Received: July 24, 2022. Revised: September 4, 2022.

International Journal of
Intelligent Engineering & Systems

http://www.inass.org/

A Hybrid Approach for Moving Object Detection Using TV-L1 Features and RPCA–MOG

Sudhir Dagar[1]* **Geeta Nijhawan[2]**

[1]*Department of Electronics and Communication Engineering,*
Chaudhary Devi Lal State Institute of Engineering & Technology, India
[2]*Department of Electronics and Communication Engineering,*
Manav Rachna International Institute of Research and Studies, India
* Corresponding author's Email: sudhirdagar2002@yahoo.com

Abstract: Robust Principal Component Analysis (RPCA) has recently been active everywhere for dimension reduction in image processing, display and pattern recognition. Methods based on low-rank sparse representations, which make some specific significant assumptions, have recently received a lot of attention in background modelling. Meanwhile, a powerful analysis framework is needed to handle background areas or foreground motion at various scales. this paper presents a hybrid approach along with total variation L1 (TV-L1) features and reproductive RPCA model in low rank background subtraction modelling and sparse matrix with mixture of Gaussians (MoG) as foreground modelling. The hybrid structure with TV-L1 features imposes a hierarchical RPCA on the singular values of the low-rank component and MOG sparsity indicators. The proposed work was evaluated on the CDnet2014 (ChangeDetection.net) dataset, obtained result as accuracy was 92.9%, 86.7% 95.7% for Highway, Escalator, and Indoor respectively. The proposed method is compared with traditional methods and obtained relative reconstruction error is 0.01529 as a lower side.

Keywords: MOG, RPCA, LAPLACIAN, GAUSSIAN, TV-L1.

1. Introduction

When we have video processing it is necessary to apply different techniques separated in steps to separate the pixels that contain information (objects or persons movement) of static objects (known as background), among these we find the background subtraction, which is a technique widely used in the processing of video for the segmentation (separate the foreground or the front of the background), obtaining as result, information-rich pixels in which different techniques can be applied. Due to the nature of this technique, it has been widely adopted in areas such as detection of movement, multimedia applications, surveillance by means of videos. and all are cases it is necessary to detect objects in motion in a scene. Due to some inconveniences that arise in videos like poor signal, noise caused by low camera resolution, compression techniques of a noisy environment, it leads to present numerous false positives.

Background subtraction is a step that aims to identify the regions of the image that have moving objects, separating them from the static areas (background). In moving object focused applications, it is common to use modeling that allow updating the background for subsequent subtraction between the video and background images at the matrix level. This update ensures that pixel values background (Red, Green, And Blue (RGB) pattern, for example) come closer to the "real" pixel values in the moment of subtraction, as changes in brightness change these values with the time, thus interfering with the subtraction.

Despite this advantage, there are several challenges associated with background modeling in the application of traffic and one of them would be the fact that if a moving object stops moving by an

Received: July 24, 2022. Revised: September 4, 2022.

over a long period of time it will be accepted as part of the background [1].

The application of image processing and computer vision techniques for the analysis of stream video sequences is able to provide data such as volume, speed, coordinates and object classification, and may be relatively successful in these determinations in different applications and environments [2-4].

Among the strategies most commonly found in the literature for identifying and tracking of objects applied to the vehicular flow, the strategy based on features (feature-based approach) that is capable of tracking even with partial occlusions, and can be used in different lighting conditions [5, 6]. This strategy has several steps, however, like most other methods, it usually relies on background subtraction to subsequent feature detection and [7, 8]. Background subtraction is a step that aims to identify the regions of the image that have moving objects, separating them from the static areas (background).

Following a moving subject after a while is a test issue in video processing. Analysts have developed many video management programs to distinguish the position of objects in each frame (that is, video lines) in a given array, thereby solving the problem of temporal grouping of addresses. In the unlikely event that this procedure is understood for each image in a specific group, and we essentially relate the position of a specific object, tracking is considered good practice [9]. However, existing surveys do not consider tracking because each site is identified independently of the others. The main reason is the lack of time data during the tracking process. The focal point in a simple drawing that an object crosses on the period scale is the direction in which the object is moving. This methodology can be extended to object states. These states are often stored in a state vector, and each step of this vector contains an approximate of a certain parameter at a certain time. Thus, the usual type of tracking process for a moving object becomes a procedure for updating the previous state. The state vector trajectory can consist of several basic points such as position, speed, and increase in speed, size, shape, shadow, area, etc. In a capture-based approach to object detection and tracking, the most common method for recognizing moving objects is background subtraction (BS). The main idea behind this methodology is to approximate an adequate representation (i.e., the background image shows it) of a given scene based on pixel variance. You can also select objects around the edges of the current video by subtracting the edges of the current video from the background image.

2. Literature review

Background subtraction system (BGS) has become a hot topic in the field of Computer Vision in recent years as a method for detecting moving objects. BGS is commonly used to track moving objects in video surveillance systems. The authors of [10] proposed a new method based on a Hybrid modelling strategy that combines two models, namely Robust Principal Component Analysis (RPCA) employing Principal Component Pursuit (PCP) and randomised Singular Value Decomposition (rSVD), to perform the aforementioned task for background subtraction. The proposed method triggers with false positive rate due to singular value decomposition (SVD). Further [11] proposed the architecture of dynamic background subtraction with limitation of RPCA where Low rank matrix cannot capture dynamic back ground. The author further introduced masked RPCA to process backdrops with shifting textures. A mask that roughly classifies moving objects and backgrounds is generated using a first-order Markov random field. The propose algorithm is sensitive to data frame and size of video input.

When considering its background modeling characteristic, the segmentation techniques into two large groups: non-recursive and recursive techniques [12]. Non-recursive techniques approximate the fund from a sampling of images stored in a buffer. As an example of them, there is the frame-by-frame subtraction (which uses frame t-1 to subtract the current frame) and the median filter, which makes a sampling for x amount of previous frames and calculates the median of the values of each pixel in order to generate a representation of the background.

Recursive techniques, in turn, update the background modelling at each new frame, using mathematical models. Among these techniques, those using Gaussian mixture model (GMM) [13]. A problem found in recursive methods such as GMM is that the addition of each new frame with objects that move slowly, or stopped, can generate a modelling erroneous background when adhering the object to the background [14].

Recently, Gaussian parameter compression approaches [15] and a combination of Mixture of Gaussian (MoG) and compressed sensing (CS) [16] have been utilised. Furthermore, object detection in UAV-sourced movies is being done using parallel implementation methodologies for algorithms [14]. Background subtraction, as previously stated, is a frequently used approach in surveillance movies, object tracking and identification, traffic or crowd monitoring, and other applications where the primary goal is to separate moving objects, i.e., foreground from stationary background [17].

There are a lot of recent research works have published on background subtraction in moving object detection. A Fully residual convolutional neural network (FR- CNN) is used in [18] for background subtraction where non-handcrafted feature extraction is utilized. The authors of [19] presented fast background subtraction with adaptive block learning (FBS-ABL) algorithm for real-time moving object detection. The major drawback of this techniques is slower background update. The authors of [20] presented an approach for moving object detection using spectral dual mode background subtraction (SDMBS). The authors of [21] presented a background subtraction model based on parallel vision and Bayesian generative adversarial networks (BSPVBGANs). One significant disadvantage of this method is that it cannot overcome the restriction of training images, and hence the technique cannot satisfy the requirement for real-time incremental training. However, due to challenges such as shadow, fluctuating illumination, occlusion, background motion, and camera movement, detecting moving objects in films or other applications remains a difficult task as well as a variety of other uncertainties such as atmospheric disturbances or noise, object overlapping outliers, and so on [22]. Statistical models are the most effective at overcoming these obstacles. Nonparametric and parametric statistical models are also possible. The Kernel Density Estimation (KDE) and Eigenvalue Decomposition are two of the most often used non-parametric approaches, although it need a lot of memory and have a high computational complexity [15].

Matrix decomposition methods, such as Robust Principal Component Analysis (RPCA), have recently emerged as a cost-effective paradigm for background reduction. These methods are designed to break down a matrix into low-rank (background) and sparse (foreground) components.

However, as the magnitude of the input data grows and there are no sparsity constraints, these approaches demonstrate poor performance in particular cases, since they are unable to manage the real-time issues, resulting in inaccurate foreground areas [23].

Based on Robust Principal Component Analysis, a new background subtraction model with logarithm rank function and structural sparsity was presented [23] to deal with the dynamic background and slow moving objects (RPCA). The segmentation and index trees were employed in this model to dynamically process the foreground, improving appearance similarity and spatial continuity between the pixels. The foreground structure lacks with the shadow removal and accurate detection of light.

Although RPCA provides a solid framework for background subtraction, due to its batch optimization, it still has a high computational cost and large memory requirements. To address this problem, TV-L1-RPCA is created, which can process such high-dimensional data using stochastic methods.

However, for the background subtraction job, the RPCA is used with imprecise assumptions for both static and dynamic variables. Components that match the video's background and respectively, foreground and background As a result, it is vital to think about the finer understanding to deal with the ever-increasing issues considering the drawbacks of RPCA-based techniques,

We propose a TV-L1 feature and rank-1 regularised RPCA in this study as background subtraction model, utilising both static and dynamic data. Constraint-based RPCA-based video representation to better encode the static component (i.e., background latent) in a video sequence, knowledge) of temporal structures is required. Second, with RPCA, we refine the dynamic component even further.

It can be thought of as the spatial-temporal superposition of smooth video foreground and sparse noise.

Total variation and L1 norm words can be used to express this.

The article presents the following materials:

1. Hybrid sparse modelling using RPCA-MOG.
2. Concatenated texture information based on the TV-L1 features for RPCA, with the low rank representing the background pattern.
3. Improved methods of extracting texture features for detecting distant objects are being developed.

The method proposed in this work consists of the following steps: (i) Definition of the sequence of video for analysis; (ii) Processing the reference segmentation (ground truth); (iii) Presentation of the proposed subtraction method; (iv) calibrationand execution of algorithms for segmentation and; (v) Comparative analysis of the results found.

3. Materials and methods

3.1 Mixture of gaussians (MoG)

It is a background and edge segmentation algorithm based on GMM. An important feature of this algorithm is that it selects the appropriate number of Gaussian distributions for each pixel. They offer

Received: July 24, 2022. Revised: September 4, 2022.

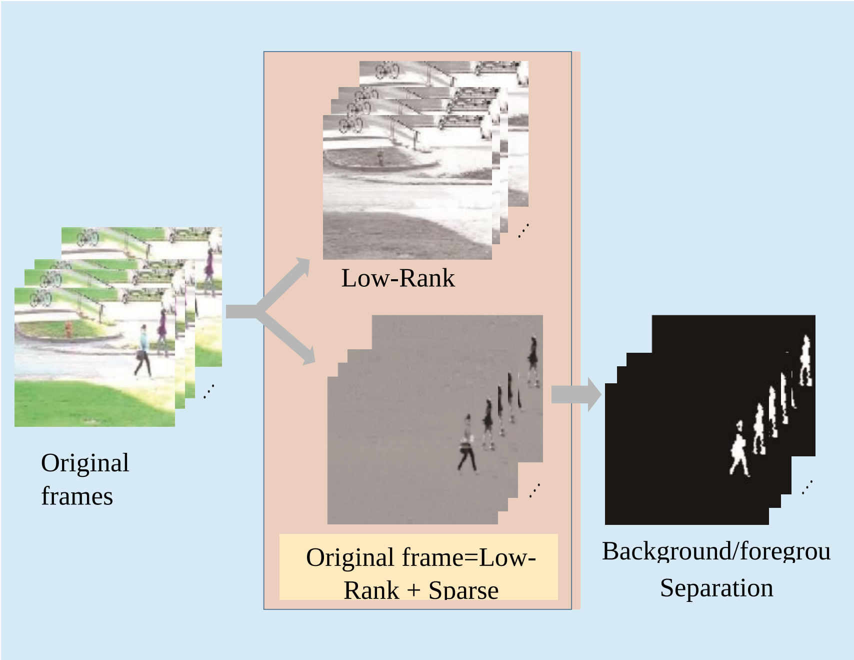

Figure. 1 Video segmentation process based on low-rank and sparse matrix decomposition

better adaptation to changing scenes due to lighting, etc.

In practice, the lighting in a scene can change gradually (weather or outdoor weather conditions) or suddenly (turn off the lights when shooting indoors). When you can remove an object from the scene. After we are done adapting to the changes, we can update the training set to add new objects and remove noise.

We chose a sensible period of time T and in the time t we have $X_T = \{x^{(t)} \dots \dots x^{(t-T)}\}$. For each new show, we update the training set XT and we went back to calculate $\hat{P}(\vec{X}|x_T, BG + FG)$. Without embargo, among the recent samples could be have values that associated to front objects and we must symbolize this approximation as $\hat{P}(\overrightarrow{X^{x(t)}}|x_T, BG + FG)$.

We utilize GMM having M components:

$$\hat{P}(\vec{X}|x_T, BG + FG) = \sum_{m=1}^{M} \hat{\pi}_m \; N(\vec{x}; \hat{\mu}_m, \hat{\sigma}_m^2 I) \quad (1)$$

Where $\hat{\mu}_1 \dots\dots \hat{\mu}_m$ are approximate average and $\hat{\sigma}_1^2 \dots \dots \dots \dots \dots \dots \hat{\sigma}_m^2$ are the approximation of variance that define the Gaussian components.Covariance matrices are supposed to be diagonals and the identity matrix I have appropriate dimensions. The mixture of weights is symbolized by $\hat{\pi}_m$ are non-negative and increase to 1. Assumed a novel demonstration $\vec{X}^{(t)}$ in time t the execution of the recursive update is given by following equations:

$$\hat{\pi}_m \leftarrow \hat{\pi}_m + \alpha(0_m^{(t)} - \hat{\pi}_m \quad (2)$$

$$\hat{\mu}_m \leftarrow \hat{\mu}_m + 0_m^{(t)} \left(\frac{\alpha}{\hat{\pi}_m}\right) \delta_m \quad (3)$$

$$\hat{\sigma}^2 \leftarrow \hat{\sigma}^2 + 0_m^{(t)} \left(\frac{\alpha}{\hat{\pi}_m}\right) (\delta_m^T \delta_m - \hat{\sigma}_m^2) \quad (4)$$

3.2 Robust principal component analysis (RPCA)

This section describes the fast principal component analysis method, focusing on the incremental PCP algorithm [24, 25] (which, in turn, is based on [26], which is used to improve the classification modeling of background video is a ubiquitous pre-processing step in many computer vision applications used to detect moving objects in digital video.

In particular, PCP is offered in [27] as a non-convex optimization problem defined by Eq. (5):

$$\arg\min_{L,S} rank(L) + \lambda \|S\|_0 \;\mathrm{s.t.}\; D = L + S \quad (5)$$

Where $D \in \mathbb{R}^{m \times n}$ is the observation video of n images, each with size $m = N_r \times N_c \times N_d$ (row, column and depth or channel), $L \in \mathbb{R}^{m \times n}$ is a lower-order matrix representing the background of the plane, and $S \in \mathbb{R}^{m \times n}$ is a sparse matrix representing the foreground (moving object).

However, RPCA still has clear limitations. As in the Eq. (5), the L1 norm is used for the characteristic S, which is ideal only for Laplace noise. Although the

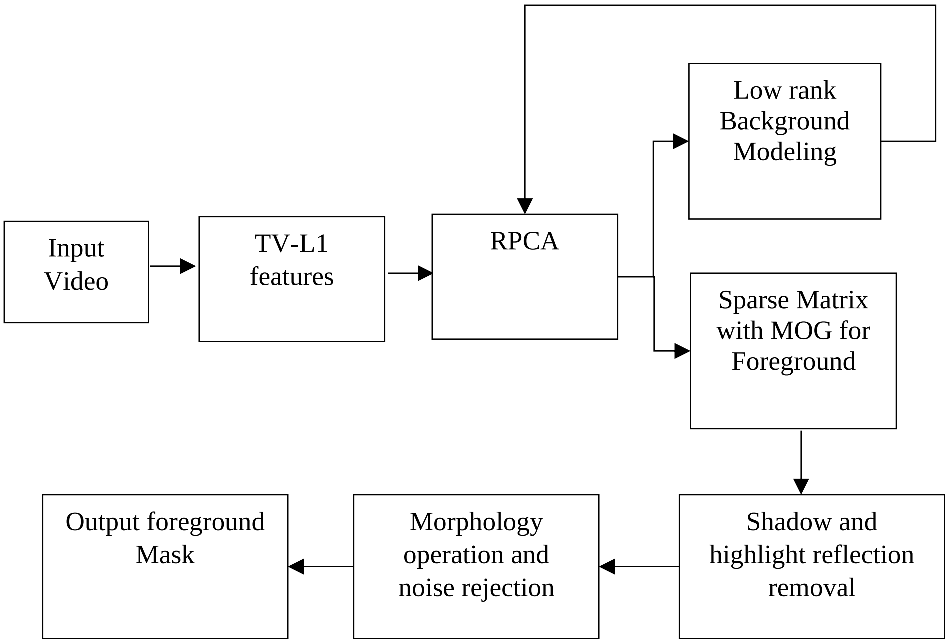

Figure. 2 Proposed diagram of background subtraction and foreground detection

L1 norm performs better than the sparse noise L2 norm, the actual noise is usually neither Gaussian nor Laplacean but it has a much more complex statistical structure.

4. Proposed method

This article introduces a new hybrid approach to RPCA that can be adapted for more complex noises.

We formulate the problem as a generative model in a Bayesian framework and model the noise in the data as a mixture of Gaussian (MoG). We then use the variation inference method to get the posterior. Since MoG is a universal approximation of any continuous probability distribution, the proposed MoG-RPCA approach can be adapted to a much wider range of real noise than existing RPCA techniques.

The Fig. 2 explains the process of background modelling and subtraction i.e., foreground detection. The first step comprises TV-L1 texture features extracted from input video frames. After that RPCA is applied on to get low rank and sparse matrix. Low rank represents the background of the moving object and sparse matrix represents the foreground information. Further sparse matrix is modelled with MOG (Mixture of Gaussian) to represent the foreground. The MOG2 function, using n number of initial frames, so as to give the starting point for the subsequent subtraction process. The sampling of these frames must take place in a period flow, so that the function is able to do a good modelling of the foreground. The dimensions of the regions are input data in the proposed algorithm. The default template adopted for this analysis was square sections of 50×50 pixels. This means that the algorithm will section each frame of the binary images, of the two methods, into squares of 50 pixels sideways to analyse. In each section (squares) it is checked if the MOG2 function, or the subtraction made by thresholding, detects the presence of foreground (vehicles or noise, white pixels). If so happens, an analysis is made to see if the foreground rates present in the two tested binary images (MOG2 and thresholding) are within the considered values acceptable for each (noise), and this value is also an input parameter in the algorithm. If in the analysed area there is the presence of a foreground above these values, in any of the subtraction strategies, the background is not updated, otherwise the background in the respective area is updated, receiving the section of the video frame referring to area of the analysed binary images. We use morphological operations, where the objective here will be to carry out an erosion followed by a dilation to eliminate this white noise on the black areas. Erosion will disappear in isolated white spots and a later one dilation will take care of restoring the change caused by erosion in the zones whites that survived. Finally, after morphological operation foreground mask is created to get the differentiation between the background and objects of interest.

4.1 Feature extraction using TV-L1 descriptors

The TV-L1 feature in images is classically obtained by using a pair of low-pass and high-pass filters to the image I by the following minimization:

$$TV - L1\ \ Feature = \min_u \left\{ \sigma^4 \int |Du|^2 + \|I - u\|_{H^{-1}}^2 \right\} \quad (6)$$

Where u represents the TV-L1 part of the image I.

Eq. (2) defines the data matrix M obtained by RPCA, $f_t(x, y)$ is the function for TV-L1 texture descriptor extracted from each image. L and S matrices are derived from matrix M by RPCA.

$$M(x, y) = \beta f_t(x, y) \quad (7)$$

4.2 Modelling of sparse for foreground

We assume that each $S_{i,j}$ in S has MOG distribution as expressed in following equation.

$$S_{i,j} \sim \sum_{m=1}^{M} \pi_m \mathrm{G}(\mathrm{s_{i,j}}\, |\mu_m, \mathrm{T}_m^{-1}) \quad (8)$$

Where, "~" is an approximation operator, π_m is mixing proportion with $\pi_m > 0$ and $\sum_{m=1}^{M} \pi_m = 1$,M is the Gaussian components number and $G(s\,|\mu, T^{-1})$ is Gaussian distribution with mean $\mu$and precision T.

Eq. (8) can be consistently stated as a two-level reproductive model by presenting the indicator variable $Z_{ijm}s$:

$$S_{i,j} \sim \prod_{m=1}^{M} G(s_{i,j}\, |\mu_m, \mathrm{T}_m^{-1})^{z_{ijm}} \quad (9)$$

Where,
$z_{ij} = (z_{ij1} \ldots\ldots\ldots\ldots z_{ijm}) \in \{0,1\}^M$ and $\sum_{m=1}^{M} z_{ij} = 1$.

To compute the Bayesian model, we introduce a priori mates to the parameters of the Gaussian component, $\mu_m s$, $\tau_m s$ and the mixing proportions,π as:

$$\mu_m, \tau_m \sim G\left(\frac{\mu_m}{\mu_0}\right), (\beta_0 \tau_m)^{-1} \gamma \left(\frac{\tau_m}{c_0}\right), d_0),$$
$$\pi \sim Dir(\frac{\pi}{\alpha_0}) \quad (10)$$

Where $\gamma\left(\frac{\tau_m}{c_0}\right), d_0$ is the gamma distribution with parameters $c_0 and d_0$ and $Dir(\frac{\pi}{\alpha_0})$ denotes the Dirichlet distribution parameterized by$\alpha_{01} \ldots\ldots\ldots\ldots, \alpha_{0m}$.

4.3 Low-rank component modelling

A simple way to model the lower-rank L component is to apply Laplace a priori to a single L value. In this article, we use ARD to model low-level components due to its high speed and good scalability. L can be formulated as $L \in \mathbb{R}^{m \times n}$ with rank $l \leq (m, n)$ as the product of $U \in \mathbb{R}^{m \times R}$ and $V \in \mathbb{R}^{n \times R}$.

$$L = VU^T = \sum_{r=1}^{R} u_r v_r^T \quad (11)$$

Where $R > l$,and $u_r, v_r$is ther^{th} column of $U(V)$. To achieve the lack of U and V columns, some U and V columns are close to zero. In this way, low L-rank can be guaranteed. This goal can be achieved by applying a priori to U and V.

$$u_r \sim G(u_r|0, \tau_r^{-1} I_m), v_r \sim G(v_r|0, \tau_r^{-1} I_n) \quad (12)$$

Where I_m denotes $m \times m$ identity matrix. The conjugate prior on each precision variable τ_r is:

$$\tau_r \sim \gamma(\tau_r|a_0, b_0) \quad (13)$$

Note that each pair of $u_r, v_r of U, V$ columns from U, V has the same sparsity outline, which is considered by a generalized precision variable τ_r. It has been confirmed that such simulations can yield highly accurate values of some $\tau_r s$and hence low-rank order L estimates. Combine the Eqs. (8) to (13) together we can build a complete RPCA Bayesian model with MoG noise called MoG-RPCA. The goal is to get posteriori all involved variables:

$$p(U, V, \mathcal{Z}, \mu, \tau, \pi, \gamma Y) \quad (14)$$

Where, $\mathcal{Z} = \{Z_{ij}\}, \mu = (\mu_1 \ldots \ldots \ldots \ldots \ldots \mu_m)$.

5. Simulation and results

In this paper, we compare four background subtraction techniques using the CDNet dataset [28]. The database contains 1700 frames with a resolution of 320×240, the first 100 frames are used for background initialization, and the rest of the images are used for background updates for object detection. Each frame has a separate ground truth. Four techniques of background subtraction are compared. Following parameters are used for evaluation of the proposed techniques.

Received: July 24, 2022. Revised: September 4, 2022.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (15)$$

$$Precision = \frac{TP}{TP+FP} \quad (16)$$

$$Sensitivity = \frac{TP}{TP+FN} \quad (17)$$

$$Specificity = \frac{TN}{TN+FN} \quad (18)$$

$$ErrorRate = \frac{FP+FN}{TP+TN+FP+FN} \quad (19)$$

$$FalsePositiveRate(FPR) = \frac{FP}{FP+TN} \quad (20)$$

$$F - Score = \frac{2TP}{2TP+FP+FN} \quad (21)$$

$$Matthews\ Correlation\ Coefficient(MCC) = \frac{(TP\times TN)-(FP\times FN)}{\sqrt{(TP+FN)(TP+FP)(TN+FN)(TN+FP)}} \quad (21)$$

$$Kappa\ Statistics = \frac{2(TP\times TN - FN\times FP)}{(TP+FP)\times(FP+TN)+(TN+FN)\times(FN+TN)} \quad (22)$$

To obtain these parameters it is necessary to compare the two images and count the number of pixels according to the following definitions:
TP (True Positive): ground truth objects that were correctly classified in fgmask as objects;
TN (True Negative): ground truth background correctly classified in fgmask as background;
FP (False Positive): pixels of objects in fgmask misclassified because represent background on ground truth;
FN (False Negative): background pixels in fgmask misclassified, as they are objects in ground truth;

The dataset is convoyed by precise ground-truth segmentation and annotations of changes / motion zones for each video frame. Similarity measures is performed by using Mahalanobis distance.

$$D^2 = (x-m)^T C^{-1}(x-m) \quad (23)$$

Where D^2=Mahalanobis distance
X=vector data
M=vector of mean values of independent variables
C^{-1} =Inverse covariance matrix of independent variables
T=Indicates vector should be transposed

5.1 Results for object identification in videos

The objective of this work is to identify four classes of objects, being them, escalator, highway moving car cyclists, moving person, Fig. 2 illustrates the steps of the proposed methodology. Initially, for the development of the activities performed in this research, the objects that move in each frame of the

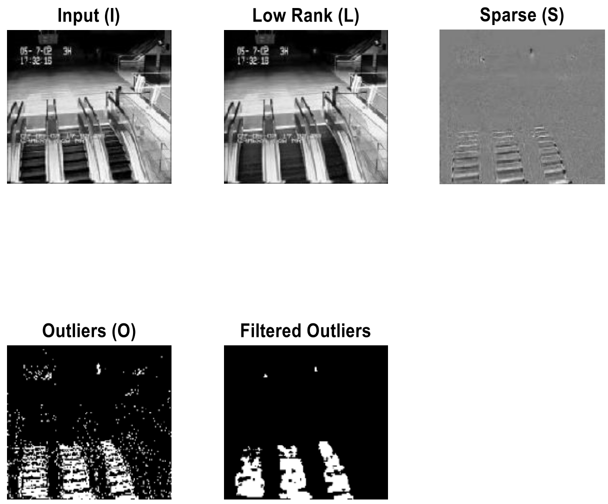

Figure. 3 Simulation results performed on escalator video

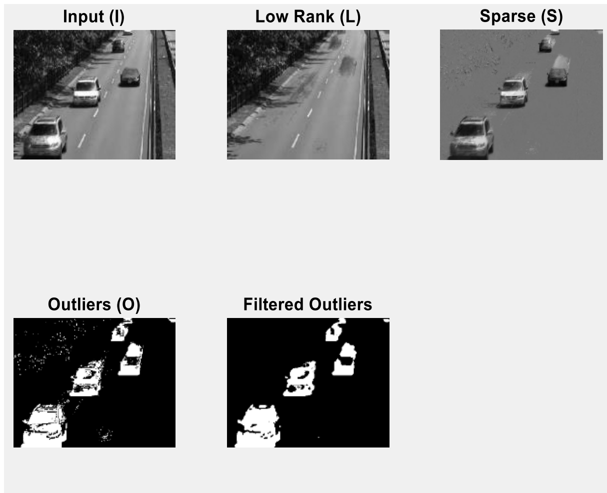

Figure. 4 Simulation results performed on highway video

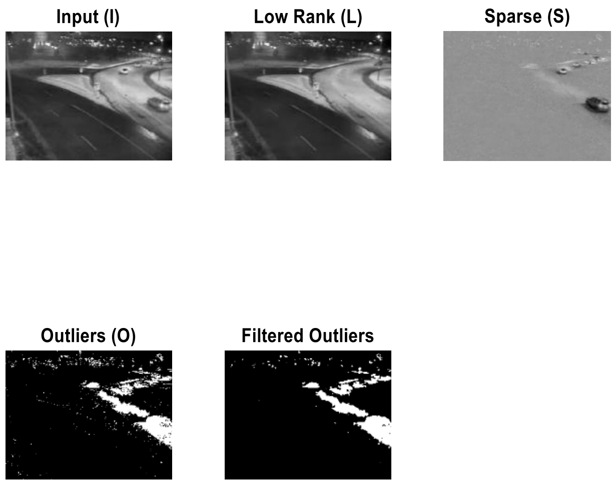

Figure. 5 Simulation results performed on highway-2 video

Received: July 24, 2022. Revised: September 4, 2022.

videos used, through the method of plane subtraction called Gaussian Mix. Due to the Gaussian Blend method not precisely delimit the moving objects, we apply a post-processing through the mathematical morphology operations called opening and closing. These operations are applied over the binary mask produced by background subtraction with the target to eliminate small components (noise), smooth the outline of larger components, and fill holes and slits present in the detection mask.

5.2 Model performance analysis

The performance of the investigated models was measured by their representativeness in identify the objects in the image. For this, after generating the background image of each model, a subtraction of the current frame with the modelled background was performed. It is difference image, called fgmask (foreground mask) represents the objects found by the detection system. This analysis method is the most common to compare segmentations by Mahalanobis distance.

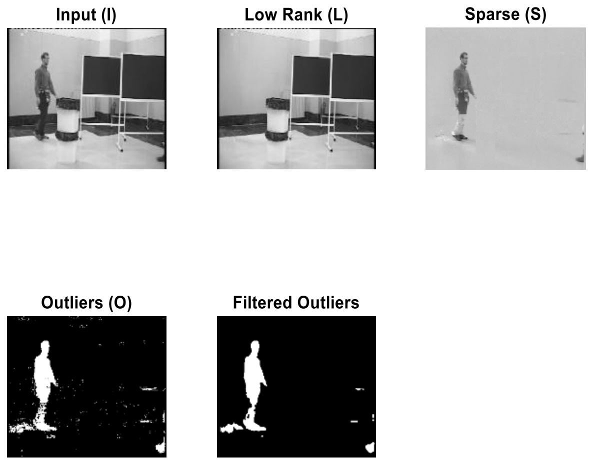

Figure. 6 Simulation results performed on indoor video

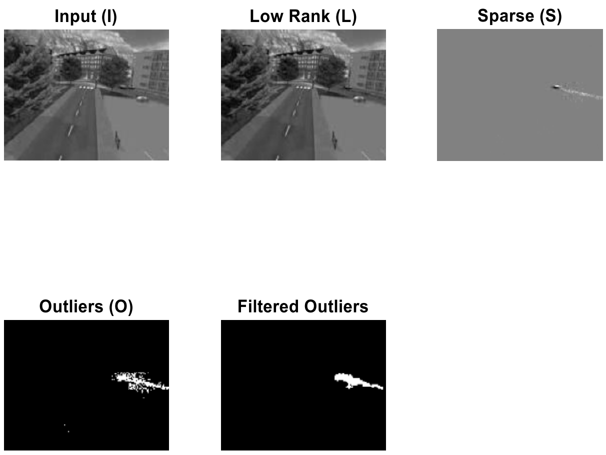

Figure. 7 Simulation results performed on highway-3 video

Table 1. Relative reconstruction error of different methods on highway video

Methods	Relative reconstruction error
GMM	0.23565
RPCA	0.1056
RPCA-MOG	0.019104
Cultural-RPCA-MOG	0.01529

Table 2. Frame-wise comparison on highway video with ground truth and segmented outcome

Highway	Accuracy	Precision	Sensitivity	F-Score
Frame 50	92.9%	87.5%	100%	93.3%
Frame 100	92.3%	86.7%	100%	92.9%
Frame 150	94.4%	90%	100%	94.7%
Frame 200	95.2%	91.3%	100%	95.5%

Table 3. Frame-wise comparison on escalator video with ground truth and segmented outcome

Escalator	Accuracy	Precision	Sensitivity	F-Score
Frame 50	86.7%	84.2%	88.9%	86.5%
Frame 100	90%	91.4%	88.9%	90.1%
Frame 150	91.5%	94.4%	89.3%	91.8%
Frame 200	90.39%	100%	84.78%	91.8%

Table 4. Frame-wise comparison on indoor video with ground truth and segmented outcome

Indoor	Accuracy	Precision	Sensitivity	F-Score
Frame 50	95.7%	100 %	92.9%	96.3%
Frame 100	96.1%	100%	95.7%	96.2%
Frame 150	92.2%	96.5%	89.9%	92.7%
Frame 200	89.4%	100%	82.1%	90.2%

We contrast our suggested approach with various cutting-edge algorithms that are listed on www.changedetection.net. We provide comprehensive overall and per-category F-Measure comparisons in Table 5. The online evaluation server posts the results. The top-ranked algorithms are FR-CNN [18], FBS-ABL [19], SDMBS [20], and BSPVBGANs [21]. Deep learning-based supervised approaches include FR-CNN, FBS-ABL, and

Received: July 24, 2022. Revised: September 4, 2022.

Table 5. Comparative analysis with previous research works

Method	F-Score							
	Baseline	Camera Jitter	Dynamic Background	Intermittent	Shadow	Thermal	Bad Weather	Low Frame Rate
FR-CNN [18]	0.9711	0.9614	0.9656	0.9112	0.9593	0.9118	0.9634	0.8508
FBS-ABL [19]	0.8910	0.8046	0.7958	0.7861	0.9143	0.6394	0.7449	0.6616
SDMBS [20]	0.9114	0.5371	0.8333	0.7639	NA	0.7673	0.8624	0.6585
BSPVBGANs [21]	0.9730	0.9890	0.9780	0.9830	0.9360	0.9760	0.9640	0.8630
Proposed	**0.9780**	**0.9968**	**0.9956**	**0.9961**	**0.9957**	**0.9949**	**0.9889**	**0.9257**

SDMBS. First off, it is clear that the top two approaches are all deep we learning-based; FR-CNN [18] even surpasses for the baseline input, the proposed hybrid algorithm with an F-Measure of 0.9780. However, as we have indicated, background subtraction algorithms for video surveillance should be unsupervised from an applications standpoint. It is debatable if these algorithms can be used in situations like the ones we think about in this paper. The suggested proposed framework, which provides state-of-the-art performance among all unsupervised approaches, actually outperforms some deep learning-based methods, including BSPVBGANs [21], as can be seen in the second observation. It is important to note that proposed model is a very adaptable framework, making it simple to change the components as needed.

6. Conclusion

We propose the noise modeling of the new RPCA method as a MoG distribution in a Bayesian structure. Compared to the current RPCA method, which assumes a certain distribution of noise (such as sparse or Gaussian noise) in the data. The proposed method demonstrates clear advantages over the previous method in terms of its ability to accurately reconstruct low-level structures and carefully extract multimodal noise patterns from video data observed in various scenarios where the proposed hybrid method has moderate accuracy of 93% on multiple videos. The relative error of recovery by the proposed method is 0.015 less than by other methods. It was concluded that the proposal had produced satisfactory results.

Conflicts of Interest

The authors declare no conflict of interest.

Author Contributions

This paper conceptualization, software simulation, verification of results and original draft preparation has been done by Sudhir Dagar. The supervision and final approval have been done by Dr.Geeta Nijhawan

References

[1] J. Wei, J. Zhao, Y. Zhao, and Z. Zhao, "Unsupervised Anomaly Detection for Traffic Surveillance based on Background Modeling", In: *Proc. of International Conf. on Computer Vision and Pattern Recognition Workshops*, pp. 129-136, 2018.

[2] S. H. Supangkat, F. Hidayat, I. A. Dahlan, and F. Hamami, "The Implementation of Traffic Analytics Using Deep Learning and Big Data Technology with Garuda Smart City Framework", In: *Proc. of Global Conf. on Consumer Electronics (GCCE)*, pp. 883-887, 2019.

[3] A. Chetouane, S. Mabrouk, I. Jemili, and M. Mosbah, "Vision-based Vehicle Detection for Road Traffic Congestion Classification", *Concurrency and Computation: Practice and Experience*, Vol. 34, No. 7, p. e5983, 2022.

[4] J. Shang, Y. Liu, H. Zhou, and M. Wang, "Moving Object Properties-based Video Saliency Detection", *Journal of Electronic Imaging*, Vol. 30, No. 2, p. 023005, 2021.

[5] R. D. Rohan, Z. A. Patel, S. C. Yadavannavar, C. Sujata, and U. Mudengudi, "Image Segmentation and Geometric Feature Based Approach for Fast Video Summarization of Surveillance Videos", In: *Proc. of Computer Vision Applications, Communications in Computer and Information Science*, Vol. 1019, pp. 79-88, 2018.

[6] A. F. Abbas, U. U. Sheikh, F. T. A. Dhief, and M. N. H. Mohd, "A Comprehensive Review of Vehicle Detection using Computer Vision", *Telecommunication Computing Electronics and Control (TELKOMNIKA)*, Vol. 19, No. 3, pp. 838-850, 2021.

[7] Z. R. Mahayuddin and A. S. Saif, "A Comparative Study of Three Corner Feature based Moving Object Detection using Aerial Images", *Malaysian Journal of Computer Science*, Special Issue 3, pp. 25-33, 2019.

[8] Y. Q. Chen, Z. L. Sun, and K. M. Lam, "An Effective Subsuperpixel-based Approach for Background Subtraction", *IEEE Transactions on Industrial Electronics*, Vol. 67, No. 1, pp. 601-609, 2019.

[9] R. Padilla, S. L. Netto, and E. A. D. Silva, "A Survey on Performance Metrics for Object-Detection Algorithms", In: *Proc. of International Conf. on Systems, Signals and Image Processing (IWSSIP)*, pp. 237-242, 2020.

[10] S. Maurya and M. S. Choudhry, "Background Subtraction Using a Hybrid Modelling Based Technique", In: *Proc. of 8th International Conf. On Signal Processing and Integrated Networks (SPIN)*, pp. 527-531, 2021.

[11] H. Ahn and M. Kang, "Dynamic background Subtraction with Masked RPCA", *Signal, Image and Video Processing*, Vol. 15, No. 3, pp. 467-474, 2021.

[12] B. Azeez and F. Alizadeh, "Review and Classification of Trending Background Subtraction-based Object Detection Techniques", In: *Proc. of 6th International Engineering Conference (IEC) on Sustainable Technology and Development*, pp. 185-190, 2020.

[13] H. Bandi, D. Bertsimas, and R. Mazumder, "Learning a Mixture of Gaussians via Mixed-Integer Optimization", *INFORMS Journal on Optimization*, Vol. 1, No. 3, pp. 221-240, 2019.

[14] D. Jaiswal and P. Kumar, "Real-Time Implementation of Moving Object Detection in UAV Videos using GPUs", *Journal of Real-Time Image Processing (JRTIP)*, Vol. 17, No. 5, pp. 1301-1317, 2020.

[15] W. Munir, A. M. Siddiqui, M. Imran, I. Tauqir, N. Zulfiqar, W. Iqbal, and A. Ahmad, "Background Subtraction in Videos Using LRMF and CWM Algorithm", *Journal of Real-Time Image Processing (JRTIP)*, Vol. 18, No. 4, pp. 1195-1206, 2021.

[16] L. Mabrouk, S. Huet, D. Houzet, S. Belkouch, A. Hamzaoui, and Y. Zennayi, "Efficient Adaptive Load Balancing Approach for Compressive Background Subtraction Algorithm on Heterogeneous CPU–GPU Platforms", *Journal of Real-Time Image Processing (JRTIP)*, Vol. 17, No. 5, pp. 1567-1583, 2020.

[17] I. Draganov and R. Mironov, "Object Motion Detection in Video by Fusion of RPCA and NMF Decompositions", In: *Proc. of New Approaches for Multidimensional Signal Processing, Smart Innovation, Systems and Technologies*, Vol. 270, pp. 35-47, 2022.

[18] M. Vijayan, P. Raguraman, and R. Mohan, "A Fully Residual Convolutional Neural Network for Background Subtraction", *Pattern Recognition Letters*, Vol. 146, pp. 63-69, 2021.

[19] V. J. Montero, W. Y. Jung, and Y. J. Jeong, "Fast Background Subtraction with Adaptive Block Learning using Expectation Value Suitable for Real-Time Moving Object Detection", *Journal of Real-Time Image Processing*, Vol. 18, No. 3, pp. 967-981, 2021.

[20] A. M. A. Tawab, M. B. Abdelhalim, and S. E. D. Habib, "Moving Object Detection over Wireless Visual Sensor Networks using Spectral Dual Mode Background Subtraction", *International Journal of Advanced Computer Science and Applications*, Vol. 13, No. 1, pp. 514-523, 2022.

[21] W. Zheng, K. Wang, and F. Y. Wang, "A novel Background Subtraction Algorithm based on Parallel Vision and Bayesian GANs", *Neurocomputing*, Vol. 394, pp. 178-200, 2020.

[22] S. Maurya and M. S. Choudhry, "Concept of Hybrid Models in Background Subtraction: A Review of Recent Trends", In: *Proc. of Data Analytics and Management, Lecture Notes on Data Engineering and Communications Technologies*, Vol. 91, pp. 241-252, 2022.

[23] Y. Ban, L. Chen, and X. Wang, "Background Subtraction based on Logarithm Rank Function and Structured Sparsity", *Multimedia Tools and Applications*, Vol. 81, pp. 1-17, 2022.

[24] W. Xiao, X. Huang, F. He, J. Silva, S. Emrani, and A. Chaudhuri, "Online robust principal component analysis with change point detection", *IEEE Transactions on Multimedia*, Vol. 22, No. 1, pp. 59-68, 2019.

[25] G. Han, J. Wang, and X. Cai, "Background Subtraction based on Modified Online Robust Principal Component Analysis", *International Journal of Machine Learning and Cybernetics*, Vol. 8, No. 6, pp. 1839-1852, 2017.

[26] A. Djerida, Z. Zhao, and J. Zhao, "Background Subtraction in Dynamic Scenes using the Dynamic Principal Component Analysis", *IET Image Processing*, Vol. 14, No. 2, pp. 245-255, 2020.

[27] P. P. Brahma, Y. She, S. Li, J. Li, and D. Wu, "Reinforced Robust Principal Component Pursuit", *IEEE Transactions on Neural Networks and Learning Systems*, Vol. 29, No. 5, pp. 1525-1538, 2017.

[28] N. Goyette, P. M. Jodoin, F. Porikli, J. Konrad, and P. Ishwar, "Changedetection.net: A New Change Detection Benchmark Dataset", In: *IEEE Computer Society Conf. On Computer Vision and Pattern Recognition Workshops, Providence*, pp. 1-8, 2012.