An Automated Unified Framework for Video Deraining and Simultaneous Moving Object Detection in Surveillance Environments

BAIJU P. S. , (Graduate Student Member, IEEE), AND SUDHISH N. GEORGE , (Member, IEEE)
National Institute of Technology Calicut, Calicut 673601, India

Corresponding author: Baiju P. S. (baijupstvm@gmail.com)

ABSTRACT In many instances, outdoor surveillance systems suffer from atrocious weather conditions such as rain, since images or videos captured by such vision systems in rainy days may undergo severe visual dilapidations. This can cause a glitch in those algorithms which are further used for object detection and tracking. Therefore, an ancillary video processing algorithm namely, video deraining is necessary prior to the implementation of object detection and tracking. This indicates the requirement of a time-consuming and complicated two-step process for the detection of moving objects in a rainy environment. This paper proposes an automated single-stage formulation for the conventional three-stage procedure such as rain steak removal, original data recovery and moving object detection as simultaneous operation in the tensor framework. The brilliance of this work is confined in the efficient formulation of an operator termed as Slice Rotational Total Variation (SRTV), to unify the different rain patterns into a common pattern so that any form of rain pattern can be effectively removed from the rainy data in a visually appealing manner by preserving the important details of the data. In this paper, SRTV regularization and tensor low-rank minimization are utilized for the effective deraining as well as efficient retrieval of clean background. Besides, l_1 norm and Tensor Total Variation (TTV) regularizers together with SRTV regularizer are employed for the faithful detection of derained moving objects. The experimental results show that the proposed method outperforms the state of the art methods in terms of deraining capability and accurate moving object detection.

INDEX TERMS Low rank approximation, moving object detection, tensor total variation and video deraining.

I. INTRODUCTION

Usually, most of the modern outdoor vision systems show adequate performance in clear climatic conditions. However, certain dynamic weather conditions such as rain will degrade the visual quality of outdoor scenes, thereby deteriorating the performance of many image processing and computer vision algorithms such as object detection, event detection, tracking, classification, scene analysis and surveillance. Hence it is essential to remove such undesirable visual artifacts caused by the rain on outdoor images or videos so that the outdoor vision system can achieve better performance and greater accuracy.

Many methods have been proposed in literature to remove the rain streaks from input rainy video. Existing deraining algorithms can be classified into four approaches such as

The associate editor coordinating the review of this manuscript and approving it for publication was Hugo Proenca .

time domain, frequency domain, learning and reconstruction approaches. Time domain based techniques basically made the use of chromatic and temporal properties of the rain for rain removal. Grag *et al.* proposed a method to detect and remove rain from video data using the physical properties of raindrops [1], [2]. However, this method is incapable of distinguishing rain streaks from moving objects, when video data contains heavy rain streaks. Starik *et al.* introduced a video deraining algorithm in which each rainy pixel value was replaced by the median of corresponding pixel values in temporal frames [3]. However, median filtering brought blurring artifacts around the moving objects in the videos. Zhang *et al.* developed a method for removing the rain streaks by exploiting the temporal and chromatic properties [4]. This method debases the visual quality of video with dynamic background. Park *et al.* designed a video deraining algorithm based on Kalman filtering [5]. This algorithm is inefficient to detect rainy pixels from videos containing moving objects.

Brewer *et al.* analyzed the shape characteristics of rain streaks to identify and remove rain streaks from the video data [6]. This technique misclassifies the rain streaks when either multiple rain streaks intersect at some point or a rain streak intersects some other scene motion. These time domain based methods require all the available frames in a video, even then all the potential rain streaks may not be detected and removed effectively.

Barnum *et al.* suggested a blurred Gaussian model in frequency domain for detecting the rain streaks [7]. This idea leads to visually unpleasant effects in image space when any changes are made in frequency space. Lin *et al.* composed a bilateral filter to disintegrate the rainy image into low frequency and high frequency parts and the high frequency part is further decomposed into rain and non-rain components [8]. Chen *et al.* employed a guided filter to transform the input image into frequency domain and extracted the rain streaks from high frequency component by means of histogram of oriented gradients [9]. These frequency domain methods may not always lead to pleasing effects in spatial domain whenever changes are made in frequency domain. Moreover, if the frequency components corresponding to rain streaks contain clutters, degraded results will be produced.

Kang *et al.* presented a dictionary learning based method that uses Morphological Component Analysis (MCA) decomposition [10]. Chiang *et al.* utilized the concept of dictionary learning to compute the rain streak map [11]. However, this method misdetects the non-rain components as rain. Li *et al.* invented a rain removal algorithm using Gaussian Mixture Models (GMM) [12]. This algorithm is inadequate on images having patches with a substantial amount of background details corrupted by many rain streaks. Mi *et al.* employed a fused dictionary for removing the rain component from videos [13]. However, this method fails to remove the entire rain streaks from videos. Ren *et al.* designed Markov Random Fields (MRFs) to distinguish between moving objects and rain streaks [14]. This method is unable to deal with the dynamic components of the video data. Zhu *et al.* suggested a method that uses three priors on two layers of rainy input image [15]. However, this method causes over-smoothening of the background details. These learning based methods have high time complexity and cause visually unpleasant results in the output.

Kim *et al.* proposed a two stage process to remove the rain streaks using rain map which is obtained by optical flow estimation process [16]. However, this method is limited when the input video contains heavy rain streaks. Jiang *et al.* developed an algorithm in which unidirectional tensor total variation regularizers are used to reconstruct the rain-free video data [17]. This algorithm is incompetent when the video is having oblique rain streaks. Wei *et al.* invented a method that encodes rain in stochastic manner as patch-based mixture of Gaussians [18]. This method is not suitable for removing heavy rain patterns. Wang *et al.* formulated group sparsity based optimization model for rain removal from videos [19]. This method is limited when the rain direction

is far away from vertical direction. Huang *et al.* invented a video rain streak removal algorithm using directional gradient priors [20]. This algorithm was inadequate for handling the residual rain artifacts and the texture features of the dynamic objects in the video are not well maintained. Sun *et al.* proposed an algorithm based on directional regularized tensor for the rain removal [21]. However, this algorithm was incapable of removing heavy rain streaks from videos.

One of the prime merits of reconstruction based deraining methods is that the derained video can be obtained from the available single input rainy video data. Moreover, the non-essentialness of prior database makes these methods are more convenient in nature and a time consuming training session which is essential for learning based approaches can be avoided. Hence, our research is mainly focused on the design of video deraining algorithm using reconstruction based approach in order to negate the above mentioned drawbacks of the existing techniques. Hence the aim of this work is to design an efficient video deraining technique which can detect and remove the unwanted rain steaks from the rainy video data and reconstruct the clean data by incorporating the missing details from the available rainy data alone. Among various reconstruction based methods, optimization approach using low rank concept with suitable regularizers is the leading and emerging technique for various video processing applications. This is the main motive of the proposed work.

Apart from rain removal methods, several Moving Object Detection (MOD) algorithms using the concept of sparse and low rank decomposition technique have been introduced in the past decade. These techniques are dominant compared with the conventional MOD techniques in terms of accurate object detection and exact recovery of background information. Candes *et al.* proposed a low rank approximation based method for background subtraction using Robust Principal Component Analysis (RPCA) [22]. Zhou *et al.* developed a classical method for object detection using low rank representation [23]. Cao *et al.* invented a modified version of RPCA for foreground detection by utilizing spatio-temporal continuity behaviour of moving objects [24]. Chen *et al.* developed a method for separating background and moving objects, by extending the matrix based RPCA into tensor domain [25]. This method is restricted to videos with less number of frames. Sajid *et al.* suggested an online tensor decomposition based method for extracting moving objects from videos, which is effective for videos containing more frames [26]. Anju *et al.* developed tensor based Low Rank with Total Variation model for MOD applications [27]. Shijila *et al.* introduced a matrix based technique for partitioning foreground and background by incorporating l_1 norm and Total Variation (TV) regularizations into RPCA [28]. The same authors extended their work [28] to detect the moving objects in noisy environment by including an extra term for inculcating the noise in the observation model [29]. In [30], an MOD method for extreme surveillance environments by using twist spatio-temporal total variation to enhance the detection performance of the foreground and tensor nuclear
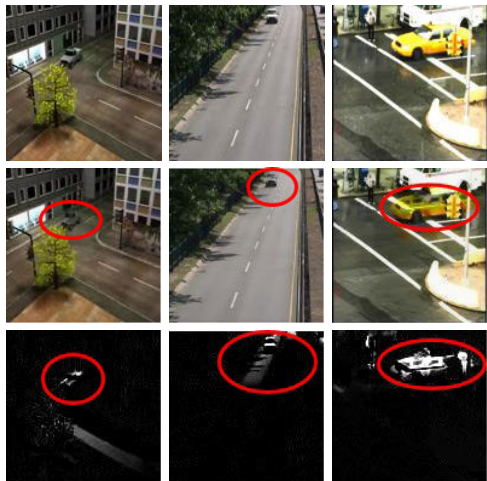
FIGURE 1. Illustration of failure cases of Fastderain [20] in rain removal process. Top row: Original frames of 'truck', 'highway' and 'traffic' videos respectively. Middle row: Corresponding derained results obtained by Fastderain [20]. Bottom row: Corresponding MOD results obtained by [29]. The circled portions correspond to the moving objects with distorted texture features.

norm minimization for efficient background separation was proposed.

However, the object detection algorithms mentioned above were actually designed for clear weather conditions. Bad weather conditions such as rain may put these detection methods into trouble. Even if the the recent reconstruction based deraining technique such as Fastderain [20] can remove the rain steaks in a better manner, it will adversely affect the texture features of the moving objects especially when small sized moving objects are being considered. This scenario is depicted in the middle row of Fig. 1. This will diminish the detection accuracy of moving objects when the derained data is fed to an MOD system as illustrated in the bottom row of Fig. 1, where the most recent MOD method [29] is used for verification. Moreover, blurring of objects may create some misinterpretation about the data. For example, it may adversely affect the number plate recognition system or similar type of computer vision applications.

In summary, even the most recent and popular deraining techniques cannot preserve texture features of moving objects and the recent MOD methods cannot detect the moving objects from a derained data in an effective manner. Thus, it is beneficial to devise an algorithm to address all these drawbacks in a single step. Hence the aim of this work is reformulated as to design an efficient single stage algorithm for three objectives such as rain streak removal, recovery of rain-free data and moving object detection in a unified framework without much computational overhead. In this work, we model the video data as a third order tensor and tensor decomposition technique is attempted to address the above mentioned objective.

The rest of the paper is structured as follows. Section II provides the mathematical preliminaries required for the better description of the proposed method. Section III explains

TABLE 1. Mathematical notations.

Notation	Description
$\mathcal{X}, \mathbf{x}, x$	Tensor, vector, scalar
$\mathcal{X}(m_1, :, m_3, m_4, ... m_N)$	A fibre of a tensor $\mathcal{X}$, defined by fixing all indices except one.
$\mathcal{X}(m_1, :, :, m_4, ... m_N)$	A slice of a tensor $\mathcal{X}$, defined by fixing all indices except two.

the proposed work in which the formulation of the optimization and the solution of the proposed model are elaborated. Section IV discusses experimental results and performance evaluation. Lastly, section V concludes the paper.

II. MATHEMATICAL PRELIMINARIES AND NOTATIONS

This section presents the preliminaries on the concept of tensor algebra used in this paper. The mathematical notations used in this paper are shown in Table 1. A tensor, $\mathcal{X} \in \mathbb{R}^{m_1 \times m_2 \times .... m_N}$ is a multi-dimensional array [31]–[36]. In this paper, Euler's script is used to denote the tensors. Slice of a tensor is a two dimensional section defined by all but two indices [32], [35], [36]. A third order tensor, $\mathcal{X} \in \mathbb{R}^{m_1 \times m_2 \times m_3}$ has horizontal, lateral and frontal slices denoted by $\mathcal{X}(k, :, :)$, $\mathcal{X}(:, k, :)$ and $\mathcal{X}(:, :, k)$ respectively, where, k is the slice number. A fiber of a tensor is a one dimensional segment defined by fixing all indices except one [37]. A third order tensor, $\mathcal{X}$ has mode-1, mode-2 and mode-3 fibers denoted by $\mathcal{X}(:, j, k)$, $\mathcal{X}(i, :, k)$ and $\mathcal{X}(i, j, :)$ respectively.

A. TENSOR PRODUCT (T-PRODUCT)

t-product of two 3rd order tensors, $\mathcal{X} \in \mathbb{R}^{m_1 \times m_2 \times m_3}$ and $\mathcal{Y} \in \mathbb{R}^{m_2 \times m_4 \times m_3}$ is defined as [38],

$$\mathcal{Z} = \mathcal{X} * \mathcal{Y}, \quad \mathcal{Z} \in \mathbb{R}^{m_1 \times m_4 \times m_3} \quad (1)$$

and,

$$\mathcal{Z}(x, y, :) = \sum_{z=1}^{m_2} \mathcal{X}(x, z, :) \odot \mathcal{Y}(z, y, :) \quad (2)$$

where, $\odot$ is the circular convolution operator.

B. TENSOR SINGULAR VALUE DECOMPOSITION (T-SVD)

t-SVD of a tensor, $\mathcal{X} \in \mathbb{R}^{m_1 \times m_2 \times m_3}$ is given by [36], [38], [39],

$$\mathcal{X} = \mathcal{U} * \Sigma * \mathcal{V}^T \quad (3)$$

where, Σ is a rectangular tensor with each frontal slice as f-diagonal of size $m_1 \times m_2 \times m_3$, $\mathcal{U}$ and $\mathcal{V}$ are orthogonal tensors of size $m_1 \times m_1 \times m_3$ and $m_2 \times m_2 \times m_3$ respectively and '$*$' denotes t-product.

C. TENSOR MULTI-RANK

The multi-rank of a tensor, $\mathcal{X} \in \mathbb{R}^{m_1 \times m_2 \times m_3}$ is a vector, $r \in \mathbb{R}^{m_3}$, whose individual elements are the ranks of frontal slices of Fourier transform of the tensor, $\mathcal{X}$ [36], [40].

$$\text{multi-rank}(\mathcal{X}) = [r_1, r_2, \ldots, r_{m_3}] \quad (4)$$

where, $r_i = rank(\mathcal{X}_f^{(i)})$, $i = 1, 2, .., m_3$ and $\mathcal{X}_f^{(i)}$ denotes the Fourier transform of ith frontal slice of $\mathcal{X}$.

D. l_1 NORM

The l_1 norm of a 3rd order tensor, $\mathcal{X} \in \mathbb{R}^{m_1 \times m_2 \times m_3}$ is given by [37],

$$\|\mathcal{X}\|_1 = \sum_{x,y,z} |\mathcal{X}(x,y,z)|. \tag{5}$$

E. TENSOR NUCLEAR NORM (TNN)

The tensor nuclear norm of a tensor, $\mathcal{X} \in \mathbb{R}^{m_1 \times m_2 \times m_3}$ is defined as [38], [39],

$$\|\mathcal{X}\|_{\circledast} = \sum_{z=1}^{m_3} \sum_{x=1}^{min(m_1,m_2)} |\Sigma_f(x,x,z)|. \tag{6}$$

where, Σ_f is the Fourier transform of a rectangular tensor, Σ with each frontal slice as f-diagonal of size $m_1 \times m_1 \times m_3$.

F. WEIGHTED TENSOR NUCLEAR NORM (WTNN)

For a tensor, $\mathcal{X} \in \mathbb{R}^{m_1 \times m_2 \times m_3}$ with $\Sigma_f \in \mathbb{R}^{\min(m_1,m_2) \times m_3}$ as its singular value tensor in Fourier domain, Weighted Tensor Nuclear Norm operator $\|.\|_{\circledast \mathcal{W}}$ is defined as [37],

$$\|\mathcal{X}\|_{\circledast \mathcal{W}} = \sum_{z=1}^{m_3} \sum_{x=1}^{min(m_1,m_2)} \mathcal{W}_{\mathcal{X}}(x,x,z)|\Sigma_f(x,x,z)| \tag{7}$$

where, $\mathcal{W}_{\mathcal{X}} \in \mathbb{R}^{\min(m_1,m_2) \times m_3}$ is a weight tensor whose individual element is given by

$$\mathcal{W}_{\mathcal{X}}(x,x,z) = \frac{1}{\Sigma_f(x,x,z)+\epsilon}, \qquad \epsilon > 0$$

G. TENSOR TOTAL VARIATION (TTV) NORM

For a given tensor, $\mathcal{X} \in \mathbb{R}^{m_1 \times m_2 \times m_3}$, anisotropic total variation is defined as [24],

$$\|\mathcal{X}\|_{TTV} = \sum_{x,y,z} (D_h\mathcal{X} + D_v\mathcal{X} + D_t\mathcal{X}) \tag{8}$$

where, $D_h\mathcal{X} = |\mathcal{X}_{(x,y,z)} - \mathcal{X}_{(x+1,y,z)}|$ and $D_v\mathcal{X} = |\mathcal{X}_{(x,y,z)} - \mathcal{X}_{(x,y+1,z)}|$ and $D_t\mathcal{X} = |\mathcal{X}_{(x,y,z)} - \mathcal{X}_{(x,y,z+1)}|$

In the following section, the proposed method for video deraining and simultaneous moving object detection by utilizing the above mentioned concepts in tensor algebra is presented.

III. PROPOSED METHOD

Video deraining task generally aims at removing the rain streaks and recovering the original data from the outdoor rainy videos. However, object detection is the main operation of many surveillance systems where an additional step is required for detecting the objects from derained videos. This two step process takes more execution time and is not suitable for real time applications. Moreover, the texture features of dynamic objects will be lost during the deraining process even for the most recent deraining algorithms which may adversely affect the detection performance of the MOD system. Hence, our method proposes a single stage optimization method for deraining and moving object detection in an effective

manner. The objectives of the this problem consists of efficient removal of the rain components without affecting the important features of the original rain-free data and accurate detection of moving objects in surveillance videos.

In the proposed method, rainy video, $\mathcal{V}$ is modeled as a third order tensor in order to effectively utilize the spatial and temporal features of the data. It is because of the fact that the application of matrix algebra ignores the temporal redundancy present in the data as it is restricted to 2D space only. However, the different operators in tensor algebra help to negate this drawback to a certain extend by considering the data as a whole instead of individual slices. The proposed method decomposes the rainy data, $\mathcal{V}$ into three 3rd order tensors as given below.

$$\mathcal{V} = \mathcal{B} + \mathcal{F} + \mathcal{R} \tag{9}$$

where, $\mathcal{B}$, $\mathcal{F}$ and $\mathcal{R}$ are background, moving objects or foreground and rain components respectively. The proposed decomposition model is illustrated in Fig 2. Hence the objective is to accurately decompose $\mathcal{B}$, $\mathcal{F}$ and $\mathcal{R}$ components for the efficient separation of background component from the foreground component for moving object detection succeeded by the integration of these components for obtaining the rain-free data.
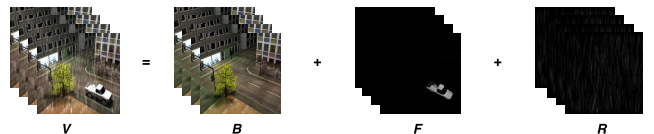

FIGURE 2. Three way decomposition of rainy video.

The significant contributions of the proposed method are summarized below.

- To the best of our understanding, this is the first attempt to introduce rain steak removal, rain-free data recovery and moving object detection for outdoor surveillance videos by solving a single objective function in a unified framework from the available rainy data in the tensor framework. This objective is accomplished by decomposing the 3rd order rainy tensor as the sum of three tensors such as rainy, background and foreground components.
- A new operator termed as Slice Rotational Total Variation (SRTV) is formulated to amalgamate the different types of rain patterns such as vertical and oblique rain streaks to a single pattern so that the further processing can become more simplified and efficient.
- A tensor low rank regularization term is incorporated in the formulated optimization model for exploiting the low rank nature of background component. The low rank regularization term together with SRTV regularization helps to increase the efficacy of exact background separation without rain steaks from the rainy input data.
- In order to extract the sparse moving objects, a greedy sparsity regularization term such as l_1 minimization

term is included. The inclusion of Tensor Total Variation (TTV) regularization helps the fine tuning of exact foreground and background separation. Moreover, the inclusion of SRTV regularization terms removes the rain steaks from the foreground component. The integration of decomposed and derained foreground and background components provides the rain-free video data.

- The formulated three-way objective function is solved using Augmented Lagrangian Method (ALM) with Alternating Direction Method of Multipliers (ADMM) technique. The most popular and recent techniques in the domain of video deraining and moving object detection are compared with the proposed scheme. It is found that the proposed method excels in majority cases compared with the other schemes. However, none of these compared techniques can perform video deraining and simultaneous MOD as a single stage operation like the proposed method.

A. THEORY BEHIND THE PROPOSED MODEL

In this section, the basic idea behind the decomposition of rainy video into foreground, background and rainy components are outlined.

1) REMOVAL OF RAINY COMPONENT

Rain streak removal is the crucial stage of deraining process. This is because of the fact that the nature of rain streaks such as light and heavy and the direction of rain streaks such as vertical and oblique may differently affect the removal process of the rain component. To diminish this problem, a unique operator called Slice Rotational Total Variation (SRTV) is introduced for the effective elimination of rain streaks from background and foreground components irrespective of nature as well as direction of the rain streaks.

a: SLICE ROTATIONAL TOTAL VARIATION (SRTV) OPERATOR

In this technique, each horizontal slice of the rainy data is transformed into frontal slice by rotating an angle of 90^0 and then two dimensional Total Variation (TV) is applied to each rotated frontal slice. Fig. 3 shows that despite the fact that the frontal slices of heavy and oblique rainy videos are having different rain pattens, their corresponding horizontal slices look alike and occupy dot-like pattern. Hence, if all the horizontal slices are converted to frontal slices by means of rotation, minimization of simple two dimensional Total Variation (TV) operator will be adequate to remove such dot-like pattern and thereafter the reverse process of slice rotation yields the clear background /foreground. Thus, SRTV operator is an efficacious method for eradicating rain component from videos having complex rain patterns such as heavy and oblique rain streaks. SRTV operator can be mathematically expressed as follows.

Let $\mathcal{Y} \in \mathbb{R}^{m_3 \times m_2 \times m_1}$ be the resultant tensor obtained by applying slice rotation on a tensor, $\mathcal{X} \in \mathbb{R}^{m_1 \times m_2 \times m_3}$. Then, the entries in each horizontal slice of $\mathcal{X}$ are same as those in
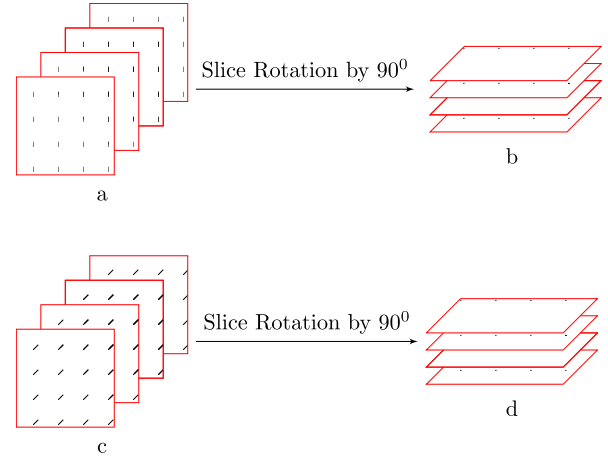

FIGURE 3. (a) and (c) Frontal slices with vertical rain streaks and oblique rain streaks respectively, (b) and (d) Horizontal slices (obtained by rotation) of (a) and (c) respectively.

the corresponding frontal slice of $\mathcal{Y}$.

$$ie. \mathcal{X}(i, :, :) \equiv \mathcal{Y}(:, :, i) \quad \text{for} \quad i = 1, 2, \ldots, m_1 \quad (10)$$

Now, SRTV of $\mathcal{X}$ is defined as,

$$\|\mathcal{X}\|_{SRTV} = \sum_z \left[\sum_{x,y} [D_h \mathcal{Y}(:, :, z) + D_v \mathcal{Y}(:, :, z)] \right] \quad (11)$$

where,

$$D_h \mathcal{Y}(:, :, z) = |\mathcal{Y}_{(x,y,z)} - \mathcal{Y}_{(x+1,y,z)}| \quad (12)$$

and

$$D_v \mathcal{Y}(:, :, z) = |\mathcal{Y}_{(x,y,z)} - \mathcal{X}_{(x,y+1,z)}| \quad (13)$$

2) SEPARATION OF BACKGROUND COMPONENT

The background component has spatial correlation within each frame and temporal correlation among the frames. This spatio-temporal consistency reveals that the background is low rank in nature [41]. The popular tensor robust principal component analysis (TRPCA) suggests the modeling of background component as a low rank tensor [42]. Hence tensor multi-rank minimization is the best solution for recovering background from videos. Since multi-rank minimization is not convex in nature, unique solution can be obtained by replacing the problem by Tensor Nuclear Norm (TNN) minimization, where, TNN is assumed to be the best convex surrogate of tensor multi-rank [40]. However, in [38], Baburaj *et al.* proposed that re-weighting the singular values in terms of their significance improves the performance of low rank recovery in a more fruitful manner. Hence, the proposed model utilizes Weighted Tensor Nuclear Norm (WTNN) minimization for enhancing the efficient recovery of background components. The WTNN minimization of foreground component together with SRTV minimization on the rainy component will yield the retrieval of rain-free component.

3) SEPARATION OF FOREGROUND COMPONENT

The moving objects in each frame usually contain lesser number of pixels. As a result, they exhibit sparse nature. Therefore, l_0 norm minimization can provide foreground component. Since l_0 norm is non-convex in nature, l_1 norm minimization is considered as the best substitute for the same [43]. Moreover, the moving objects have salient and continuous change in intensity along the spatial as well as temporal directions. Hence the more accurate foreground detection can be achieved by exploiting this spatio-temporal continuity of the data. Tensor Total Variation (TTV) minimization is the superior method for detecting moving objects by imposing the spatio-temporal continuity constraints [24]. Thus, the combined usage of l_1 norm and TTV norm minimization along with SRTV regularization will provide the detection of rain-free foreground components.

B. FORMULATION OF OPTIMIZATION MODEL

The aforementioned mathematical ingredients for the proposed method are recapped in terms of the following short comments;

- Introduction of the joint processing by applying WTNN minimzation on background component and SRTV norm minimization on rainy component accomplishes mutual exclusiveness between background and rain streaks. The joint action of these two regularization terms results a visually pleasing background video without rain steaks.
- Since it is assumed that the foreground component poses sparse nature, sparsity greedy l_1 norm minimization is incorporated. In order to ensure the piecewise smoothness, TTV norm minimization term is included so that the fine tuning of exact foreground component can be ensured. Moreover, the formulated SRTV norm minimization term ensures the removal of rain steaks from the foreground component.

Considering all the above-mentioned points, an optimization model can be formulated as,

$$\underset{\mathcal{B},\mathcal{F},\mathcal{R}}{\text{argmin}} \ \|\mathcal{B}\|_{\circledast\mathcal{W}} + \lambda_1 \|\mathcal{F}\|_1 + \lambda_2 \|\mathcal{G}\|_{TTV} + \lambda_3 \|\mathcal{R}\|_{SRTV}$$
$$\text{s.t } \mathcal{V} = \mathcal{B} + \mathcal{F} + \mathcal{R} \quad \text{and } \mathcal{F} = \mathcal{G} \tag{14}$$

where, $\|.\|_{\circledast\mathcal{W}}$, $\|.\|_1$, $\|.\|_{TTV}$ and $\|.\|_{SRTV}$ represent WTNN, l_1, TTV and SRTV oprerators respectively.

The Augmented Lagrangian function for Eq. (14) is defined as,

$$\mathcal{L}_{\mu,\lambda_1,\lambda_2,\lambda_3,}(\mathcal{R},\mathcal{F},\mathcal{G},\mathcal{B},\mathcal{Y}_1,\mathcal{Y}_2)$$
$$= \|\mathcal{B}\|_{\circledast\mathcal{W}} + \lambda_1 \|\mathcal{F}\|_1$$
$$+ \lambda_2 \|\mathcal{G}\|_{TTV} + \lambda_3 \|\mathcal{R}\|_{SRTV}$$
$$+ <\mathcal{Y}_1, \mathcal{V} - \mathcal{B} - \mathcal{F} - \mathcal{R}> + <\mathcal{Y}_2, \mathcal{G} - \mathcal{F}$$
$$> + \frac{\mu}{2} \|\mathcal{V} - \mathcal{B} - \mathcal{F} - \mathcal{R}\|_F^2 + \frac{\mu}{2} \|\mathcal{G} - \mathcal{F}\|_F^2 \tag{15}$$

The problem defined in Eq. (15) can be solved by using Alternating Direction Method of Multipliers (ADMM) in which the whole problem is divided into four subproblems and each one is solved in an iterative manner as follows [44].

1) $\mathcal{B}$ SUBPROBELM

$$\mathcal{B}^{[n+1]} = \underset{\mathcal{B}}{\text{argmin}} \ \|\mathcal{B}\|_{\circledast\mathcal{W}}$$
$$+ \frac{\mu}{2} \left\| \mathcal{B} - \left[\mathcal{V} - \mathcal{R}^{[n]} - \mathcal{F}^{[n]} + \frac{\mathcal{Y}_1^{[n]}}{\mu} \right] \right\|_F^2 \tag{16}$$

The closed form solution of Eq. (16) is given by,

$$\mathcal{B}^{[n+1]} = \mathcal{D}_{\frac{\mathcal{W}^{[n]}}{\mu}} \left[\mathcal{V} - \mathcal{R}^{[n]} - \mathcal{F}^{[n]} + \frac{\mathcal{Y}_1^{[n]}}{\mu} \right] \tag{17}$$

where, $\mathcal{D}_\delta(\mathcal{X})$ is the singular value convolution operator defined on a tensor $\mathcal{X} \in \mathbb{R}^{m_1 \times m_2 \times m_3}$ having t-SVD $\mathcal{X} = \mathcal{U} * \Sigma * \mathcal{V}^T$ given by [38],

$$\mathcal{D}_\delta(\mathcal{X}) = \mathcal{U} * \mathcal{C}_\delta(\Sigma) * \mathcal{V}^T \tag{18}$$

where, $\mathcal{C}_\delta(\Sigma) = \Sigma * \mathcal{J}$ and $\mathcal{J}$ is an f-diagonal tensor whose diagonal element in Fourier domain is

$$\mathcal{J}_f(x, x, z) = \left(1 - \frac{\delta}{\Sigma_f(x, x, z)} \right) \tag{19}$$

here, Σ_f is the Fourier transform of Σ and $1 \leq x \leq m_1$, $1 \leq z \leq m_3$, $m_1 < m_2$

2) $\mathcal{F}$ SUBPROBELM

$$\mathcal{F}^{[n+1]} = \underset{\mathcal{F}}{\text{argmin}} \ \lambda_1 \|\mathcal{F}\|_1 + \frac{\mu}{2} \left\| \mathcal{F} - \mathcal{K}^{[n+1]} \right\|_F^2 \tag{20}$$

where,

$$\mathcal{K}^{[n+1]} = \left[\frac{\mathcal{V} - \mathcal{B}^{[n+1]} - \mathcal{R}^{[n]} + \mathcal{G}^{[n]}}{2} \right] + \left[\frac{\mathcal{Y}_1^{[n]} + \mathcal{Y}_2^{[n]}}{2\mu} \right] \tag{21}$$

The closed form solution of Eq. (20) is given by,

$$\mathcal{F}^{[n+1]} = \mathcal{S}_{\frac{\lambda_1}{2\mu}} \left[\mathcal{K}^{[n+1]} \right] \tag{22}$$

where, $\mathcal{S}_\tau(\mathcal{X}) = sgn(\mathcal{X}) \times \max(\mathcal{X} - \tau, 0)$ is the soft thresholding operation [42].

3) $\mathcal{G}$ SUBPROBELM

$$\mathcal{G}^{[n+1]} = \underset{\mathcal{G}}{\text{argmin}} \ \lambda_2 \|\mathcal{G}\|_{TTV} + \frac{\mu}{2} \left\| \mathcal{G} - \left[\mathcal{F}^{[n+1]} - \frac{\mathcal{Y}_2^{[n]}}{\mu} \right] \right\|_F^2 \tag{23}$$

The Eq. (23) can be solved as,

$$\mathcal{G}^{[n+1]} = TTV_{\frac{\lambda_2}{\mu}} \left[\mathcal{F}^{[n+1]} - \frac{\mathcal{Y}_2^{[n]}}{\mu} \right] \tag{24}$$

where, $TTV[.]$ denotes tensor TV minimization algorithm in [45].

4) $\mathcal{R}$ SUBPROBELM

$$\mathcal{R}^{[n+1]}$$
$$= \underset{\mathcal{R}}{\arg\min} \ \lambda_3 \left\| \mathcal{R} \right\|_{SRTV}$$
$$+ \frac{\mu}{2} \left\| \mathcal{R} - \left[\mathcal{V} - \mathcal{B}^{[n+1]} - \mathcal{F}^{[n+1]} + \frac{\mathcal{Y}_1^{[n]}}{\mu} \right] \right\|_F^2 \quad (25)$$

The solution of the Eq. (25) is given by,

$$\mathcal{R}^{[n+1]} = SRTV_{\frac{\lambda_3}{\mu}} \left[\mathcal{V} - \mathcal{B}^{[n+1]} - \mathcal{F}^{[n+1]} + \frac{\mathcal{Y}_1^{[n]}}{\mu} \right] \quad (26)$$

where, $SRTV[.]$ is the slice rotational TV minimization operator defined by,

$$SRTV[\mathcal{X}] = \mathcal{P} \uparrow$$

where, $\mathcal{P} \uparrow$ is the frontal to horizontal slice rotated version of the tensor,

$$\mathcal{P} = TV[\mathcal{X} \downarrow].$$

where, $\mathcal{X} \downarrow$ is the horizontal to frontal slice rotated version of the tensor $\mathcal{X}$ and $TV[.]$ is the two dimensional TV minimization procedure applied to each frontal slice of $\mathcal{X} \downarrow$ [46]. Finally, the Lagrangian multipliers are updated by the following equations,

$$\mathcal{Y}_1^{[n+1]} = \mathcal{Y}_1^{[n]} + \mu \left[\mathcal{V} - \mathcal{B}^{[n+1]} - \mathcal{F}^{[n+1]} - \mathcal{R}^{[n+1]} \right] \quad (27)$$

$$\mathcal{Y}_2^{[n+1]} = \mathcal{Y}_2^{[n]} + \mu \left[\mathcal{G}^{[n+1]} - \mathcal{F}^{[n+1]} \right] \quad (28)$$

Here, the final updates of $\mathcal{B}$ and $\mathcal{F}$ can provide clear background and foreground components respectively. Since, the resultant $\mathcal{F}$ component does not contain rain streaks, our algorithm is able to precisely detect the moving objects in rainy environments. Moreover, the original rain-free video can be effectively regained by simply adding the last updates of $\mathcal{B}$ and $\mathcal{F}$ components together.

The algorithm description for solving the aforementioned optimization problem is summarized in **Algorithm 1**.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

The performance evaluation of the proposed method is done on synthetic and real video sequences. In order to assess the efficacy of the proposed method for removing rain streaks from videos, we compare our method with latest state of the art deraining methods such as Kim *et al.*'s Low Rank Matrix Completion (LRMC) method [16], Wei *et al.*'s method [18], Huang *et al.*'s Fast Derain method [20] and Sun *et al.*'s method [21]. Morevover, the capability of our method for detecting moving objects is validated by comparing with the recent MOD techniques such as Sajid *et al.*'s Online Tensor Decomposition (OTD) method [26], Chen *et al.*'s method [25] and Shijila *et al.*'s method [29]. None of these method can perform the deraining and moving object detection as a simultaneous action.

The proposed method is implemented on the platform of Linux 14.04 and Matlab (2016a) with an Intel(R) Xenon(R)

Algorithm 1 Rain Removal and Foreground Extraction

Data: Rainy video, $\mathcal{V} \in \mathbb{R}^{m_1 \times m_2 \times m_3}$
Result: Foreground video, $\mathcal{F}$ and Derained video, $\mathcal{B} + \mathcal{F}$
Initialization: $\lambda_1 > 0$, $\lambda_2 > 0$, $\mu^{(0)} > 0$, $\mu_{max} = 10^4$, $\rho > 0$, $\mathcal{R}^{(0)} = \mathcal{F}^{(0)} = \mathcal{B}^{(0)} = \mathcal{Y}_1^{(0)} = \mathcal{Y}_2^{(0)} = 0$
converged = false
while *not converged* **do**

$\mathcal{B}^{[n+1]} = \mathcal{D}_{\frac{\mathcal{W}^{[n]}}{\mu}} \left[\mathcal{V} - \mathcal{R}^{[n]} - \mathcal{F}^{[n]} + \frac{\mathcal{Y}_1^{[n]}}{\mu} \right]$

$\mathcal{K}^{[n+1]} = \left[\frac{\mathcal{V} - \mathcal{B}^{[n+1]} - \mathcal{R}^{[n]} + \mathcal{G}^{[n]}}{2} \right] + \left[\frac{\mathcal{Y}_1^{[n]} + \mathcal{Y}_2^{[n]}}{2\mu} \right]$

$\mathcal{F}^{[n+1]} = \mathcal{S}_{\frac{\lambda_1}{2\mu}} \left[\mathcal{K}^{[n+1]} \right]$

$\mathcal{G}^{[n+1]} = TTV_{\frac{\lambda_2}{\mu}} \left[\mathcal{F}^{[n+1]} - \frac{\mathcal{Y}_2^{[n]}}{\mu} \right]$

$\mathcal{R}^{[n+1]} = SRTV_{\frac{\lambda_3}{\mu}} \left[\mathcal{V} - \mathcal{B}^{[n+1]} - \mathcal{F}^{[n+1]} + \frac{\mathcal{Y}_1^{[n]}}{\mu} \right]$

$\mathcal{Y}_1^{[n+1]} = \mathcal{Y}_1^{[n]} + \mu \left[\mathcal{V} - \mathcal{B}^{[n+1]} - \mathcal{F}^{[n+1]} - \mathcal{R}^{[n+1]} \right]$

$\mathcal{Y}_2^{[n+1]} = \mathcal{Y}_2^{[n]} + \mu \left[\mathcal{G}^{[n+1]} - \mathcal{F}^{[n+1]} \right]$

$\mu^{[n+1]} = \min \left(\rho \mu^{[n]}, \mu_{max} \right)$

if $\frac{\left\| \mathcal{V} - \mathcal{B}^{[n+1]} - \mathcal{F}^{[n+1]} - \mathcal{R}^{[n+1]} \right\|_F}{\|\mathcal{V}\|_F} \leq threshold$ **then**
$\quad converged = true$
end
$n = n + 1$
end

E5-1620 CPU at 3.70GHz and 8GB RAM. Since our method is capable of removing rain streaks as well as detecting moving objects, we need to analyse both deraining performance and object detection performance separately.

A. DERAINING PERFORMANCE ANALYSIS

In order to validate the capacity of our method for removing rain streaks, synthetic videos with different rain patterns such as light, heavy and oblique, and certain real rainy videos are utilized.

1) SYNTHETIC DATA

Synthetic rainy videos are obtained by adding three types of simulated rain streaks such as light rain, heavy rain and rain at an arbitrary angle with ground truth videos named 'truck', 'highway', 'pedestrian', 'park', 'backdoor' and 'bungalows', which are collected from CD.net and SABS datasets. The size of the videos is set to $256 \times 256 \times 3 \times 100$. The parameters, ρ and ϵ are experimentally set as, $\rho = 2.5$ and $\epsilon = 0.01$

For the comparison purpose of our method with existing video deraining techniques such as [16], [18], [20] and [21], the quality metrics such as peak signal-to-noise ratio (PSNR) in dB, structural similarity (SSIM) [47] and inverse Relative Error (iRSE) in dB [38] are calculated for each frame of the aforementioned synthetic videos and the corresponding mean values of such parameters along with execution time are entered in Table 2. As shown in Table 2, the proposed

TABLE 2. Quantitative comparisons of rain removal results on synthetic videos.

Video	Method	Light				Heavy				Oblique			
Rain Types		PSNR	SSIM	iRSE	Time(s)*	PSNR	SSIM	iRSE	Time(s)*	PSNR	SSIM	iRSE	Time(s)*
Truck	Rainy	51.05	0.9691	22.55	-	41.32	0.9120	14.35	-	44.52	0.9581	17.15	-
	LRMC [16]	52.77	0.9705	26.16	194.3	42.44	0.9206	15.88	192.15	45.11	0.9602	18.66	197.14
	Wei et al. [18]	56.05	0.9840	30.08	177.66	44.12	0.9383	18.72	185.15	46.63	0.9696	20.11	180.25
	Fast Derain [20]	57.55	0.9965	31.16	**11.42**	45.93	0.9404	21.41	**9.5**	47.86	0.9701	21.87	**8.9**
	Sun et al. [21]	53.05	0.9790	29.11	202.23	43.88	0.9242	16.05	210.54	46.01	0.9632	19.56	217.16
	Proposed	**59.18**	**0.9971**	**32.41**	62.67	**46.31**	**0.9484**	**21.85**	64.82	**48.05**	**0.9704**	**22.04**	66.97
Highway	Rainy	53.34	0.9810	27.22	-	41.43	0.9007	16.56	-	44.48	0.9652	19.77	-
	LRMC [16]	55.06	0.9833	30.11	155.02	42.59	0.9207	17.11	162.89	45.07	0.9693	20.01	167.22
	Wei et al. [18]	56.59	0.9901	29.77	167.74	44.03	0.9569	20.53	172.22	46.13	0.9705	21.14	169.44
	Fast Derain [20]	58.84	0.9915	32.05	**10.05**	44.89	0.9655	21.35	**9.01**	47.05	0.9787	21.85	**11.07**
	Sun et al. [21]	55.92	0.9874	31.12	234.79	43.04	0.9309	18.14	241.20	45.80	0.9718	20.68	237.96
	Proposed	**59.32**	**0.9968**	**34.67**	57.52	**45.05**	**0.9685**	**21.56**	59.79	**47.55**	**0.9804**	**22.36**	59.12
Pedestrian	Rainy	46.51	0.9708	28.35	-	40.47	0.9445	19.18	-	42.19	0.9606	23.16	-
	LRMC [16]	48.22	0.9747	30.09	123.46	40.82	0.9486	19.77	110.62	42.67	0.9628	23.77	119.87
	Wei et al. [18]	52.78	0.9844	34.66	112.12	41.68	0.9639	22.23	94.21	43.89	0.9697	24.98	102.14
	Fast Derain [20]	55.48	0.9947	36.96	**12.05**	42.36	0.9842	23.01	**7.8**	45.02	0.9722	26.04	**11.88**
	Sun et al. [21]	51.44	0.9802	32.31	202.76	41.08	0.9511	20.75	212.23	43.04	0.9669	24.28	208.79
	Proposed	**56.26**	**0.9977**	**38.79**	27.96	**42.92**	**0.9877**	**23.89**	28.37	**45.65**	**0.9777**	**26.88**	28.49
Park	Rainy	54.48	0.9754	22.89	-	41.66	0.8937	16.58	-	43.37	0.9543	18.76	-
	LRMC [16]	55.11	0.9887	26.18	222.14	42.13	0.9177	17.77	227.13	45.88	0.9646	19.53	231.34
	Wei et al. [18]	56.45	0.9823	32.45	188.43	43.78	0.9642	19.09	195.26	45.19	0.9772	21.19	192.74
	Fast Derain [20]	57.66	0.9939	34.05	**10.34**	44.99	0.9744	19.89	**8.1**	46.91	0.9828	22.22	**9.56**
	Sun et al. [21]	55.92	0.9899	28.88	250.62	42.55	0.9272	18.03	255.57	46.39	0.9697	20.11	259.66
	Proposed	**58.01**	**0.9968**	**34.70**	58.55	**45.66**	**0.9759**	**20.28**	60.05	**47.02**	**0.9868**	**22.78**	61.37
Backdoor	Rainy	57.95	0.9698	26.56	-	47.82	0.9395	18.79	-	47.58	0.9493	17.41	-
	LRMC [16]	58.12	0.9702	29.19	286.86	48.09	0.9412	18.92	288.44	48.03	0.9533	18.28	292.32
	Wei et al. [18]	59.95	0.9815	31.14	201.76	50.56	0.9562	19.81	205.79	49.28	0.9681	19.38	211.65
	Fast Derain [20]	61.19	0.9909	33.23	**11.12**	52.12	0.9702	21.55	**9.6**	51.55	0.9799	21.55	**10.77**
	Sun et al. [21]	58.67	0.9744	30.22	311.21	48.86	0.9577	19.09	320.66	46.39	0.9697	18.74	302.21
	Proposed	**62.82**	**0.9956**	**34.19**	72.96	**53.37**	**0.9792**	**22.40**	71.25	**52.16**	**0.9840**	**22.88**	72.18
Bungalows	Rainy	52.08	0.9593	20.41	-	44.18	0.9496	17.54	-	45.47	0.9386	16.23	-
	LRMC [16]	52.84	0.9641	20.88	195.66	45.56	0.9509	18.01	201.72	46.12	0.9491	16.86	193.45
	Wei et al. [18]	54.07	0.9725	22.65	177.76	47.66	0.9571	18.68	180.79	46.97	0.9743	18.06	183.65
	Fast Derain [20]	55.29	0.9811	23.06	**9.62**	48.18	0.9615	19.78	**8.7**	47.78	0.9796	18.74	**11.97**
	Sun et al. [21]	53.55	0.9689	21.04	220.96	46.39	0.9548	18.22	212.85	46.52	0.9622	17.59	218.99
	Proposed	**56.08**	**0.9865**	**24.14**	40.56	**48.72**	**0.9697**	**20.14**	43.42	**48.17**	**0.9819**	**19.31**	41.98

* The computational time of proposed method is the total time taken for the video deraining and simultaneous moving object detection. For all other methods, it represents only the computational time of video deraining.
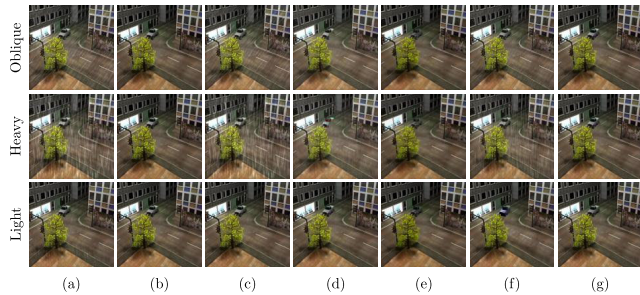

FIGURE 4. (a) Sample frames of 'truck' video with light, heavy and oblique rain streaks, (b) corresponding frames of rain-free video, deraining results obtained by (c) LRMC [16], (d) Wei *et al.* [18], (e) Fast Derain [20], (f) Sun *et al.* [21], (g) Proposed method.

FIGURE 5. (a) Sample frames of 'highway' video with light, heavy and oblique rain streaks, (b) corresponding frames of rain-free video, deraining results obtained by (c) LRMC [16], (d) Wei *et al.* [18], (e) Fast Derain [20], (f) Sun *et al.* [21], (g) Proposed method.
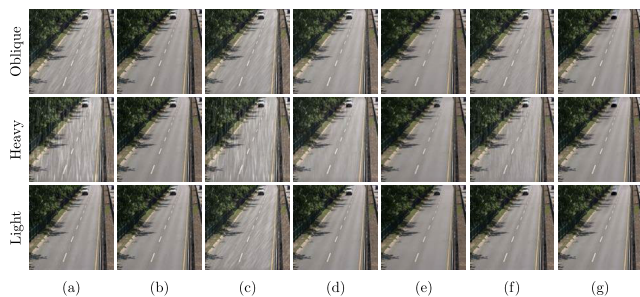

FIGURE 6. (a) Sample frames of rainy 'traffic', 'wall' and 'banana' videos, deraining results obtained by (b) LRMC [16], (c) Wei *et al.* [18], (d) Fast Derain [20], (e) Sun *et al.* [21], (f) Proposed method.

TABLE 3. Quantitative comparisons of rain removal results on real videos.

Video	Metric	LRMC [16]	Wei et al. [18]	Sun et al. [21]	Fast Derain [20]	Proposed Mehod
Traffic	NIQE	6.15	4.28	5.25	5.14	**2.21**
	BRISQUE	28.11	31.10	30.55	30.72	**35.42**
	SI	48.53	53.13	50.66	52.01	**59.14**
Wall	NIQE	16.32	8.2	11.21	5.94	**5.20**
	BRISQUE	20.35	26.85	24.94	32.72	**33.46**
	SI	23.11	28.82	25.22	30.12	**31.24**
Banana	NIQE	8.42	5.82	7.28	4.41	**3.22**
	BRISQUE	18.76	24.98	20.22	29.15	**31.14**
	SI	26.22	29.79	28.94	31.47	**32.22**
Backyard	NIQE	9.44	4.47	6.58	2.92	**2.05**
	BRISQUE	20.75	23.62	22.20	25.16	**28.78**
	SI	38.67	44.84	40.23	48.26	**49.62**

method remarkably outperformed the state of the art methods in terms of chosen quality metrics. However, the execution time for Fast Derain [20] is very less compared with other methods including pro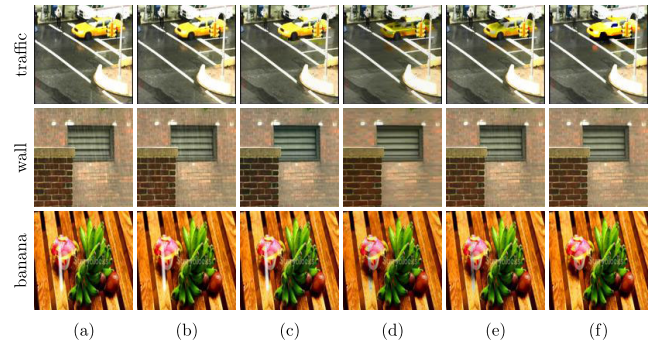posed method. However, the proposed method can provide dual solutions such as deraining and MOD in a reasonable computational time. Fig 4 and 5 represent derained results and comparison of sample frames taken from synthetic videos named 'truck' and 'highway' respectively with three different rain patterns as light, heavy and oblique. It can be obviously seen that rain removal task is effectively carried out by the proposed method.
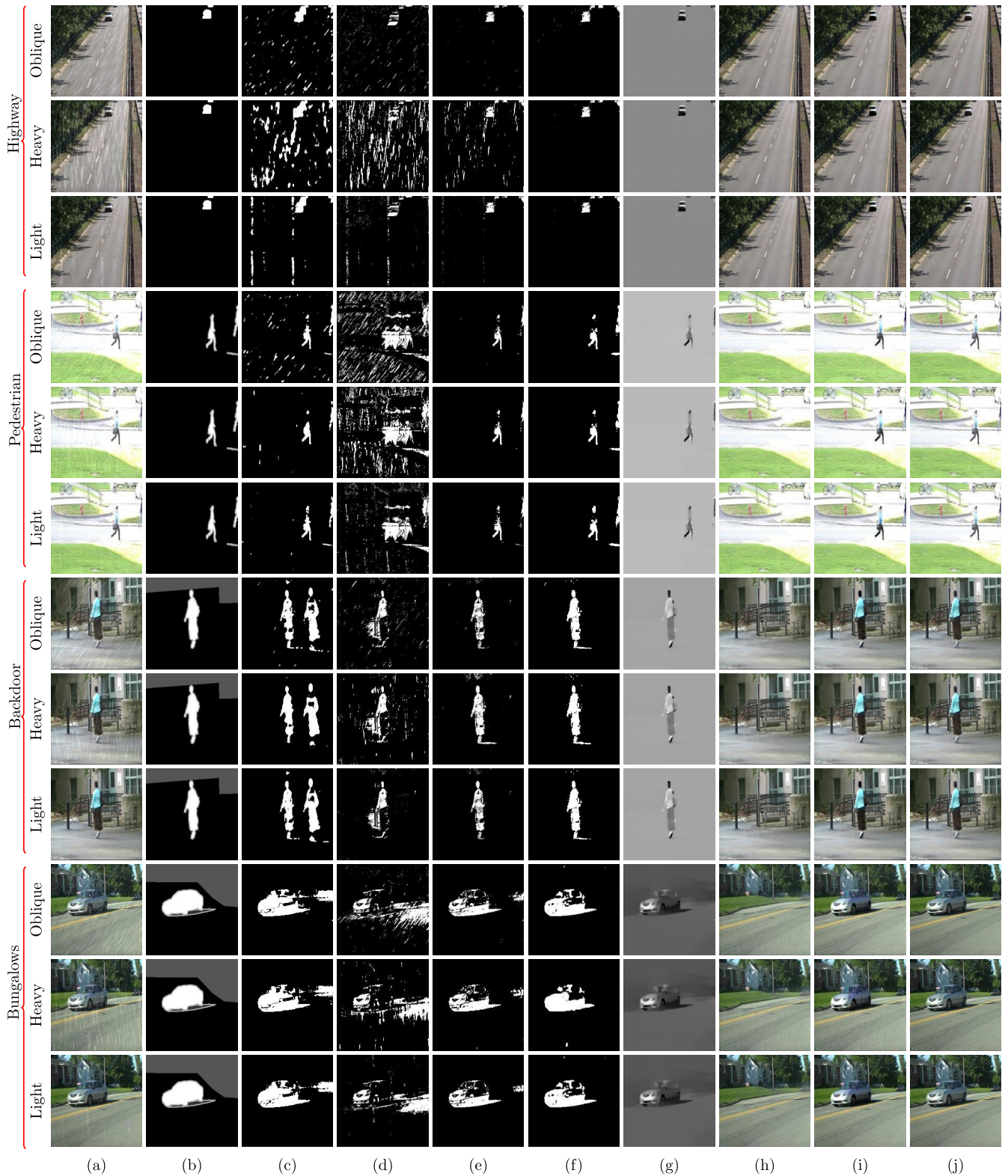
FIGURE 7. (a) Sample frames of rainy vidoes named 'highway', 'pedestrain', 'Backdoor' and 'Bungalows' with light, heavy and oblique rain streaks, (b) Corresponding frames of groundtruth, foreground masks obtained by (c) Sajid *et al.* [26], (d) Chen *et al.* [25], (e) Shijila *et al.* [29], (f) Proposed method, (g)-(i) Corresponding frames of foreground, background and derained result obtained by proposed method, (j) Original frame.

2) REAL DATA

Natural rainy videos such as 'traffic', 'wall', 'banana' and 'backyard' of size $256 \times 256 \times 3 \times 100$ are collected from various database including 'Storyblocks' and are utilized to verify how the proposed method performs rain removal for real data. To substantiate the ability of proposed method to

TABLE 4. Quantitative comparison of proposed method with different object detection algorithms.

Video	Rain streaks	Sajid et al. [26]			Chen et al. [25]			Shijila et al. [29].			Proposed method		
		f_0	f_1	f_j	f_0	f_1	f_j	f_0	f_1	f_j	f_0	f_1	f_j
Highway	Light	0.5155	0.6485	0.5744	0.8780	0.7421	0.8044	0.9504	0.8322	0.8874	0.9795	0.8899	**0.9326**
	Heavy	0.5026	0.4005	0.4458	0.6988	0.5442	0.6119	0.8905	0.6707	0.7651	0.9621	0.8630	**0.9099**
	Oblique	0.5205	0.6055	0.5598	0.8103	0.6898	0.7452	0.9604	0.8794	0.9181	0.9764	0.9001	**0.9367**
Pedestrian	Light	0.8814	0.7929	0.8348	0.1512	0.1054	0.1242	0.9371	0.8506	0.8918	0.9438	0.8634	**0.9018**
	Heavy	0.8289	0.7592	0.7925	0.1041	0.1021	0.1031	0.9265	0.8422	0.8823	0.9402	0.8566	**0.8965**
	Oblique	0.7468	0.6028	0.6671	0.0888	0.0764	0.0821	0.9273	0.8429	0.8831	0.9320	0.8499	**0.8891**
Backdoor	Light	0.3541	0.2537	0.2956	0.6214	0.4993	0.5537	0.9581	0.8632	0.9082	0.9618	0.8848	**0.9217**
	Heavy	0.3420	0.2327	0.2770	0.6201	0.3146	0.4174	0.9589	0.8799	0.9177	0.9604	0.8817	**0.9194**
	Oblique	0.3456	0.2153	0.2653	0.6077	0.4011	0.4832	0.9544	0.8770	0.9141	0.9599	0.8802	**0.9183**
Bungalows	Light	0.8116	0.6982	0.7506	0.4939	0.3996	0.4418	0.8657	0.7399	0.7979	0.9413	0.7719	**0.8482**
	Heavy	0.8285	0.7015	0.7597	0.4670	0.3705	0.4132	0.8445	0.7311	0.7837	0.9524	0.7733	**0.8536**
	Oblique	0.7915	0.6932	0.7391	0.4581	0.3282	0.3824	0.8512	0.7431	0.7935	0.9515	0.7726	**0.8528**

remove rain streaks on real videos, visual quality parameters called Natural Image Quality Evaluator (NIQE) [48], Blind or Referenceless Image Spatial Quality Evaluator (BRISQUE) [49] and Sharpness Index (SI) [50] are calculated for each frame and the corresponding mean values of such parameters for whole video are entered in Table 3. Smaller values of NIQE and higher values of BRISQUE and SI indicate better performance. As seen from Table 3, the proposed method gives better results for all the four real videos. Fig 6 shows the visual comparison of our method with existing methods as [16], [18], [20] and [21]. It is clear that our method outperforms the other state of the art methods.

B. OBJECT DETECTION PERFORMANCE ANALYSIS

In order to analyse the performance effectiveness of our method for detecting moving objects, we have done a comparison with recent object detection methods such as [25], [26] and [29]. Videos used for comparison, such as 'highway', 'pedestrian', 'backdoor' and 'bungalows' are collected from CD.net dataset and simulated rain streaks with different patterns are added. The performance evaluation for background as well as foreground detection can be measured in terms of parameters such as f_0 measure, f_1 measure and f_j measure [30], [51]. These parameters can be formulated as follows.

$$f_0 \; measure = \frac{2R_0 P_0}{R_0 + P_0} \tag{29}$$

$$f_1 \; measure = \frac{2R_1 P_1}{R_1 + P_1} \tag{30}$$

$$f_j \; measure = \frac{2f_0 f_1}{f_0 + f_1} \tag{31}$$

where,

$$R_0 = \frac{TN}{TN + FP}, \quad R_1 = \frac{TP}{TP + FN} \tag{32}$$

and

$$P_0 = \frac{TN}{TN + FN}, \quad P_1 = \frac{TP}{TP + FP} \tag{33}$$

Here, FP, FN, TP and TN indicate false positives, false negatives, true positives, and true negatives, respectively. By using the above expressions, the selected parameters are calculated for each frame of the videos and the corresponding mean values are entered in Table 4. It is clear that the proposed method shows better performance than the state of the art methods

with respect to the selected quality metrics. Fig 7 shows the visual comparison of our method with existing object detection methods. From Fig 7, it is seen that the proposed method can efficiently detect the foreground irrespective of the nature of rain streaks.

V. CONCLUSION

A new method for video deraining and simultaneous moving object detection is proposed and implemented in this paper. A 3-way tensor decomposition model is introduced for the retrieval of rain-free background and extraction of foreground from the rainy video data. A new operator termed as Slice Rotational Total Variation (SRTV) norm is formulated for transforming the all different rain patterns into unique dot-like pattern. SRTV combined with Weighted Tensor Nuclear Norm (WTNN) regularizations on rainy and low rank components are employed for effectively removing the rain streak outlier from background. Moreover, l_1 norm minimization along with TTV regularizers are incorporated for the proper detection of moving objects from rainy video. The quantitative comparison of the proposed method for deraining performance with the existing methods such as, LRMC method [16], Wei et al. [18], Fast Derain [20], Sun et al. [21] was done. Moreover, the proposed method was compared for analysing the object detection performance with the recent MOD techniques such as, Javed et al. [26], Chen et al. [25] and Shijila et al. [29]. In both the above cases, our method brought good results. The main contributions of this work are the formulation of the new operator, SRTV norm for addressing the all types of rain patterns and the simultaneous video deraining and moving object detection by iteratively solving a single optimization model without any training phase. In this work, we have considered only videos with static backgrounds. In future, we can reformulate the optimization model for addressing videos with dynamic backgrounds.

REFERENCES

[1] K. Garg and S. K. Nayar, "Detection and removal of rain from videos," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, vol. 1, Jun. 2004, pp. 1–8.

[2] K. Garg and S. K. Nayar, "Vision and rain," *Int. J. Comput. Vis.*, vol. 75, no. 1, pp. 3–27, Jul. 2007.

[3] S. Starik and M. Werman, "Simulation of rain in videos," in *Proc. Texture Workshop ICCV*, vol. 2, 2003, pp. 406–409.

[4] X. Zhang, H. Li, Y. Qi, W. Leow, and T. Ng, "Rain removal in video by combining temporal and chromatic properties," in *Proc. IEEE Int. Conf. Multimedia Expo*, Jul. 2006, pp. 461–464.

[5] W.-J. Park and K.-H. Lee, "Rain removal using Kalman filter in video," in *Proc. Int. Conf. Smart Manuf. Appl.*, Apr. 2008, pp. 494–497.

[6] N. Brewer and N. Liu, "Using the shape characteristics of rain to identify and remove rain from video," in *Proc. Joint IAPR Int. Workshops Stat. Techn. Pattern Recognit. (SPR) Struct. Syntactic Pattern Recognit. (SSPR)*. Berlin, Germany: Springer, 2008, pp. 451–458.

[7] P. C. Barnum, S. Narasimhan, and T. Kanade, "Analysis of rain and snow in frequency space," *Int. J. Comput. Vis.*, vol. 86, nos. 2–3, pp. 256–274, Jan. 2010.

[8] Y.-H. Fu, L.-W. Kang, C.-W. Lin, and C.-T. Hsu, "Single-frame-based rain removal via image decomposition," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2011, pp. 1453–1456.

[9] D.-Y. Chen, C.-C. Chen, and L.-W. Kang, "Visual depth guided color image rain streaks removal using sparse coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 24, no. 8, pp. 1430–1455, Aug. 2014.

[10] L.-W. Kang, C.-W. Lin, and Y.-H. Fu, "Automatic single-image-based rain streaks removal via image decomposition," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 1742–1755, Apr. 2012.

[11] S.-H. Sun, S.-P. Fan, and Y.-C.-F. Wang, "Exploiting image structural similarity for single image rain removal," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Oct. 2014, pp. 4482–4486.

[12] Y. Li, R. T. Tan, X. Guo, J. Lu, and M. S. Brown, "Rain streak removal using layer priors," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 2736–2744.

[13] Z. Mi, J. Shang, H. Zhou, and M. Wang, "Image fusion-based video deraining using sparse representation," *Electron. Lett.*, vol. 52, no. 18, pp. 1528–1529, Sep. 2016.

[14] W. Ren, J. Tian, Z. Han, A. Chan, and Y. Tang, "Video desnowing and deraining based on matrix decomposition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 2838–2847.

[15] L. Zhu, C.-W. Fu, D. Lischinski, and P.-A. Heng, "Joint bi-layer optimization for single-image rain streak removal," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 2526–2534.

[16] J.-H. Kim, J.-Y. Sim, and C.-S. Kim, "Video deraining and desnowing using temporal correlation and low-rank matrix completion," *IEEE Trans. Image Process.*, vol. 24, no. 9, pp. 2658–2670, Sep. 2015.

[17] T.-X. Jiang, T.-Z. Huang, X.-L. Zhao, L.-J. Deng, and Y. Wang, "A novel tensor-based video rain streaks removal approach via utilizing discriminatively intrinsic priors," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 4057–4066.

[18] W. Wei, L. Yi, Q. Xie, Q. Zhao, D. Meng, and Z. Xu, "Should we encode rain streaks in video as deterministic or stochastic?" in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 2516–2525.

[19] Y.-T. Wang, X.-L. Zhao, T.-X. Jiang, L.-J. Deng, T.-H. Ma, Y.-T. Zhang, and T.-Z. Huang, "A total variation and group sparsity based tensor optimization model for video rain streak removal," *Signal Process., Image Commun.*, vol. 73, pp. 96–108, Apr. 2019.

[20] T.-X. Jiang, T.-Z. Huang, X.-L. Zhao, L.-J. Deng, and Y. Wang, "FastDeRain: A novel video rain streak removal method using directional gradient priors," *IEEE Trans. Image Process.*, vol. 28, no. 4, pp. 2089–2102, Apr. 2019.

[21] Z. Sun, S. Xiong, and R. Wen Liu, "Directional regularized tensor modeling for video rain streaks removal," 2019, *arXiv:1902.07090*. [Online]. Available: http://arxiv.org/abs/1902.07090

[22] E. J. Candès, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis?" *J. ACM*, vol. 58, no. 3, pp. 11–20, 2011.

[23] X. Zhou, C. Yang, and W. Yu, "Moving object detection by detecting contiguous outliers in the low-rank representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 3, pp. 597–610, Mar. 2013.

[24] X. Cao, L. Yang, and X. Guo, "Total variation regularized RPCA for irregularly moving object detection under dynamic background," *IEEE Trans. Cybern.*, vol. 46, no. 4, pp. 1014–1027, Apr. 2016.

[25] L. Chen, Y. Liu, and C. Zhu, "Iterative block tensor singular value thresholding for extraction of lowrank component of image data," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Mar. 2017, pp. 1862–1866.

[26] S. Javed, T. Bouwmans, and S. K. Jung, "Stochastic decomposition into low rank and sparse tensor for robust background subtraction," in *Proc. 6th Int. Conf. Imag. Crime Prevention Detection (ICDP)*, 2015, pp. 1–6.

[27] A. J. Tom and S. N. George, "Tensor total variation regularized moving object detection for surveillance videos," in *Proc. Int. Conf. Signal Process. Commun. (SPCOM)*, Jul. 2018, pp. 327–331.

[28] B. Shijila, A. J. Tom, and S. N. George, "Moving object detection by low rank approximation and l1-TV regularization on RPCA framework," *J. Vis. Commun. Image Represent.*, vol. 56, pp. 188–200, Oct. 2018.

[29] B. Shilila, A. J. Tom, and S. N. George, "Simultaneous denoising and moving object detection using low rank approximation," *Future Gener. Comput. Syst.*, vol. 90, pp. 198–210, Jan. 2019.

[30] A. J. Tom and S. N. George, "Video completion and simultaneous moving object detection for extreme surveillance environments," *IEEE Signal Process. Lett.*, vol. 26, no. 4, pp. 577–581, Apr. 2019.

[31] P. Comon, "Tensors: A brief introduction," *IEEE Signal Process. Mag.*, vol. 31, no. 3, pp. 44–53, May 2014.

[32] J. D. Carroll and J.-J. Chang, "Analysis of individual differences in multidimensional scaling via an N-way generalization of 'Eckart-Young' decomposition," *Psychometrika*, vol. 35, no. 3, pp. 283–319, Sep. 1970.

[33] E. Kernfeld, M. Kilmer, and S. Aeron, "Tensor–tensor products with invertible linear transforms," *Linear Algebra Appl.*, vol. 485, pp. 545–570, Nov. 2015.

[34] M. E. Kilmer, K. Braman, N. Hao, and R. C. Hoover, "Third-order tensors as operators on matrices: A theoretical and computational framework with applications in imaging," *SIAM J. Matrix Anal. Appl.*, vol. 34, no. 1, pp. 148–172, Jan. 2013.

[35] L. De Lathauwer, B. De Moor, and J. Vandewalle, "A multilinear singular value decomposition," *SIAM J. Matrix Anal. Appl.*, vol. 21, no. 4, pp. 1253–1278, 2000.

[36] C. D. Martin, R. Shafer, and B. LaRue, "An order-p tensor factorization with applications in imaging," *SIAM J. Sci. Comput.*, vol. 35, no. 1, pp. A474–A490, Jan. 2013.

[37] B. Madathil and S. N. George, "DCT based weighted adaptive multi-linear data completion and denoising," *Neurocomputing*, vol. 318, pp. 120–136, Nov. 2018.

[38] B. Madathil and S. N. George, "Twist tensor total variation regularized-reweighted nuclear norm based tensor completion for video missing area recovery," *Inf. Sci.*, vol. 423, pp. 376–397, Jan. 2018.

[39] Z. Zhang, G. Ely, S. Aeron, N. Hao, and M. Kilmer, "Novel methods for multilinear data completion and de-noising based on tensor-SVD," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 3842–3849.

[40] M. Baburaj and S. N. George, "Tensor based approach for inpainting of video containing sparse text," *Multimedia Tools Appl.*, vol. 78, no. 2, pp. 1805–1829, Jan. 2019.

[41] H. Wang, F. Nie, and H. Huang, "Low-rank tensor completion with spatio-temporal consistency," in *Proc. 28th AAAI Conf. Artif. Intell.*, 2014, pp. 2846–2852.

[42] C. Lu, J. Feng, Y. Chen, W. Liu, Z. Lin, and S. Yan, "Tensor robust principal component analysis: Exact recovery of corrupted low-rank tensors via convex optimization," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 5249–5257.

[43] W. Hu, Y. Yang, W. Zhang, and Y. Xie, "Moving object detection using tensor-based low-rank and saliently fused-sparse decomposition," *IEEE Trans. Image Process.*, vol. 26, no. 2, pp. 724–737, Feb. 2017.

[44] S. Boyd, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, 2010.

[45] S. Yang, J. Wang, W. Fan, X. Zhang, P. Wonka, and J. Ye, "An efficient ADMM algorithm for multidimensional anisotropic total variation regularization problems," in *Proc. 19th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining (KDD)*, 2013, pp. 641–649.

[46] M. Sakurai, S. Kiriyama, T. Goto, and S. Hirano, "Fast algorithm for total variation minimization," in *Proc. 18th IEEE Int. Conf. Image Process.*, Sep. 2011, pp. 1461–1464.

[47] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

[48] A. Mittal, R. Soundararajan, and A. C. Bovik, "Making a 'completely blind' image quality analyzer," *IEEE Signal Process. Lett.*, vol. 20, no. 3, pp. 209–212, Mar. 2013.

[49] A. Mittal, A. K. Moorthy, and A. C. Bovik, "No-reference image quality assessment in the spatial domain," *IEEE Trans. Image Process.*, vol. 21, no. 12, pp. 4695–4708, Dec. 2012.

[50] G. Blanchet and L. Moisan, "An explicit sharpness index related to global phase coherence," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Mar. 2012, pp. 1065–1068.

[51] D. Giveki, G. A. Montazer, and M. A. Soltanshahi, "Atanassov's intuitionistic fuzzy Histon for robust moving object detection," *Int. J. Approx. Reasoning*, vol. 91, pp. 80–95, Dec. 2017.

BAIJU P. S. (Graduate Student Member, IEEE) received the B.Tech. degree in electronics and communication engineering and the M.Tech. degree in signal processing from the University of Kerala, India, in 2005 and 2013, respectively. He is currently pursuing the Ph.D. degree with the National Institute of Technology Calicut. His research interests include image and video processing, low rank approximation, video deweathering.

SUDHISH N. GEORGE (Member, IEEE) received the B.Tech. degree from the Mar Athanasius College of Engineering, Mahatma Gandhi University, Kottayam, India, in 2004, the M.Tech. degree from the College of Engineering Thiruvananthapuram, University of Kerala, Thiruvananthapuram, India, in 2007, and the Ph.D. degree from the National Institute of Technology Calicut, Kozhikode, India, in 2014. He is currently an Assistant Professor with the Department of Electronics and Communication Engineering, National Institute of Technology Calicut. His research interests include sparse signal processing and computer vision.

. . .