, fighting, resting, browsing, slumping, group meeting, leaving baggage, shop entering and shop exiting)
SABS (Stuttgart Artificial Background Subtraction)	2011	Basic surveillance scenes, Dynamic background, Gradual illumination changes, Sudden illumination changes, Shadow, bootstrapping, Camouflage, Video compression, and video noise.
CDnet(2012)	31(2012)	Thermal, Shadows, Intermittent Object Detection, Camera Jitter, Dynamic Background, Baseline
BMC	29(2012)	Sudden illumination changes, Casted shadows, Challenging weather, Big foreground, and Dynamic background
SBM-RGBD	33(2013)	Depth Camouflage, Intermittent motion, Out of sensor range, Shadows, Bootstrapping
Fish4Knowledge	14(2010-2013)	Occlusion, Complex textures of background, Low contrasted video frames, Camouflage, Dynamic background, and Illumination changes
CDnet(2014)	53(2014)	Bad Weather, Low Framerate Night Videos, PTZ Camera Thermal, Shadows, Intermittent Object Detection, Camera Jitter, Dynamic Background, Baseline, Turbulence
SBMnet (Scene Background Modelling. NET)	79(2016)	Basic, Intermittent motion, Clutter, Jitter, Illumination changes, Background motion, Very long, and Very short
CAMO-UOW	10(2017)	Camouflage
GTFD (Gray Scale Thermal Foreground Detection)	25(2017)	Intermittent Motion, Bad Weather, Low illumination, Intense shadow, Dynamic Scene, Background Clutter, Thermal Crossover

CHAPTER 5
RESULTS AND CONCLUSION

5.1 EXPERIMENTAL ANALYSIS

The experiments were conducted on three different datasets[3] namely *Highway* sequence from Baseline category of CDnet (2014) dataset, *walking people* video sequence of BMC 2012 Background Models Challenge Dataset and a random YouTube video namely *monkey-giraffe*. We have used a Intel Core i7 processor with 8 GB RAM to run the experiments on python 3.6 environment.

Highway is a video sequence of a highway video surveillance camera with slow moving cars in it. The foreground objects here are slow and medium sized moving objects. The next video sequence is a surveillance clip monitoring people roaming around. Frames contains people roaming at different time intervals. Here also the foreground objects are slowly moving but their size is smaller. The final dataset is a YouTube video showing two toys. The foreground objects of this video sequence are moving at a fast speed as compared to the other two video sequences.

The figures 8,9 and 10 below shows the experimental results of our model on all three video sequences. Resolution of the frames is 0.75 in the experiments as it was the highest, we can work with without the storage issues of data matrix. As evident from the results, our model performs well with slow moving foreground objects irrespective of their sizes, although with fast moving objects the detected foreground is not so accurate with some disturbance. Also, experiments shows that our model takes less time to achieve this process as compared to the traditional PCP model based on RPCA. Table 9 shows the comparative results of both models with respect to time. The comparison is performed on different sequences with different resolutions as videos with more frames needs more space on device, therefore the resolution of large video is kept less to deal with storage issues. Figure 11 presents the visual comparison of time difference of the two methods compared in the table.

[3]http://sites.google.com/site/backgroundsubtraction/testsequences

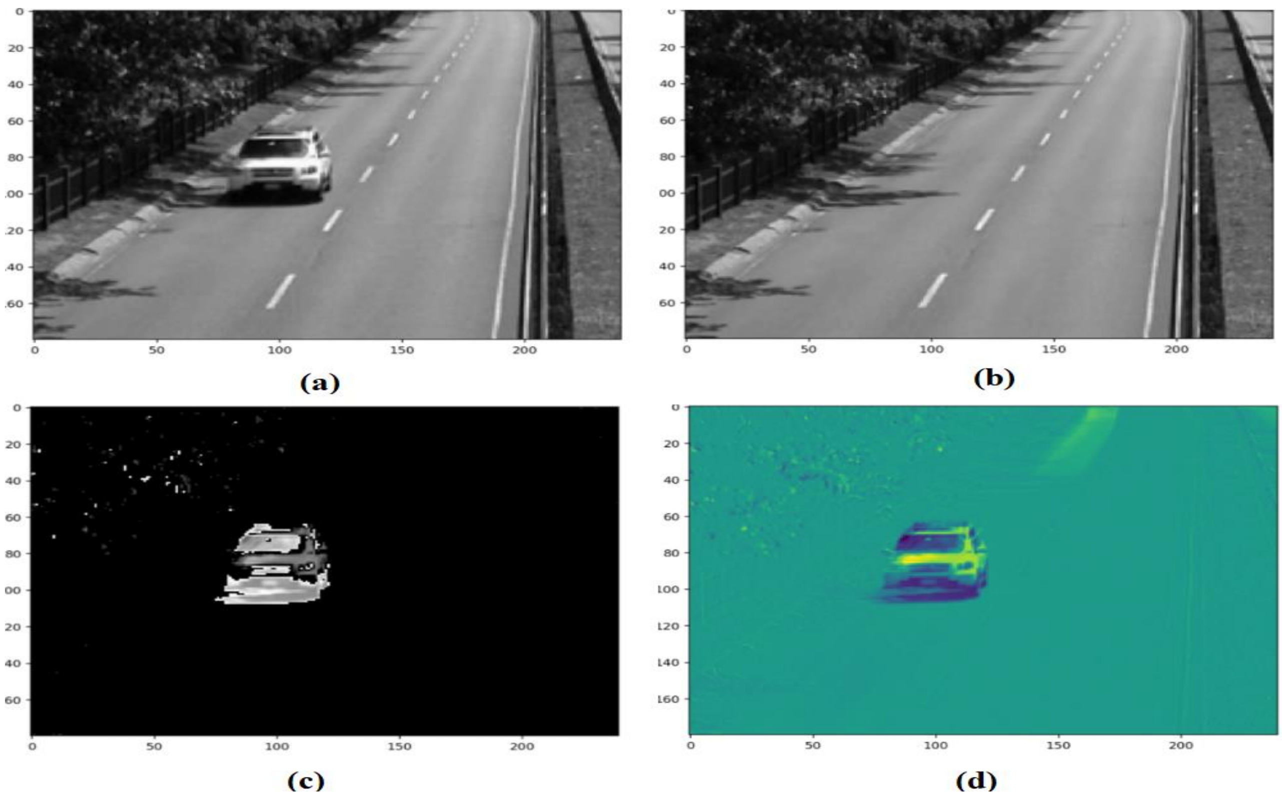

Fig. 9 Proposed method results on the Highway Dataset (a) Original frame of video sequence, (b) Background image obtained, (c) Foreground mask and (d) Extracted foreground.

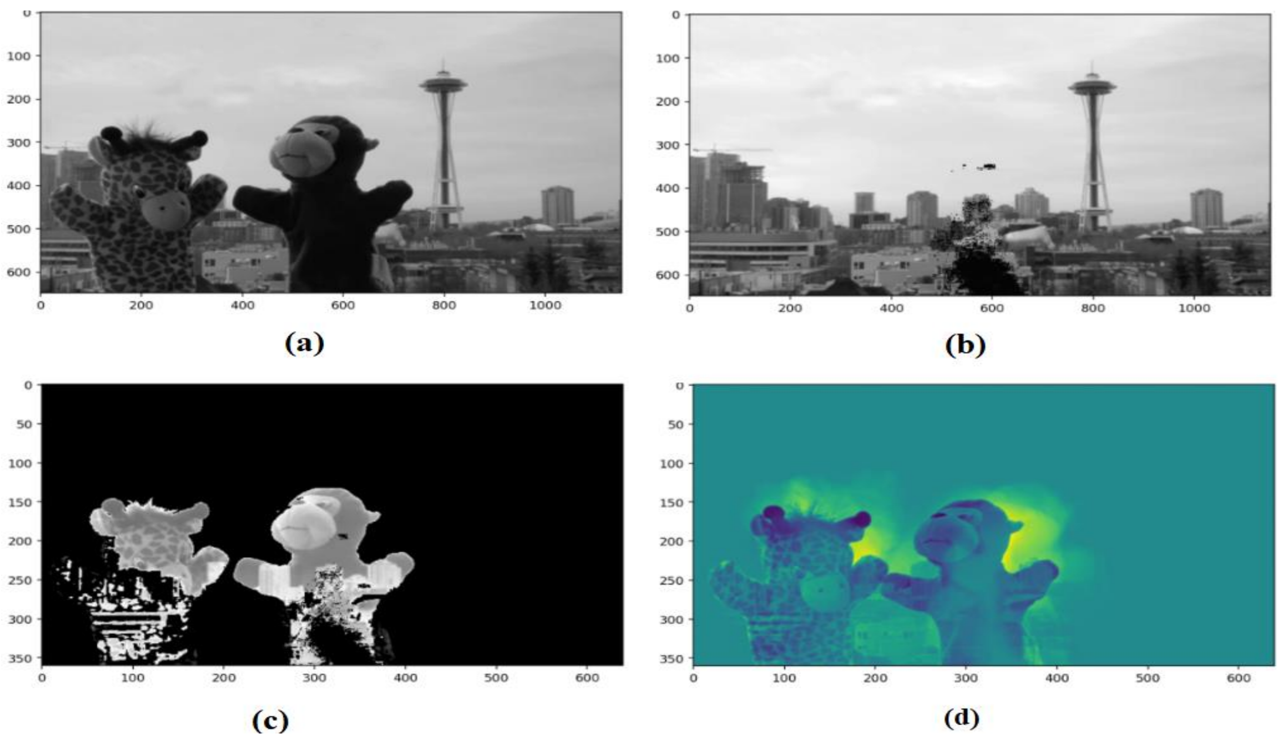

Fig. 10 Proposed method results on the monkey-giraffe Dataset (a) Original frame of video sequence, (b) Background image obtained, (c) Foreground mask and (d) Extracted foreground.

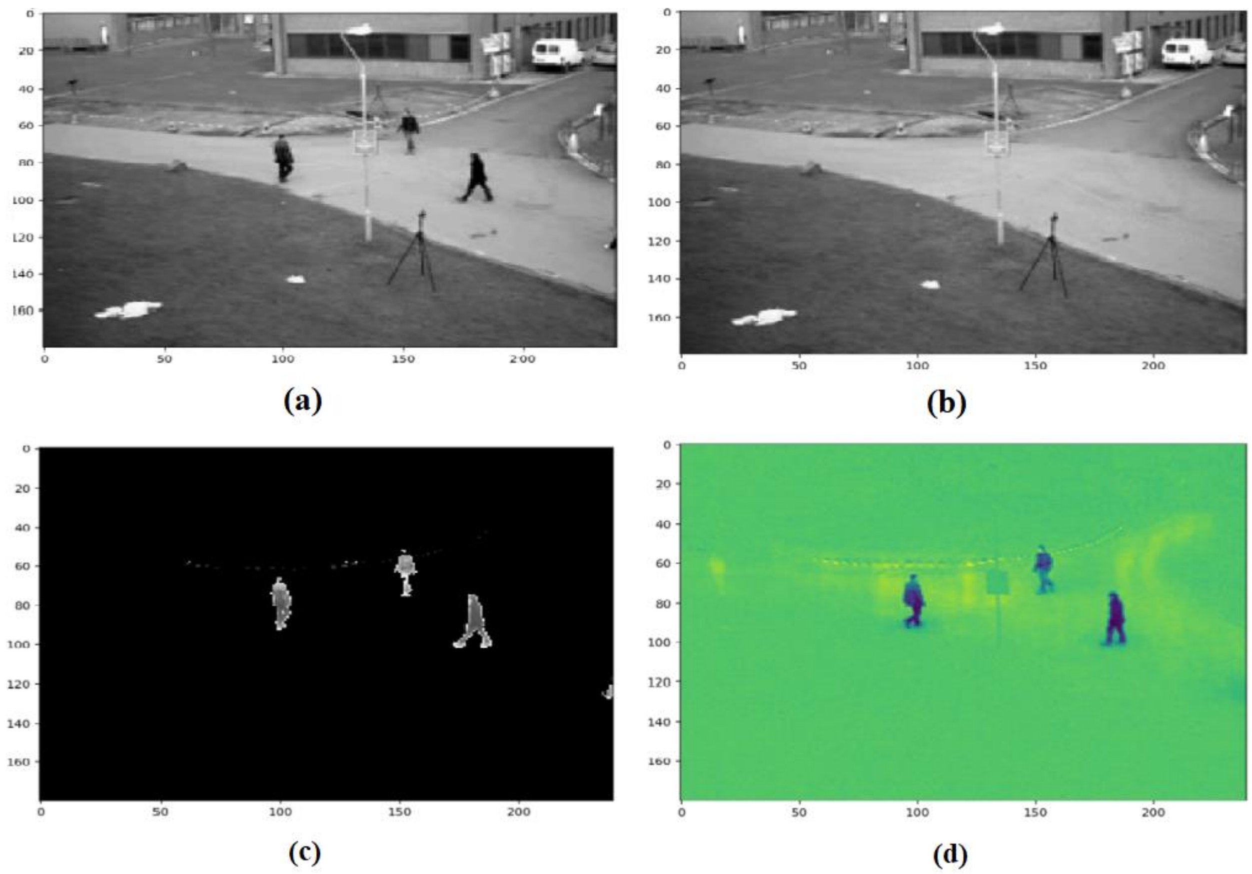

Fig. 11 Proposed method results on the Walking People Dataset (a) Original frame of video sequence, (b) Background image obtained, (c) Foreground mask and (d) Extracted foreground.

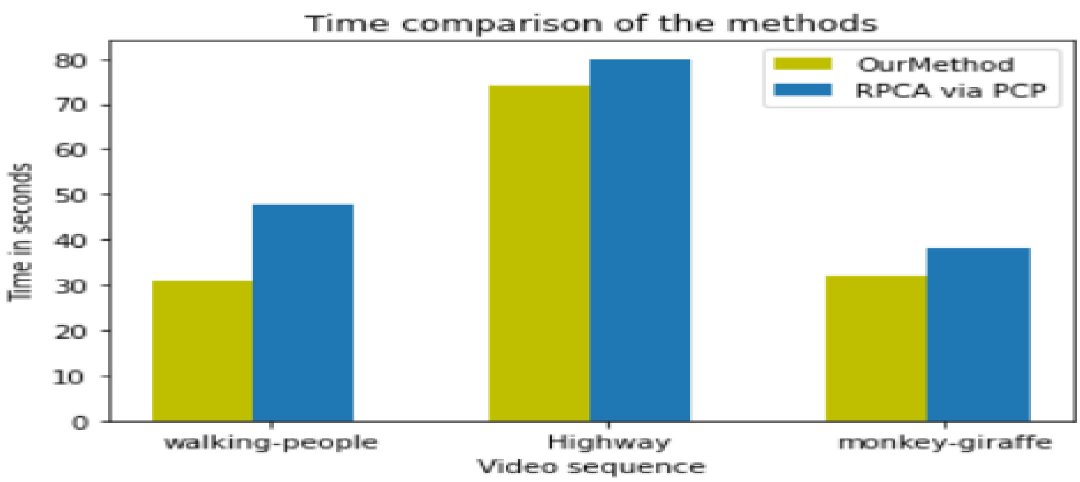

Fig. 12 Visual Comparison of the methods mentioned in Table 9.

Table 9 Comparison of proposed model with the traditional PCP model.

Video Sequence	Resolution	No. of frames	Time in seconds	
			Proposed Method	RPCA via PCP
Walking-people	0.35	50	31	48
Highway	0.25	180	74	80
Monkey-giraffe	0.75	10	32	38

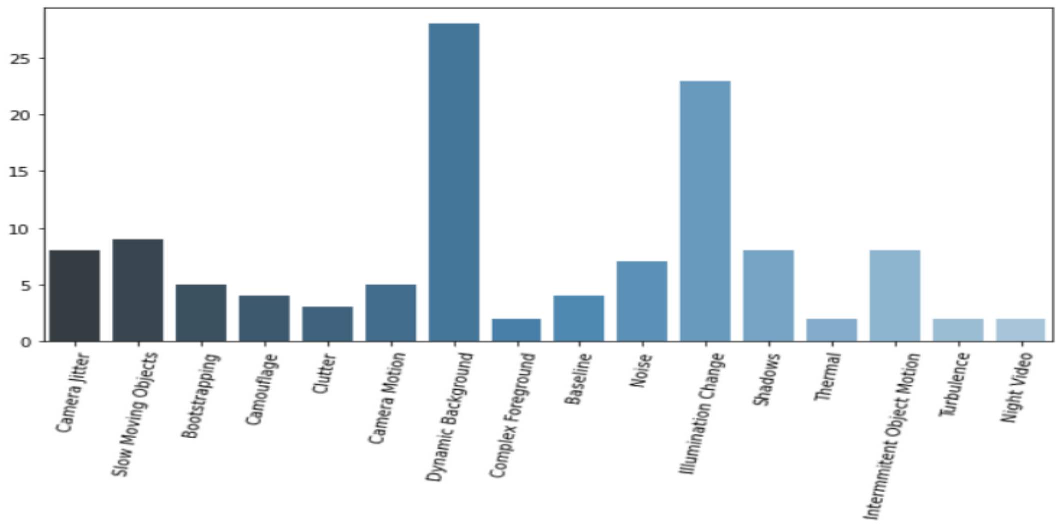

Fig. 13 Plot showing the various challenges addressed by the number of models.

5.2 CONCLUSION AND FUTURE SCOPE

In this report, a new hybrid modeling-based strategy that combines two well-known models, RPCA via PCP and rSVD is introduced. The combination of the two models resulted in a more robust BGS model with lower computing complexity. Advantages of the proposed model over the traditional PCP model are as follows: (i) The approach works effectively with video sequences featuring slow moving objects, regardless of their size, (ii) The time taken by the proposed model to calculate the results of BGS is comparatively less compared to the PCP model and (iii) Model provides accurate and fast results for dynamic background video sequences. A comparison of the traditional RPCA via PCP model with the proposed one in terms of processing time demonstrates that the latter is preferable to the former. When it comes to the concept of fast-moving foreground objects, however, the proposed model has certain drawbacks.

In addition to that, it can be concluded that the hybrid models are successful in addressing most of the difficult issues such as complex background, camouflage etc. Algorithm formed using randomised matrix can also be used for other linear algebra-based models based specially those using SVD. The use of DMD is an interesting topic as it can give fast results but on the cost of high precision and it can be presented as RPCA's fast approximation. AME+SOS comes in rescue when the question is about dealing with complex videos and shaking camera with pretty convincing accuracy. Issue of insufficient boundary observation and holes in extraction of foreground can be handled effectively by LSCNC+RPCA. Most of the issues with dynamic background is solved by HMAO with a performance better than many available states of art methods. Although the models reviewed here are latest in the literature and they try to address some critical issues but still they are not enough when issues like night video, thermal, turbulence etc. comes into picture.

The model's performance will be improved in the future to address other BGS concerns. Moreover, in-depth exploration of deep neural networks may be a future research trend, as CNN-based models are commonly used to manage complex background. Many challenges, such as moving camera, PTZ videos, dark videos, turbulence, and less frame rate, are less touched in the field and require more attention. Furthermore, while models like STLBP2 and some hybrid models attempt to address a variety of challenges, they fall short in the presence of a camouflage environment and complex scenes. As a result, a model that is resistant to all challenges is still missing from the literature, and it is something that will be worked on in the future. Thus, Future work may include development of a more robust model handling a greater number of issues along with the critical ones as well.

REFERENCES

[1] R. J. Radke, S. Andra, O. Al-Kofahi and B. Roysam, "Image change detection algorithms: a systematic survey," in IEEE Transactions on Image Processing, vol. 14, no. 3, pp. 294-307, March 2005, doi: 10.1109/TIP.2004.838698.

[2] X. Cui, D. Metaxas, Q. Liu and S. Zhang,” Temporal spectral residual for fast salient motion detection”, Neurocomputing, vol. 86, 2012, pp. 24-32, https://doi.org/10.1016/j.neucom.2011.12.033.

[3] A. Mc Ivor, “Background subtraction techniques, in: International Conference on Image and Vision Computing”, New Zealand, IVCNZ 2000, November 2010.

[4] M. Piccardi, "Background subtraction techniques: a review," 2004 IEEE International Conference on Systems, Man and Cybernetics (IEEE Cat. No.04CH37583), The Hague, Netherlands, 2004, pp. 3099-3104 vol.4. doi: 10.1109/ICSMC.2004.1400815.

[5] S. Cheung and C. Kamath, “Robust background subtraction with foreground validation for urban traffic video”, EURASIP J. Adv. Signal Process. 2005, 726261 (2005), https://doi.org/10.1155/ASP.2005.2330.

[6] M. Cristani, M. Farenzena, D. Bloisi and V. Murino, ”Background subtraction for automated multisensor surveillance: A comprehensive review”, EURASIP J. Adv. Signal Process. 2010, https://doi.org/10.1155/2010/343057.

[7] T. Bouwmans, F. El-Baf and B. Vachon, “Statistical background modeling for foreground detection: A survey”, in: Handbook of Pattern Recognition and Computer Vision, vol. 4(2),World Scientific Publishing, 2010, pp. 181–199.

[8] Thierry Bouwmans and El Hadi Zahzah, “Robust PCA via Principal Component Pursuit: A review for a comparative evaluation in video surveillance”, Computer Vision and Image Understanding, vol. 122, 2014, pp. 22–34, https://doi.org/10.1016/j.cviu.2013.11.009.

[9] Thierry Bouwmans**, “**Traditional and recent approaches in background modeling for foreground detection: An overview” Computer Science Review, vol. 11-12, 2014, pp. 31-66, https://doi.org/10.1016/j.cosrev.2014.04.001.

[10] Thierry Bouwmans, Andrews Sobral , Sajid Javed , Soon Ki Jung and El-Hadi Zahzah, “Decomposition into low-rank plus additive matrices for background/foreground separation: A review for a comparative evaluation with a large-scale dataset”, Computer Science Review,Volume 23, 2017, pp. 1-71, https://doi.org/10.1016/j.cosrev.2016.11.001.

[11] Thierry Bouwmans , Sajid Javed , Maryam Sultana and Soon Ki Jung, “Deep neural network concepts for background subtraction: A systematic review and comparative evaluation”, Neural Networks, vol. 117, 2019, pp. 8–66, https://doi.org/10.1016/j.neunet.2019.04.024.

[12] Belmar Garcia-Garcia, Thierry Bouwmans and Alberto Jorge Rosales Silva, “Background subtraction in real applications: Challenges, current models and future directions”, Computer Science Review, vol. 35, 2020.

[13] E. J. Candès, X. Li, Y. Ma, and J. Wright, ``Robust principal component analysis?'' J. ACM, vol. 58, no. 3, p. 11, May 2011.

[14] Jiashi Feng, Huan Xu, and Shuicheng Yan, ”Online Robust PCA via stochastic optimization”, In Proceedings of the 26th International Conference on Neural Information Processing Systems - Volume 1, 2013, Curran Associates Inc., Red Hook, NY, USA, 404–412.

[15] G. Han, J. Wang and X. Cai,” Background subtraction based on modified online robust principal component analysis”, Int. J. Mach. Learn. & Cyber, vol. 8, 2017.

[16] J. Xue, Y. Zhao, W. Liao and J. Cheung-Wai Chan, "Total Variation and Rank-1 Constraint RPCA for Background Subtraction," in IEEE Access, vol. 6, pp. 49955-49966, 2018, doi: 10.1109/ACCESS.2018.2868731.

[17] S. Wang, Y. Wang, Y. Chen, P. Pan, Z. Sun and G. He, "Robust PCA Using Matrix Factorization for Background/Foreground Separation," in IEEE Access, vol. 6, pp. 18945-18953, 2018, doi: 10.1109/ACCESS.2018.2818322.

[18] B. E. Moore, C. Gao and R. R. Nadakuditi, "Panoramic Robust PCA for Foreground–Background Separation on Noisy, Free-Motion Camera Video," in IEEE Transactions on Computational Imaging, vol. 5, no. 2, pp. 195-211, June 2019, doi: 10.1109/TCI.2019.2891389.

[19] S. E. Ebadi and E. Izquierdo, "Foreground Segmentation with Tree-Structured Sparse RPCA," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 40, no. 9, pp. 2273-2280, 1 Sept. 2018, doi: 10.1109/TPAMI.2017.2745573.

[20] W. Cao, Y. Wang, J. Sun, D. Meng, C. Yang, A. Cichocki and Z. Xu, "Total Variation Regularized Tensor RPCA for Background Subtraction From Compressive Measurements," in IEEE Transactions on Image Processing, vol. 25, no. 9, pp. 4075-4090, Sept. 2016, doi: 10.1109/TIP.2016.2579262

[21] Javed Sajid, Soon Ki Jung, Arif Mahmood and Thierry Bouwmans, ”Motion-Aware Graph Regularized RPCA for Background Modeling of Complex Scenes”, 23rd International Conference on Pattern Recognition (ICPR) Cancún Center, 2016,doi: 10.1109/ICPR.2016.7899619.

[22] Chen R, Tong Y, Yang J and Wu M, “Video Foreground Detection Algorithm Based on Fast Principal Component Pursuit and Motion Saliency”, Comput Intell Neurosci. 2019 Feb 3;2019:4769185. doi: 10.1155/2019/4769185.

[23] Srivatsa Prativadibhayankaram, Huynh Van Luong, Thanh Ha Le and Andre Kaup, ”Compressive Online Video Background–Foreground Separation Using Multiple Prior Information and Optical Flow”, Journal of Imaging.,2018.

[24] H. Yong, D. Meng, W. Zuo and L. Zhang, "Robust Online Matrix Factorization for Dynamic Background Subtraction," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 40, no. 7, pp. 1726-1740, 1 July 2018, doi: 10.1109/TPAMI.2017.2732350.

[25] Min Wu, Yubao Sun, Renlong Hang, Qingshan Liu and Guangcan Liu, “Multi-Component Group Sparse RPCA Model for Motion Object Detection under Complex Dynamic Back-ground”, Neurocomputing (2018), doi:10.1016/j.neucom.2018.03.001

[26] Aihua Zheng, Minghe Xu, Bin Luo, Zhili Zhou and Chenglong Li, “CLASS: Collaborative Low-Rank and Sparse Separation for Moving Object Detection”. Cogn Comput , vol. 9**,** 2017 pp. 180–193. https://doi.org/10.1007/s12559-017-9449-5

[27] Athisha S., Keerthi Krishnan K. and Sreelekshmi P.S., “Separating Moving Objects from Stationary Background using Dynamic Mode Decomposition”, International Journal of Scientific Research in Network Security and Communication, Vol.6, Issue.3, pp.58-64, 2018.

[28] S. Javed, A. Mahmood, T. Bouwmans and S. K. Jung, "Spatiotemporal Low-Rank Modeling for Complex Scene Background Initialization," in IEEE Transactions on Circuits and Systems for Video Technology, vol. 28, no. 6, pp. 1315-1329, June 2018, doi: 10.1109/TCSVT.2016.2632302.

[29] M. Ma, R. Hu, S. Chen, J. Xiao and Z. Wang, "Robust background subtraction method via low-rank and structured sparse decomposition," in China Communications, vol. 15, no. 7, pp. 156-167, July 2018, doi: 10.1109/CC.2018.8424611.

[30] B. Rezaei and S. Ostadabbas, "Moving Object Detection Through Robust Matrix Completion Augmented With Objectness," in IEEE Journal of Selected Topics in Signal Processing, vol. 12, no. 6, pp. 1313-1323, Dec. 2018, doi: 10.1109/JSTSP.2018.2869111.

[31] L. Zhu, Y. Hao and Y. Song, "$L_{1/2}$ Norm and Spatial Continuity Regularized Low-Rank Approximation for Moving Object Detection in Dynamic Background," in IEEE Signal Processing Letters, vol. 25, no. 1, pp. 15-19, Jan. 2018, doi: 10.1109/LSP.2017.2768582.

[32] L. Chen, X. Jiang, X. Liu, T. Kirubarajan and Z. Zhou, "Outlier-Robust Moving Object and Background Decomposition via Structured L_P Regularized Low-Rank Representation," in IEEE Transactions on Emerging Topics in Computational Intelligence, doi: 10.1109/TETCI.2019.2935747.

[33] Aihua Zheng, Tian Zou, Yumiao Zhao, Bo Jiang, Jin Tang and Chenglong Li, “Background subtraction with multi-scale structured low-rank and sparse factorization”, Neurocomputing, vol. 328, 2019, pp. 113-121, ISSN 0925-2312.

[34] Zhenzhen Yang, Lu Fan, Yongpeng Yang, Zhen Yang and Guan Gui, "Generalized nuclear norm and Laplacian scale mixture based low-rank and sparse decomposition for video foreground-background separation", Signal Processing, vol. 172, 2020, 107527, ISSN 0165-1684.

[35] Y. Yang, Z. Yang, J. Li and L. Fan, "Foreground-Background Separation via Generalized Nuclear Norm and Structured Sparse Norm Based Low-Rank and Sparse Decomposition," in IEEE Access, vol. 8, pp. 84217-84229, 2020, doi: 10.1109/ACCESS.2020.2992132.

[36] Huaxin Xiao, Yu Liu and Maojun Zhang, “Fast ℓ 1-minimization algorithm for robust background subtraction”’ J Image Video Proc., vol. 45, 2016. https://doi.org/10.1186/s13640-016-0150-5

[37] Z. Zeng, J. Jia, D. Yu, Y. Chen and Z. Zhu, "Pixel Modeling Using Histograms Based on Fuzzy Partitions for Dynamic Background Subtraction," in IEEE Transactions on Fuzzy Systems, vol. 25, no. 3, pp. 584-593, June 2017, doi: 10.1109/TFUZZ.2016.2566811.

[38] T. Yu, J. Yang and W. Lu, "Dynamic Background Subtraction Using Histograms Based on Fuzzy C-Means Clustering and Fuzzy Nearness Degree," in IEEE Access, vol. 7, pp. 14671-14679, 2019, doi: 10.1109/ACCESS.2019.2893771.

[39] Ali Darwich, Pierre-Alexandre Hébert, André Bigand and Yasser Mohanna, "Background Subtraction Based on a New Fuzzy Mixture of Gaussians for Moving Object Detection", J. Imaging 4, 2018, no. 7: 92.

[40] K. Kima, T. H. Chalidabhongseb, D. Harwooda and L.Davisa, "Real Time Foreground-Background Segmentation Using a Modified Codebook Model," in Advanced Video and Signal Based Surveillance, 2009. AVSS '09. Sixth IEEE International Conference, Genova, 2009.

[41] R. Krungkaew and W. Kusakunniran, "Foreground segmentation in a video by using a novel dynamic codebook," 2016 13th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), Chiang Mai, Thailand, 2016, pp. 1-6, doi: 10.1109/ECTICon.2016.7561253.

[42] M. Qin, Y. Lu, H. Di and W. Huang, "A Background Basis Selection-Based Foreground Detection Method," in IEEE Transactions on Multimedia, vol. 18, no. 7, pp. 1283-1296, July 2016, doi: 10.1109/TMM.2016.2557729.

[43] G. Gemignani and A. Rozza, "A Robust Approach for the Background Subtraction Based on Multi-Layered Self-Organizing Maps," in *IEEE Transactions on Image Processing*, vol. 25, no. 11, pp. 5239-5251, Nov. 2016, doi: 10.1109/TIP.2016.2605004.

[44] Yao Wang, Zujun Yu and Liqiang Zhu,” Foreground Detection with Deeply Learned Multi-Scale Spatial-Temporal Features”, Sensors, vol. 18, no. 12, p. 4269. Dec. 2018.

[45] Amirreza Farnoosh, Behnaz Rezaei and Sarah Ostadabbas, ”DeepPBM: Deep Probabilistic Background Model Estimation from Video Sequences”, Computer Vision and Pattern Recognition, 2019, arXiv:1902.00820.

[46] T. Yu, J. Yang, and W. Lu, “Refinement of Background-Subtraction Methods Based on Convolutional Neural Network Features for Dynamic Background,” Algorithms, vol. 12, no. 7, p. 128, Jun. 2019.

[47] Mithulaya Vijayan and R. Mohan,”A Universal Foreground Segmentation Technique using Deep-Neural Network”, Multimed Tools Appl , vol. 79, 2020.

[48] P. St-Charles, G. Bilodeau and R. Bergevin, "Universal Background Subtraction Using Word Consensus Models," in IEEE Transactions on Image Processing, vol. 25, no. 10, pp. 4768-4781, Oct. 2016, doi: 10.1109/TIP.2016.2598691.

[49] Maryam Sultana, Arif Mahmood, Sajid Javed and Soon Ki Jung, "Unsupervised deep context prediction for background estimation and foreground segmentation", Machine Vision and Applications, vol. 30, 2019.

[50] G. Shi, T. Huang, W. Dong, J. Wu and X. Xie, "Robust Foreground Estimation via Structured Gaussian Scale Mixture Modeling," in IEEE Transactions on Image Processing, vol. 27, no. 10, pp. 4810-4824, Oct. 2018, doi: 10.1109/TIP.2018.2845123.

[51] Thangarajah Akilan, Q.M. Jonathan Wu and Yimin Yang, "Fusion-based foreground enhancement for background subtraction using multivariate multi-model Gaussian distribution", Information Sciences, vol. 430–431, 2018, pp. 414-431,ISSN 0020-0255.

[52] Guian Zhang, Zhiyong Yuan, Qianqian Tong, Mianlun Zheng and Jianhui Zhao, "A novel framework for background subtraction and foreground detection", Pattern Recognition, vol. 84, 2018, pp. 28-38, ISSN 0031-3203.

[53] Y. Liu and D. A. Pados, "Compressed-Sensed-Domain L1-PCA Video Surveillance," in IEEE Transactions on Multimedia, vol. 18, no. 3, pp. 351-363, March 2016, doi: 10.1109/TMM.2016.2514848.

[54] H. Huang, X. Liu, T. Zhang and B. Yang, "Regression PCA for Moving Objects Separation," GLOBECOM 2020 - 2020 IEEE Global Communications Conference, Taipei, Taiwan, 2020, pp. 1-6, doi: 10.1109/GLOBECOM42002.2020.9322471.

[55] H. Mahdian Toroghi, M. Mirzarezaee and B. Najar Araabi. "Foreground Extraction Using Hilbert-Schmidt Independence Criterion and Particle Swarm Optimization Independent Component Analysis". International Journal of Engineering, vol. 33, no. 5, pp. 1020-1026, 2020. doi: 10.5829/ije.2020.33.05b.36.

[56] I. Kajo, N. Kamel and Y. Ruichek, "Incremental Tensor-Based Completion Method for Detection of Stationary Foreground Objects," in IEEE Transactions on Circuits and Systems for Video Technology, vol. 29, no. 5, pp. 1325-1338, May 2019, doi: 10.1109/TCSVT.2018.2841825.

[57] S. Pei, L. Li, L. Ye and Y. Dong, "A Tensor Foreground-Background Separation Algorithm Based on Dynamic Dictionary Update and Active Contour Detection,"

in IEEE Access, vol. 8, pp. 88259-88272, 2020, doi: 10.1109/ACCESS.2020.2992494.

[58] L. Guo, D. Xu and Z. Qiang, "Background Subtraction Using Local SVD Binary Pattern," *2016 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, Las Vegas, NV, USA, 2016, pp. 1159-1167, doi: 10.1109/CVPRW.2016.148.

[59] P. St-Charles, G. Bilodeau and R. Bergevin, "SuBSENSE: A Universal Change Detection Method With Local Adaptive Sensitivity," in *IEEE Transactions on Image Processing*, vol. 24, no. 1, pp. 359-373, Jan. 2015, doi: 10.1109/TIP.2014.2378053.

[60] Zhenzhen Yang, Lu Fan, Yongpeng Yang, Zhen Yang and Guan Gui, "Generalized singular value thresholding operator based nonconvex low-rank and sparse decomposition for moving object detection", Journal of the Franklin Institute, vol. 356, issue 16, 2019, pp. 10138-10154.

[61] I. Kajo, N. Kamel, Y. Ruichek and A. S. Malik, "SVD-Based Tensor-Completion Technique for Background Initialization," in *IEEE Transactions on Image Processing*, vol. 27, no. 6, pp. 3114-3126, June 2018, doi: 10.1109/TIP.2018.2817045.

[62] W. He, Y. K. Kim, H. Ko, J. Wu, W. Li and B. Tu, "Local Compact Binary Count Based Nonparametric Background Modeling for Foreground Detection in Dynamic Scenes," in IEEE Access, vol. 7, pp. 92329-92340, 2019, doi: 10.1109/ACCESS.2019.2927745.

[63] Zuofeng Zhong, Jiajun Wen, Bob Zhang and Yong Xu, "A general moving detection method using dual-target nonparametric background model", Knowledge-Based Systems, vol. 164, 2019, pp. 85-95.

[64] N. Benjamin Erichson and Carl Donovan, "Randomized low-rank Dynamic Mode Decomposition for motion detection", Computer Vision and Image Understanding, vol. 146, 2016, pp. 40-50.

[65] L. Li, Q. Hu and X. Li, "Moving Object Detection in Video via Hierarchical Modeling and Alternating Optimization," in IEEE Transactions on Image Processing, vol. 28, no. 4, pp. 2021-2036, April 2019, doi: 10.1109/TIP.2018.2882926.

[66] Y. Wang, H. Wei, X. Ding and J. Tao, "Video Background/Foreground Separation Model Based on Non-Convex Rank Approximation RPCA and Superpixel Motion

Detection," in IEEE Access, vol. 8, pp. 157493-157503, 2020, doi: 10.1109/ACCESS.2020.3018705.

[67] Y. Tocker, R. R. Hagege and J. M. Francos, "Dynamic Spatial Predicted Background," in IEEE Transactions on Image Processing, vol. 29, pp. 5517-5530, 2020, doi: 10.1109/TIP.2020.2983598.

[68] T. Thenmozhi and A.M. Kalpana, "Adaptive motion estimation and sequential outline separation based moving object detection in video surveillance system", Microprocessors and Microsystems, vol. 76, 2020, 103084.

[69] D. Mohanapriya and K. Mahesh, ”A novel foreground region analysis using NCP-DBP texture pattern for robust visual tracking”, Multimed Tools Appl, vol. 76, 2017.

[70] Badri Narayan Subudhi, Susmita Ghosh, Sung-Bae Cho and Ashish Ghosh, "Integration of fuzzy Markov random field and local information for separation of moving objects and shadows", Information Sciences, vol. 331, 2016, pp. 15-31, ISSN 0020-0255.

[71] Danilo Avola , Marco Bernardi , Luigi Cinque , Cristiano Massaroni and Gian Luca Foresti “Fusing Self-Organized Neural Network and Keypoint Clustering for Localized Real-Time Background Subtraction.” International journal of neural systems vol. 30,4 (2020): 2050016. doi:10.1142/S0129065720500161.

[72] Saqib Umer, Hassan Dawood, Muhammad Haroon Yousaf, Hussain Dawood and Haseeb Ahmad, "Efficient foreground object segmentation from video by Probability Weighted Moments, Optik, vol. 229, 2021.

[73] N.B . Erichson, S.L. Brunton and J.N. Kutz, “Compressed dynamic mode decomposition for background modeling. J Real-Time Image Proc, vol. 16, 1479–1492, 2019. https://doi.org/10.1007/s11554-016-0655-2.

[74] R. Zhao and V. Y. F. Tan, "Online Nonnegative Matrix Factorization With Outliers," in IEEE Transactions on Signal Processing, vol. 65, no. 3, pp. 555-570, 1 Feb.1, 2017, doi: 10.1109/TSP.2016.2620967.

[75] Lin Zhu, Xiurong Jiang, Jianing Li, Yuanhong Hao, and Yonghong Tian, “Motion-Aware Structured Matrix Factorization for Foreground Detection in Complex Scenes”, ACM Trans. Multimedia Comput. Commun. Appl. 16, 4, Article 123 , January 2021. 23 pages. DOI:https://doi.org/10.1145/3407188

[76] Jingyu Yang, Wen Shi, Huanjing Yue, Kun Li, Jian Ma and Chunping Hou, "Spatiotemporally scalable matrix recovery for background modeling and moving object detection", Signal Processing, vol. 168,2020,107362,ISSN 0165-1684.

[77] Y. Lin, Y. Tong, Y. Cao, Y. Zhou and S. Wang, "Visual-Attention-Based Background Modeling for Detecting Infrequently Moving Objects," in IEEE Transactions on Circuits and Systems for Video Technology, vol. 27, no. 6, pp. 1208-1221, June 2017, doi: 10.1109/TCSVT.2016.2527258.

[78] I. Chelly, V. Winter, D. Litvak, D. Rosen and O. Freifeld, "JA-POLS: A Moving-Camera Background Model via Joint Alignment and Partially-Overlapping Local Subspaces," 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, 2020, pp. 12582-12591, doi: 10.1109/CVPR42600.2020.01260.

[79] X. Zhang, C. Zhu, S. Wang, Y. Liu and M. Ye, "A Bayesian Approach to Camouflaged Moving Object Detection," in IEEE Transactions on Circuits and Systems for Video Technology, vol. 27, no. 9, pp. 2001-2013, Sept. 2017, doi: 10.1109/TCSVT.2016.2555719.

[80] Kumar S. Ray and Soma Chakraborty, "Object detection by spatio-temporal analysis and tracking of the detected objects in a video with variable background", Journal of Visual Communication and Image Representation, vol. 58, 2019, pp. 662-674,ISSN 1047-3203.

[81] Subrata Kumar Mohanty and Suvendu Rup, “An adaptive background modeling for foreground detection using spatio-temporal features”, Multimedia Tools and Applications, vol. 80, 2021, https://doi.org/10.1007/s11042-020-09552-8.

[82] Caroline Silva, Thierry Bouwmans and Carl Frélicot, "Superpixel-based online wagging one-class ensemble for feature selection in foreground/background separation", Pattern Recognition Letters, vol. 100, 2017, pp. 144-151, ISSN 0167-8655.

[83] Y. Chen, Z. Sun and K. Lam, "An Effective Subsuperpixel-Based Approach for Background Subtraction," in IEEE Transactions on Industrial Electronics, vol. 67, no. 1, pp. 601-609, Jan. 2020, doi: 10.1109/TIE.2019.2893824.

[84] S. Javed, A. Mahmood, T. Bouwmans and S. K. Jung, "Background–Foreground Modeling Based on Spatiotemporal Sparse Subspace

[85] Clustering," in IEEE Transactions on Image Processing, vol. 26, no. 12, pp. 5840-5854, Dec. 2017, doi: 10.1109/TIP.2017.2746268.

[86] Israr Ul Haq, Keisuke Fujii and Yoshinobu Kawahara, "Dynamic mode decomposition via dictionary learning for foreground modeling in videos", Computer Vision and Image Understanding, vol. 199, 2020, 103022, ISSN 1077-3142.

[87] Shahidha Banu S. and Maheswari N., "An effective foreground segmentation using adaptive region based background modelling", Information Discovery and Delivery, Vol. 48 No. 1, pp. 23-34, 2020.

[88] Achraf Djerida, Zhonghua Zhao and Jiankang Zhao, "Robust background generation based on an effective frames selection method and an efficient background estimation procedure (FSBE)", Signal Processing: Image Communication, vol. 78, 2019, pp. 21-31,ISSN 0923-5965.

[89] S. Li, D. Florencio, W. Li, Y. Zhao and C. Cook, "A Fusion Framework for Camouflaged Moving Foreground Detection in the Wavelet Domain," in IEEE Transactions on Image Processing, vol. 27, no. 8, pp. 3918-3930, Aug. 2018, doi: 10.1109/TIP.2018.2828329.

[90] Suman Kumar Choudhury, Pankaj Kumar Sa, Kim-Kwang Raymond Choo and Sambit Bakshi, “Segmenting foreground objects in a multi-modal background using modified Z-score”, J Ambient Intell Human Comput , 2017, DOI 10.1007/s12652-017-0480-x.

[91] S. Jeeva and M. Sivabalakrishnan, “Twin background model for foreground detection in video sequence”, Cluster Comput, vol. 22, 11659–11668 ,2019. https://doi.org/10.1007/s10586-017-1446-7.

[92] Fakhri Alam Khan, Muhammad Nawaz, Muhammad Imran, Arif Ur Rahman and Fawad Qayum, ” Foreground detection using motion histogram threshold algorithm in high-resolution large datasets”, Multimedia Systems, 2020, https://doi.org/10.1007/s00530-020-00676-3.

[93] R. Kalsotra and S. Arora, "A Comprehensive Survey of Video Datasets for Background Subtraction," in IEEE Access, vol. 7, pp. 59143-59171, 2019, doi: 10.1109/ACCESS.2019.2914961.

[94] T. Bouwmans, “Background subtraction for visual surveillance:A fuzzy approach”, in: Handbook on Soft Computing for Video Surveillance, vol. 5, Taylor and Francis Group, 2012.

[95] N. Goyette, P.-M. Jodoin, F. Porikli, J. Konrad, and P. Ishwar, “changedetection.net: A new change detection benchmark dataset”, in Proc. IEEE Workshop on Change Detection (CDW-2012) at CVPR-2012, Providence, RI, 16-21 Jun., 2012.

[96] E. J. Candès, X. Li, Y. Ma, and J. Wright, ``Robust principal component analysis?'' J. ACM, vol. 58, no. 3, p. 11, May 2011.

[97] N. Halko, P.G. Martinsson and J. A. Tropp, “Finding Structure With Randomness: Probabilistic Algorithms For Constructing approximate Matrix Decompositions”, arXiv:0909.4061 [math NA], 2010