Video Smoke Removal based on Low-rank Tensor Completion via Spatial-Temporal Continuity Constraint

Hu Zhu[1] | Guoxia Xu[2] | Lu Liu[1] | Lizhen Deng*[3]

[1]Jiangsu Province Key Lab on Image Processing and Image Communication, Nanjing University of Posts and Telecommunication, Nanjing, China

[2]Department of Computer Science, Norwegian University of Science and Technology, Gjovik, Norway

[3]National Engineering Research Center of Communication and Network Technology, Nanjing University of Posts and Telecommunication, Nanjing, China

Correspondence
*Lizhen Deng, Email: alicedenglzh@gmail.com

Summary

Smoke has a very bad effect on the outdoor vision system. Not only are the videos with poor visual effects obtained, but also the quality and structure of the videos are reduced. In this paper, we propose a video smoke removal method based on low-rank tensor completion via spatial-temporal continuity constraint. The proposed method is based on the smoke mixing model and consider the sparseness of smoke and the global and local consistency of clean video. Then, the optimal solution of the smoke removal algorithm model is quickly realized by the Alternating Direction Method of Multiplier. Finally, we evaluate the experiment results of real-world data and simulated data from the visual effects and objective indicators. And the experiment results show that our proposed algorithm can achieve better smoke removal results.

KEYWORDS:
Spatial continuity, smoke removal, tensor completion

1 | INTRODUCTION

Smoke would reduce the visibility of video so that smoke removal is of great significance. After that, it has practical applications in target recognition[1] and tracking[2][3][4], disaster processing[5] and so on. Therefore, video smoke removal is an important issue in outdoor vision systems and has extremely important practical significance.

However, researchers have focused more on smoke detection, with little consideration given to smoke removal. And smoke detection[6][7][8][9] is mainly used to prevent fires and ensure people safety. With the development of computer vision, smoke removal algorithms are slowly emerging and have received widespread attention from researchers.Until now, most of the current smoke removal technologies and algorithms are based on several existing frameworks or models for a small amount of improvement. These algorithms only enhance the original models to further reduce the proportion of smoke in photos or videos.

It was originally based on wavelet transform proposed by Cheng et al. [10]. The method was based on the characteristics of space-frequency domain and multi-resolution analysis of wavelet transform. And this method was mainly effective for smoke removal of solid rocket motor images, but not suitable for smoke images of other scenes. After, Lu et al. [11] proposed a fast joint trigonometric filtering dehazing algorithm. Describes a method for enhancing underwater images through image defogging. Yu et al. [12] proposed a novel removal algorithm to single image based on a self-learning framework and structured sparse representation. This algorithm can also remove rain streaks component while preserving most of geometric details. However, the processed picture would be a little fuzzy. Xu et al. [13] proposed a smoke removal method based on spectral separation technology to remove smoke from optical remote sensing images. The algorithm first detected the smoke component by using a spectral blending analysis to generate a sub-pixel smoke fraction mask, and then subtracted the smoke component from each smoke pixel. Finally, the attenuated signal is recovered by re-adjusting the abundance of the classes presented in each pixel. This method is highly
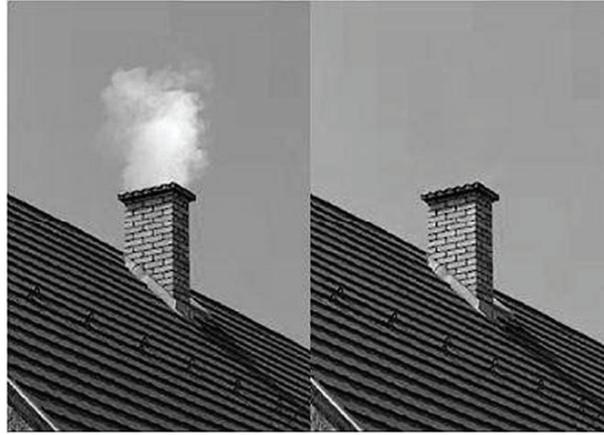
FIGURE 1 Smoke removal example applied on a simulated video frame. Left: A smoky video frame. Right: Smoke removal by our algorithm

feasible and does not depend on other auxiliary data. Enlightened by Nishino et al.[14], Kotwal et al.[15] and Baild et al.[16] suggest Bayesian inference-based laparoscopy image desmoking, accompanied with denoising and specularity removal in addition to denoising, respectively. The authors present the undamaged image as a Markov random field, and use the maximum posterior estimation to obtain the enhanced version. Chen et al.[17] considered that there were some limitations in the method of dehaze for smoke removal. Thus, they proposed a new smoke removal algorithm based on machine learning and smoke detection technology. The training algorithm corresponding to the model of each color channel was proposed based on the difference in the intensity distribution of the different color channels observed in the smoke image and the smoke is removed from the RGB channel, respectively. With the development of artificial intelligence(AI), deep learning[18] and reinforcement learning[19] has adopted into many fields. Lu et al.[20]developed an intelligent learning model called "Brain Intelligence (BI)". Furthermore, the integration of the ocean network with artificial intelligence has become a topic of increasing interest for oceanology researchers. The cognitive ocean network (CONet)[21] will become the mainstream of future ocean science and engineering developments. Specifically, Bolkar et al.[5] proposed a smoke removal solution based on deep neural network, which can improve the quality of surgical video frames in real time. In addition, Li et al.[22] explored the contrast of the Laplacian sharpening first histogram equalization algorithm to clear the smoke image, and verified the algorithm from both subjective and objective perspectives.

In addition to smoke extraction, there are many low-level optimizations. Liu et al.[23] and others proposed a novel low-rank representation (LRR) method to solve the single-person sample (SSPP) problem in face recognition by using both subspace and block sparse structure. By representing both the subspace and the block sparse structure as a unified objective function, the differentiation of the representation can be further promoted. Gao et al.[24] proposed a method called Multi-layer Locally Constrained Structure Orthogonal Procrustes Regression (MLCSOPR), which learns the pose discriminative representation features and enhances the consistency between LR and HR image spaces. It is effective in cross-resolution face matching with pose changes. Moreover, many works have exploited the adaptive relaxation methods[25] to handle the sparsity and low rank for specific works, Yu et al.[26] made the $L_{2,1}$ norm to explore the exact approximation for the affinity feature. Motivated by these discussions, We plan to use the low rank and sparsity property to achieve the desmoke task from the perspective of optimization.

Figure 1 shows the smoke removal result of one of our simulated data. It can be seen that our algorithm has achieved a very good visual effect. In the subsequent experimental part, we will measure this result from the objective evaluation index of the image. The smoke removal algorithm in this paper is mainly aiming at the smoke removal under the outdoor environment. In this paper, we propose a novel video smoke removal algorithm model based on low-rank tensor completion via temporal continuity constraint. Firstly, we consider that the clean video is low-rank. This paper is the first time to introduce the concept of low rank into the smoke removal algorithm. In general, this concept is used in many other fields. For example, the hyper-spectral image denoise[27]. Secondly, we consider that the clean video is smooth and continuous in spatial-temporal dimension like the work in[28], we introduce the concept of spatial-temporal continuity constraint[29]. into the smoke removal algorithm. Finally, the Alternating Direction Method of Multiplier can accelerate the proposed smoke removal algorithm in this paper. It has an extremely large advantage for the solution of convex optimization.

Based the above aspects, we proposed a novel video smoke removal algorithm model based on low-rank tensor completion via temporal continuity constraint. The specific algorithm will be explained in detail in the following sections. In this paper, there are mainly three contributions:

- This paper introduces the concept of low rank into the proposed smoke removal algorithm.

- This paper introduces the concept of spatial-temporal continuity constraint into the proposed smoke removal algorithm.

- This paper uses the Alternating Direction Method of Multipliers (ADMM) to optimize the solution of the proposed smoke removal algorithm model.

The paper is organized as follows. In Section 2, we review our proposed smoke removal model theory background and construct a smoke removal model based on this background. The proposed smoke removal model based on low rank tensor completion via temporal continuity constraint is presented in Section 3. The alternating direction method of Multipliers is adopted to optimize the whole algorithm model in Section 4. In Section 5, some experimental results and discussion are presented. Finally, conclusions are drawn in Section 6.

2 | MOTIVATION

In the most advanced video-based smoke detection and removal methods, the representation of smoke is primarily dependent on the visual information in the current image frame. Although these methods can deal with the problem of vision-based smoke detection and removal to a certain extent, there are common problems between them, which are that they only use various visual information of the smoke area. A notable feature of smoke is transparency. Regardless of the color of the smoke, we can observe that the background is covered by smoke. In other words, if the features of the smoke are extracted based on the image with light smoke, the visual information of the original background will inevitably affect the extracted features. These extracted features reflect the true visual characteristics of the smoke and therefore represent smoke more reliably. Based on the above observation, in order to demonstrate the superiority of smoke detection and removal using smoke components, we introduce an image separation proposed by Tian et al. [27] to detect and remove smoke in a fixed captured image sequence.

A hybrid image model was constructed, primarily a linear combination of background and smoke components, while also defining smoke opacity to describe the contribution of smoke components to the composite image. And the hybrid image model is written as:

$$I = \alpha S + (1 - \alpha)B + n \tag{1}$$

where $\alpha \in [0, 1]$ represents the contribution of pure smoke image to the input image. B represents the background of the image, S represents the pure smoke, n represents the additional noise.

According to such a model, algorithm optimization is performed to solve the smoke opacity and smoke components. The algorithm of this chapter is based on this model. By analyzing the characteristics of smoke sparse in video images and combining the global and local consistency of clean video, a video smoke removal method based on spatial-temporal consistency is proposed. The experimental results of the simulation data verify the effectiveness of the algorithm in this chapter.

3 | PROPOSED MODEL

We promote the Equation (1) to the smoke video, so the smoky video can be presented as:

$$\mathcal{I} = \alpha S + (1 - \alpha)\mathcal{B} + \mathcal{N} \tag{2}$$

where $\mathcal{I}$, S, $\mathcal{B}$ represent the smoky video, clean video, and pure smoke. They are all three order tensor and have the same size of $m \times n \times t$. In addition, the original image needs to be removed additional noise by Gauss filter.

In this section, our smoke removal algorithm based on spatial continuity method is presented. According to the characteristics of smoke sparse in video images and combining the global and local consistency of clean video, the smoke removal model can be written as:

$$\min_{\mathcal{B}} \tau \|\mathcal{I} - \alpha S - (1 - \alpha)\mathcal{B}\|_2^2 + rank(\mathcal{B}) + \mu\|S\|_1 + \frac{\lambda}{2}\|D_i\mathcal{B}\|_F^2 \tag{3}$$

The first item of the model is the fidelity item, the second represents the low rank of the clean video, the third represents the sparseness of the smoke, and the last represents the local consistency of the clean video, where D_t is an auxiliary matrix. Then the entire smoke removal model was constructed.

4 | OPTIMIZATION

For the smoke removal model (3) proposed in last chapter, the Alternating Direction Method of Multipliers (ADMM) is used to optimize the solution. First, we introduce auxiliary variables, and $\mathcal{X}$, $\mathcal{Y}$ and add some constraints $\mathcal{Z}$, then model (3) can be written as:

$$
\begin{aligned}
\min_{\mathcal{B}} \ & \tau \|\mathcal{X}\|_2^2 + \|\mathcal{Y}\|_* + \frac{\lambda}{2}\|\mathcal{Z}\|_F^2 + \mu\|\mathcal{S}\|_1 \\
s.t. \ & \mathcal{X} = \mathcal{I} - \alpha\mathcal{S} - (1-\alpha)\mathcal{B}, \\
& \mathcal{Y} = \mathcal{B} \\
& \mathcal{Z} = D_t\mathcal{B}
\end{aligned}
\tag{4}
$$

Then the augmented Lagrangian function of model (4) can be obtained as:

$$
\begin{aligned}
L(\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \mathcal{S}, \mathcal{B}) =& \tau \|\mathcal{X}\|_2^2 + \|\mathcal{Y}\|_* + \frac{\lambda}{2}\|\mathcal{Z}\|_F^2 + \mu\|\mathcal{S}\|_1 \\
& + <\Lambda_1, \mathcal{X} - \mathcal{I} + \alpha\mathcal{S} + (1-\alpha)\mathcal{B}> \\
& + \frac{\beta_1}{2}\|\mathcal{X} - \mathcal{I} + \alpha\mathcal{S} + (1-\alpha)\mathcal{B}\|_F^2 \\
& + <\Lambda_2, \mathcal{Y} - \mathcal{B}> + \frac{\beta_2}{2}\|\mathcal{Y} - \mathcal{B}\|_F^2 \\
& + <\Lambda_3, \mathcal{Z} - D_t\mathcal{B}> + \frac{\beta_3}{2}\|\mathcal{Z} - D_t\mathcal{B}\|_F^2
\end{aligned}
\tag{5}
$$

where β_1, β_2, β_3 are parameter greater than zero, and Λ_1 , Λ_2, Λ_3 are augmented Lagrangian multipliers.

Then the model (3) is transformed from solving the clean video into a clean video by coordinating the auxiliary variables, and get the global solution. Then the solution process can be divided into the following parts:

- About the $\mathcal{X}$ sub-problem:

$$
\min_{\mathcal{X}} \tau \|\mathcal{X}\|_2^2 + <\Lambda_1, \mathcal{X}> + \frac{\beta_1}{2}\|\mathcal{X} - \mathcal{I} + \alpha\mathcal{S} + (1-\alpha)\mathcal{B}\|_F^2
\tag{6}
$$

So the optimal solution of $\mathcal{X}$:

$$
\mathcal{X}^{k+1} = \frac{\beta_1(\mathcal{I} - \alpha\mathcal{S}^k - (1-\alpha)\mathcal{B}^k) - \Lambda_1^k}{2\tau + \beta_1}
\tag{7}
$$

- About the $\mathcal{Y}$ sub-problem:

$$
\min_{\mathcal{Y}} \ \|\mathcal{Y}\|_* + <\Lambda_2, \mathcal{Y}> + \frac{\beta_2}{2}\|\mathcal{Y} - \mathcal{B}\|_F^2
\tag{8}
$$

Here the Tensor Nuclear Norm is used to solve the low rank problem. And the above formulation can be simplified to:

$$
\min_{\mathcal{Y}} \ \|\mathcal{Y}\|_* + <\Lambda_2, \mathcal{Y}> + \frac{\beta_2}{2}\|\mathcal{Y} - \mathcal{B}\|_F^2
\tag{9}
$$

So the optimal solution of $\mathcal{Y}$:

$$
\mathcal{Y}^{k+1} = \frac{1}{3}\sum_{i=1}^{3}\text{fold}_i\left[\text{SVD}_+\left(\text{Unfold}_i(\mathcal{B}^k - \frac{\Lambda_2^k}{\beta_2})\right)\right]
\tag{10}
$$

- About the $\mathcal{Z}$ sub-problem:

$$
\min_{\mathcal{Z}} \ \frac{\lambda}{2}\|\mathcal{Z}\|_F^2 + <\Lambda_3, \mathcal{Z}> + \frac{\beta_3}{2}\|\mathcal{Z} - D_t\mathcal{B}\|_F^2
\tag{11}
$$

This is a question about the least squares method, which is easy to solve. So the optimal solution of $\mathcal{Z}$ is

$$
\mathcal{Z}^{k+1} = \frac{\beta_3 D_t\mathcal{B}^k - \Lambda_3^k}{\lambda + \beta_3}
\tag{12}
$$

- About the $\mathcal{S}$ sub-problem:

$$\min_{\mathcal{S}} \quad \mu \|\mathcal{S}\|_1 + \frac{\beta_1}{2} \|\mathcal{X} - \mathcal{I} + \alpha\mathcal{S} + (1-\alpha)\mathcal{B}\|_F^2 \tag{13}$$

The solution of the L1 norm can be handled by a closed soft thresholding. So the optimal solution of $\mathcal{S}$ is

$$\alpha\mathcal{S}^{k+1} = \mathrm{Shrink}_{\frac{\mu}{\alpha\beta_1}}\left(\mathcal{I} - \mathcal{X} - (1-\alpha)\mathcal{B} + \frac{\Lambda_1^k}{\beta_1} \right) \tag{14}$$

- About the $\mathcal{B}$ sub-problem:

$$\begin{aligned}
\min_{\mathcal{B}} \quad &\left\langle (1-\alpha)\Lambda_1 - \Lambda_2 - D_t^T\Lambda_3, \mathcal{B} \right\rangle \\
&+ \frac{\beta_1}{2} \|\mathcal{X} - \mathcal{I} + \alpha\mathcal{S} + (1-\alpha)\mathcal{B}\|_F^2 \\
&+ \frac{\beta_2}{2} \|\mathcal{Y} - \mathcal{B}\|_F^2 + \frac{\beta_3}{2} \|\mathcal{Z} - D_t\mathcal{B}\|_F^2
\end{aligned} \tag{15}$$

Involving the difference operator, we need to solve with the help of fast Fourier transform (FFT) and inverse fast Fourier transform (IFFT) to get the solution of $\mathcal{B}$:

$$\mathcal{B}^{k+1} = \mathcal{F}^{-1}\left(\frac{\mathcal{F}(\mathcal{M}_1^k)}{\mathcal{F}(\mathcal{M}_2^k)} \right) \tag{16}$$

where $\mathcal{F}(\mathcal{M})$ represents Fast Fourier Transform of $\mathcal{M}$, and $\mathcal{F}^{-1}(\mathcal{M})$ represents Inverse Fast Fourier Transform of $\mathcal{M}$, and

$$\mathcal{M}_1 = D_t^T\Lambda_3 + \Lambda_2 - (1-\alpha)\Lambda_1 + \beta_1(1-\alpha)(\mathcal{I} - \mathcal{X} - \alpha\mathcal{S}) + \beta_2\mathcal{Y} + \beta_3 D_t^T\mathcal{Z} \tag{17}$$

$$\mathcal{M}_2 = \beta_1(1-\alpha)^2 + \beta_2 + \beta_3 D_t^T D_t \tag{18}$$

- About the Λ sub-problem of update:

According to the multiplier update rule of ADMM, the update is performed according to the following formula:

$$\begin{aligned}
\Lambda_1^{k+1} &= \Lambda_1^k + \beta_1(\mathcal{X}^k - \mathcal{I} - \alpha\mathcal{S}^k - (1-\alpha)\mathcal{B}), \\
\Lambda_2^{k+1} &= \Lambda_2^k + \beta_2\left(\mathcal{Y}^k - \mathcal{B}^k\right), \\
\Lambda_3^{k+1} &= \Lambda_3^k + \beta_3\left(\mathcal{Z}^k - D_t\mathcal{B}^k\right).
\end{aligned} \tag{19}$$

The proposed smoke removal algorithm flow is given as Algorithm 1

Algorithm 1　Smoke Removal Algorithm Based on Spatial Continuit

---

1: **Input:** Smoky video $\mathcal{I}$
2: **When** $\|\mathcal{B}^{k+1} - \mathcal{B}^k\| / \|\mathcal{B}^k\| > \varepsilon$, where ε is a predetermined value, **Do:**
3: 　　step1: Update $\mathcal{X}$ by (7)
4: 　　step2: Update $\mathcal{Y}$ by (10)
5: 　　step3: Update $\mathcal{Z}$ by (12)
6: 　　step4: Update $\mathcal{S}$ by (14)
7: 　　step5: Update $\mathcal{B}$ by (15)
8: 　　step6: Update the Multipliers of ADMM by (19)
9: 　　step7: $k = k + 1$
10: **End**

---

5 │ EXPERIMENTAL RESULTS AND DISCUSSION

The performances of the proposed method from both simulated data and real-world data will be discussed in the section. What's more, the objective evaluation criteria for the images of the two kinds of data are also discussed below.

(a) Smoky Video1   (b) Smoky Video2

(c) Smoky Video3   (d) Smoky Video4
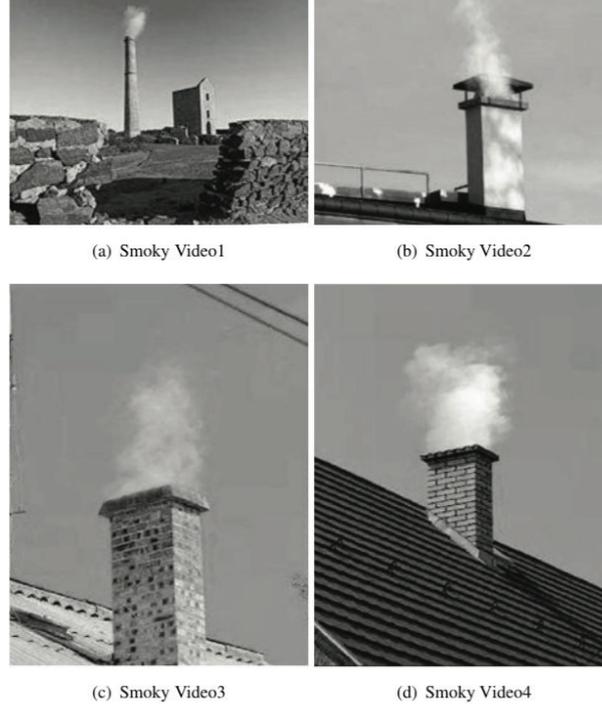
FIGURE 2 Selected frame randomly from four Simulated Smoky Videos

5.1 | Experimental Setting

Firstly, four simulated data shown in Figure 2 are used to test. In the simulated experiments, we use the image quality assessment index PSNR, SSIM[30], and FSIM [31] to measure the quality of reconstruction of the method.

(a) *Peak Signal-to-Noise Ratio(PSNR)*

$$\text{PSNR} = 10 \times \log_{10} \frac{mn}{\sum_{x=1}^{m} \sum_{y=1}^{n} [u(x, y) - \hat{u}(x, y)]^2} \tag{20}$$

(b) *Structural Similarity Index Measurement (SSIM)*

$$\text{SSIM} = \frac{(2\mu_u \mu_{\hat{u}} + C_1)(2\sigma_{u\hat{u}} + C_2)}{(\mu_u{}^2 + \mu_{\hat{u}}{}^2 + C_1)(\sigma_u{}^2 + \sigma_{\hat{u}}{}^2 + C_2)} \tag{21}$$

where the reference and the restored image is denoted by u and $\hat{u}$ respectively. μ_u and $\mu_{\hat{u}}$ represent the average values of image's u and $\hat{u}$.The variances are denoted by σ_u and $\sigma_{\hat{u}}$, and $\sigma_{u\hat{u}}$ is the covariance between u and $\hat{u}$. What's more, higher SSIM and PSNR values show the better restoration results.

(c) *Feature Similarity Index Measurement (FSIM)*

FSIM is a measure of similarity from two phases: the first phase calculates the similarity map for the local similarity of the phase consistency and the local similarity of the gradient, and the second phase combines the two similarity maps to a similarity evaluation.

(1) The first phase:

$$S_{PC}(x) = \frac{2PC_u(x) \cdot PC_{\hat{u}}(x) + T_u}{PC_u^2(xx) + PC_{\hat{u}}^2(x) + T_u} \tag{22}$$

$$S_G(x) = \frac{2G_u(ux) \cdot G_{\hat{u}}(ux) + T_{\hat{u}}}{G_u^2(x) + G_{\hat{u}}^2(u(x, y)) + T_{\hat{u}}} \tag{23}$$

(2) The second phase:

$$S_L(x) = [S_{PC}(x)]^\alpha \cdot [S_G(x)]^\beta \tag{24}$$

(a) Dry leaf smoke   (b) sParkingLot
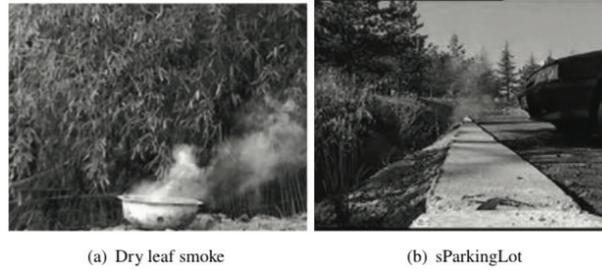
FIGURE 3 Selected frame randomly from two Real-World Smoky Videos

Weighting is performed using the value of the structure information at this point in response to the value of the structure information at this pixel:

$$\text{FSIM} = \frac{\Sigma_{x \in \Omega} S_L(x) \cdot PC_m(x)}{\Sigma_{x \in \Omega} PC_m(x)} \tag{25}$$

where α and β are the parameters influencing the relative importance of phase consistency (PC) and amplitude (GM), generally take the value of 1, Ω refers to the range of values of the entire image.

Secondly, the real-world data shown in Fig.3 are used to test. In the real-world experiments, we use the mean value and the contrast of the image to evaluate the reconstruction results.

(a) *Mean Value*

The mean value reflects the brightness of the image. The larger the mean value, the larger the brightness of the image, the smaller the opposite.

$$M = \frac{1}{mn} \sum_{xx=1}^{m} \sum_{y=1}^{n} u(x, y) \tag{26}$$

(b) *Contrast*

The contrast of the image characterizes the contrast of the grayscale of a pair of images. It is usually used to represent the sharpness of edges and textures. The greater the contrast, the clearer the image and the sharper the color of the image.

$$C = \frac{1}{mn} \sum_{xx=1}^{m} \sum_{y=1}^{n} (u(x, y) - M) \tag{27}$$

5.2 | Simulated Experiments

This section uses simulated smoky data to verify the effectiveness of our proposed smoke removal algorithm, mainly from two aspects: subjective visual evaluation and objective image measurement evaluation index value of PSNR, SSIM, FSIM. We selected four smoke videos for comparison experiments, and the sizes of the four videos are $420 \times 630 \times 12$, $400 \times 602 \times 12$, $418 \times 678 \times 12$ and $424 \times 298 \times 12$, and one frame image of the four smoke videos is shown in Fig.2.

This section selects the following four methods of haze removal for comparative experiments:

(1) DCP: the dehaze method based on Dark Chanel Prior, proposed by He et al.[32] in 2009, on single image haze removal using dark channel prior.Although this method was proposed earlier, it was a very important algorithm in the field of smoke removal at that time.

(2) DehazeNet: the multi-scale Neural Network Haze Removal method, proposed by Cai et al.[33] on an end-to-end system for single image haze removal.He learns directly from fog tiles and corresponding transmission rates.

(3) NLD: the non-local Image Dehaze method, proposed by Berman et al.[34] on non-local image dehazing.

(4) MLP: a single image defogging method based on multi-layer perceptron, proposed by Salazar-Colores et al.[35] on single image dehazing using a multilayer perceptron.

Fig.4 show the experimental results of four sets of simulated smoke video. In these figures, (a)-(e) represent the smoke removal effects of the DCP method, the DehazeNet method, the NLD method, and the MLP method, respectively, and (f) represents the ground true of the video data. Intuitively, the other four contrast methods are not able to remove the smoke from the smoke video. Our proposed algorithm slightly reduces the contrast of the smoke video, but completely removes the smoke. The DCP method makes the density of the smoke lighter, but it cannot effectively remove the smoke, and reduces the contrast of the
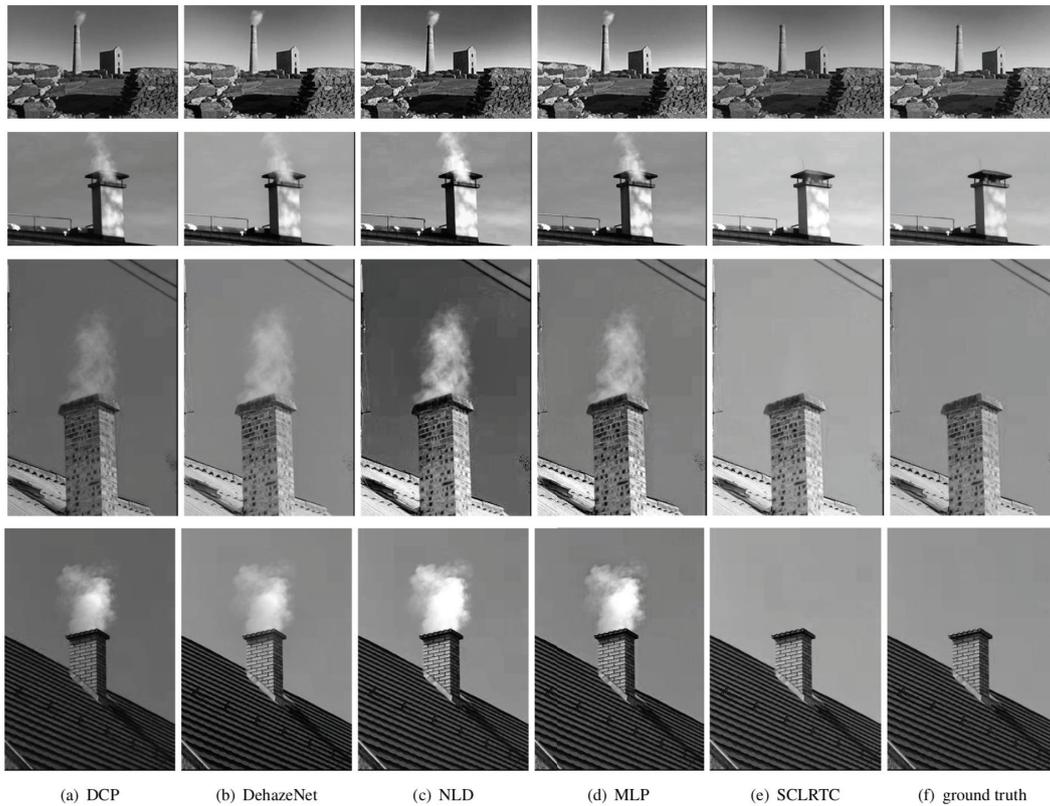
FIGURE 4 The smoke removal results of four sets simulated smoke video. (a) using DCP method, (b) using DehazeNet method, (c) using NLD method, (d) using MLP method, (e) using our proposed method, (f) ground truth.

image background. The NLD method deepens the concentration of smoke and cannot achieve the effect of removing smoke. The DehazeNet and MLP methods have less influence on the smoke video, and the experimental results show that the smoke still exists.

Fig.5 - Fig.8 show the evaluated results of quality assessment index PSNR, SSIM and FSIM corresponding to the four simulated smoke videos, respectively. In these figures, (a)-(c) compare the PSNR value, the SSIM value, and the FSIM value of selected frames from the recovered video, respectively. Obviously, it can be seen that our proposed algorithm achieves the highest PSNR value, SSIM value and FSIM value than other algorithms, which means that one smoky data processed by our algorithm can remove the smoke to the greatest extent, and maintain the largest structural and feature similarity with the original image.

Combining visual effects and evaluation index curves discussed above, we can conclude that our proposed algorithm has certain advantages in removing smoke, and also confirms the effectiveness of the algorithm.

5.3 │ Real-World Experiments

In this section, we uses real-world smoky data to verify the effectiveness of our proposed smoke removal algorithm, mainly from two aspects: subjective visual evaluation and objective image measurement evaluation index values of Mean Value and Contrast. We selected two smoke videos for comparison experiments. And they are dry leaf smoke and sParkingLot respectively. What's more, the sizes of the two videos are both $240 \times 320 \times 10$. One frame image of the four smoke videos is shown in Fig.3.

Fig.9 and Fig.10 visually show the experiment results of two kinds of real-world smoky videos. Fig.9 (a) is the selected original smoky frame of Dry leaf smoke video, (b)-(f) are the experiment results of DCP method, DehazeNet method, NLD method, MLP method, and our proposed smoke removal method respectively. Fig.10 (a) is the selected original smoky frame of sParkingLot video, (b)-(g) are the experiment results of DCP method, DehazeNet method, NLD method, MLP method, and our proposed smoke removal method respectively. From these figures, it is obviously that our proposed method better achieves the
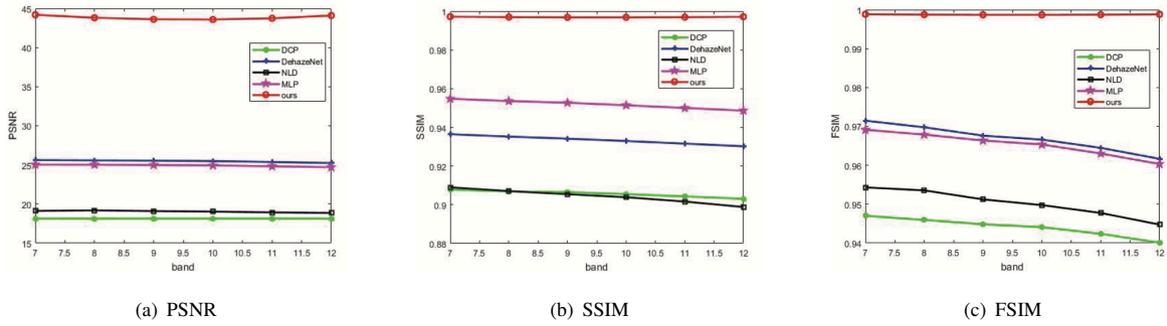
FIGURE 5 Quality assessment index of smoke removed results for the simulated Smoky Video (1).

(a) PSNR   (b) SSIM   (c) FSIM
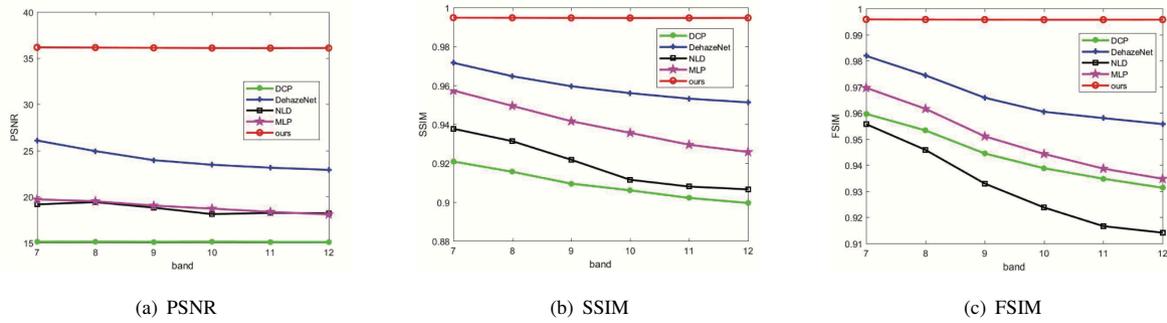
FIGURE 6 Quality assessment index of smoke removed results for the simulated Smoky Video (2).

(a) PSNR   (b) SSIM   (c) FSIM
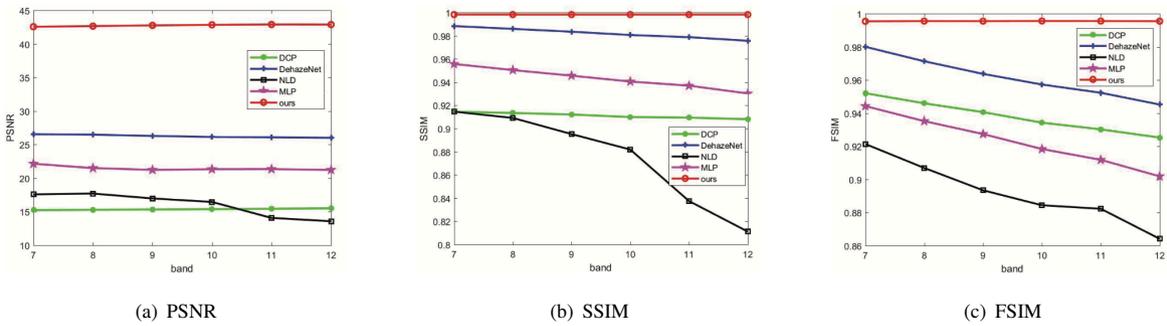
FIGURE 7 Quality assessment index of smoke removed results for the simulated Smoky Video (3).

(a) PSNR   (b) SSIM   (c) FSIM
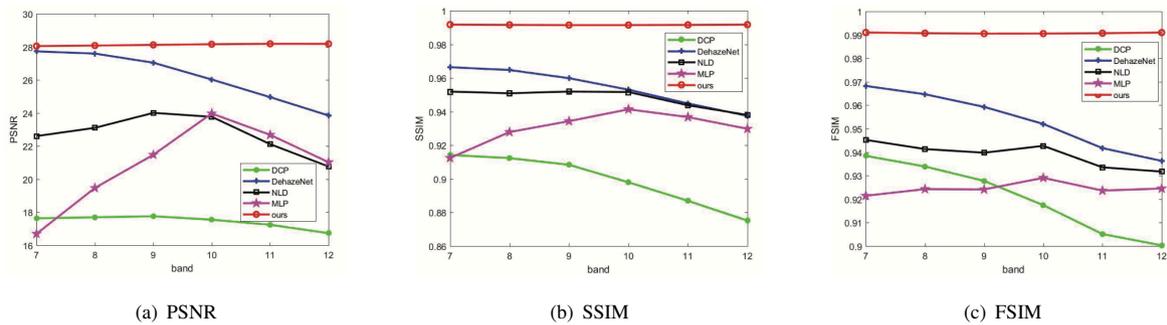
FIGURE 8 Quality assessment index of smoke removed results for the simulated Smoky Video (4).

(a) smoky frame   (b) DCP   (c) DehazeNet

(d) NLD   (e) MLP   (f) Ours
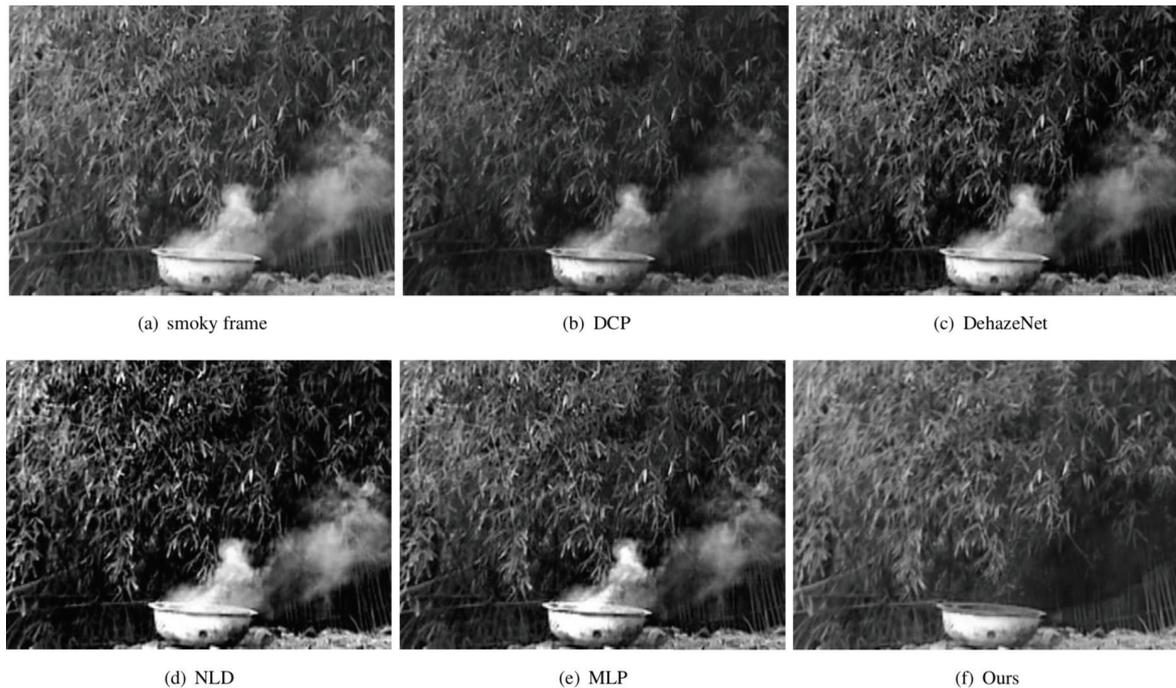
FIGURE 9 The smoke removal experiments of Dry leaf smoke video. (a) selected original smoky frame, (b) DCP method's smoke removal result, (c) DehazeNet method's smoke removal result, (d) NLD method's smoke removal result, (e) MLP method's smoke removal result, (f) our proposed method's smoke removal result.

TABLE 1 Mean Value of Two Smoky Data

Video Method frame		Mean value				
	DCP	62.446107	62.749909	61.855651	62.668346	62.577682
	DN	73.494734	73.250115	74.529303	74.299131	73.469587
sParkingLot	NLD	73.588242	74.152305	74.126172	72.659635	72.110456
	MLP	76.938061	77.138262	77.025695	77.064374	77.124972
	Ours	**127.126730**	**127.134182**	**127.135976**	**127.1318446**	**127.122953**
	DCP	61.361341	61.505495	61.640365	61.500990	61.330391
	DN	63.693823	63.859486	63.992342	63.875260	63.510371
Dry leaf Smoke	NLD	61.043841	61.852682	61.032031	61.336862	61.403307
	MLP	70.448134	69.565666	69.668507	70.965741	71.076125
	Ours	**78.632741**	**77.772004**	**77.233356**	**77.260031**	**77.998769**

purpose of smoke removal. DCP, DehazeNet, and MLP method have slight smoke removal effects. NLD method strengthens the contrast between background and smoke. And can not completely remove the smoke.

In addition, we select the mean and contrast of last five consecutive frames of the two smoky data as an objective measure of smoke removal results as shown in Table 1 and Table 2. From Table 1, we can see that our proposed smoke removal method achieved the highest mean values. This also shows that our smoke removal algorithm has better performance on smoke removal and also improves the brightness of video to some extent. From Table 2, we can see that our proposed smoke removal method achieved the highest contrast values.The greater the contrast, the clearer the image and the sharper the color of the image.

After the above comparison results in subjective visual evaluation and objective image measurement evaluation index values, it is obvious to show that our proposed smoke removal algorithm occupies a certain advantage compared to other algorithms.
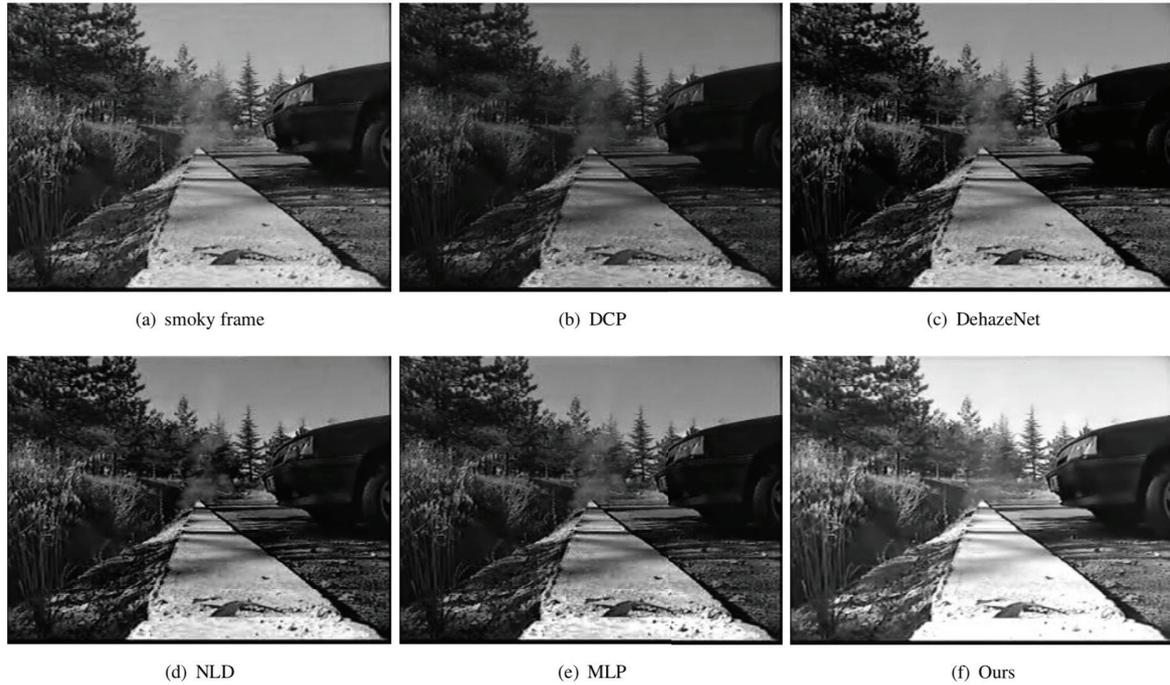
FIGURE 10 The smoke removal experiments of sParkingLot video. (a) selected original smoky frame, (b) DCP method's smoke removal result, (c) DehazeNet method's smoke removal result, (d) NLD method's smoke removal result, (e) MLP method's smoke removal result, (f) our proposed method's smoke removal result.

TABLE 2 Contrast Value of Two Smoky Data

Video Method frame		Contrast value				
	DCP	0.244074	0.245261	0.241766	0.244942	0.244588
	DN	0.287258	0.286302	0.291301	0.290402	0.287159
sParkingLot	NLD	0.287623	0.289828	0.289726	0.283994	0.281847
	MLP	0.300716	0.301499	0.301059	0.301210	0.301447
	Ours	**0.496910**	**0.496910**	**0.496917**	**0.496901**	**0.496866**
	DCP	0.239834	0.240397	0.240924	0.240380	0.239713
	DN	0.248950	0.249598	0.250117	0.249660	0.248233
Dry leaf Smoke	NLD	0.238593	0.241754	0.238547	0.239738	0.239998
	MLP	0.275350	0.271901	0.272303	0.277373	0.277804
	Ours	**0.307340**	**0.303976**	**0.301870**	**0.301975**	**0.304862**

5.4 | The Analysis of Computational Complexity

Here, we discuss the computational complexity of our model with a detailed specification of the running environment. We define that the experimental video's size is $H \times W \times N$, h and w are the rows and columns of the input video, n is its channel number.

As is shown in the Algorithm 1, the computational complexity is mainly decided by step 5. To get the solution of $\mathcal{B}$, we introduce fast Fourier transform (FFT) and inverse fast Fourier transform (IFFT). As a result, the complexity is $O(h^2 \times w^2 \times n)$. We perform the time consumption test using MATLAB R2016a on a laptop with an Intel Core i7-4270HQ CPU and 16 GB of RAM.
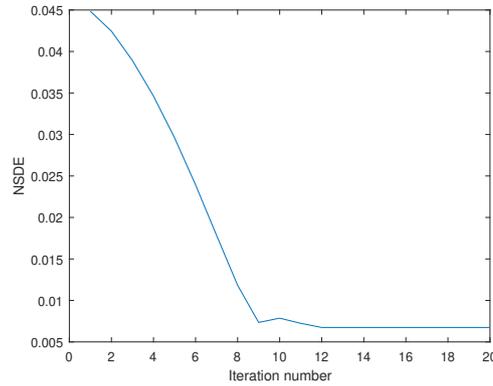
FIGURE 11 The convergence of our proposed algorithm

5.5 | Parameters Setting and Convergence Analysis

In this subsection, we discuss the parameters that we set in the proposed method and preform the analysis of convergence of our algorithm.

As mentioned above, the fidelity of the video recovery is controlled by α and τ. The value of μ and λ determines the sparseness of the smoke and local consistency of the clean video respectively. After comparing the experimental results by setting parameters many times, we find that when we set the value of α, τ, μ,λ to 0.1, 60, 50, 40 separately, the removal results are the best.

To demonstrate the convergence of our proposed algorithm, we use the value of normalized step difference energy (NSDE) to evaluate our approach. The lower the value is, the more similar the corresponding features of the processed video and the original video are, and the better the smoke removal effect is. We illustrate the convergence result in Fig.11. We can observe that our proposed method converges very quickly. After many experiments, we get the conclusion that the best result is preformed when it iterates 10 to 12 times.

6 | CONCLUSION

By analyzing the characteristics of smoke, based on the smoke mixing model, considering the sparseness of smoke and the global and local consistency of clean video, a video image smoke removal method based on space-time consistency is proposed in this paper. Then, the optimal solution of the smoke removal algorithm model is quickly implemented by the alternating direction multiplier method (ADMM). Finally, the visual effects and quality of the restored video images are evaluated from the subjective and objective aspects by simulation data.

ACKNOWLEDGMENT

References

1. Dong G, Wang N, Kuang G. Sparse representation of monogenic signal: With application to target recognition in SAR images. *IEEE Signal Processing Letters* 2014; 21(8): 952–956.

2. Comaniciu D, Ramesh V, Meer P. Kernel-based object tracking. *IEEE Transactions on Pattern Analysis & Machine Intelligence* 2003(5): 564–575.

3. Xu G, Khan S, Zhu H, Han L, Ng MK, Yan H. Discriminative tracking via supervised tensor learning. *Neurocomputing* 2018; 315: 33–47.

4. Xu G, Zhu H, Deng L, Han L, Li Y, Lu H. Dilated-aware discriminative correlation filter for visual tracking. *World Wide Web* 2019; 22(2): 791–805.

5. Bolkar S, Wang C, Cheikh FA, Yildirim S. Deep smoke removal from minimally invasive surgery videos. In: IEEE. ; 2018: 3403–3407.

6. Maruta H, Nakamura A, Kurokawa F. A new approach for smoke detection with texture analysis and support vector machine. In: IEEE. ; 2010: 1550–1555.

7. Piccinini P, Calderara S, Cucchiara R. Reliable smoke detection in the domains of image energy and color. In: IEEE. ; 2008: 1376–1379.

8. Chen TH, Yin YH, Huang SF, Ye YT. The smoke detection for early fire-alarming system base on video processing. In: IEEE. ; 2006: 427–430.

9. Çelik T, Özkaramanlı H, Demirel H. Fire and smoke detection without sensors: Image processing based approach. In: IEEE. ; 2007: 1794–1798.

10. Cheng J YJ. Smoke Removal Method for Ground Test Image of Solid Rocket Engine Based on Wavelet Transform. *Journal of System Simulation* 2004; 16(11): 2490–2492.

11. Serikawa S, Lu H. Underwater image dehazing using joint trilateral filter. *Computers & Electrical Engineering* 2014; 40(1): 41–50.

12. Yu S, Ou W, You X, Mou Y, Jiang X, Tang Y. Single image rain streaks removal based on self-learning and structured sparse representation. In: IEEE. ; 2015: 215–219.

13. Xu M, Jia X, Pickering M, Roberts D. Spectral unmixing for fire smoke detection and removal. In: IEEE. ; 2016: 806–808.

14. Nishino K, Kratz L, Lombardi S. Bayesian defogging. *International journal of computer vision* 2012; 98(3): 263–278.

15. Kotwal A, Bhalodia R, Awate SP. Joint desmoking and denoising of laparoscopy images. In: IEEE. ; 2016: 1050–1054.

16. Baid A, Kotwal A, Bhalodia R, Merchant S, Awate SP. Joint desmoking, specularity removal, and denoising of laparoscopy images via graphical models and Bayesian inference. In: IEEE. ; 2017: 732–736.

17. Chen WT, Yuan SY, Tsai GC, Wang HC, Kuo SY. Color Channel-Based Smoke Removal Algorithm Using Machine Learning for Static Images. In: IEEE. ; 2018: 2855–2859.

18. Zhu H, Qiao Y, Xu G, Deng L, Yu-Feng Y. DSPNet: A lightweight Dilated Convolution Neural Networks for Spectral Deconvolution with Self-Paced Learning. *IEEE Transactions on Industrial Informatics* 2019.

19. Lu H, Li Y, Mu S, Wang D, Kim H, Serikawa S. Motor anomaly detection for unmanned aerial vehicles using reinforcement learning. *IEEE internet of things journal* 2017; 5(4): 2315–2322.

20. Lu H, Li Y, Chen M, Kim H, Serikawa S. Brain intelligence: go beyond artificial intelligence. *Mobile Networks and Applications* 2018; 23(2): 368–375.

21. Lu H, Wang D, Li Y, et al. CONet: A cognitive ocean network. *IEEE Wireless Communications* 2019; 26(3): 90–96.

22. S L. Research on smoke removal algorithm for grayscale image of fire field. In: National University Safety Science and Engineering Annual Meeting Committee. ; 2018: 1.

23. Liu F, Ding Y, Rui T, Xu F. Learning Low-Rank Representation with Block-Sparse Structure for Single Sample Face Recognition. In: Association for Computing Machinery; 2016; New York, NY, USA.

24. Gao G, Yu Y, Yang M, Chang H, Huang P, Yue D. Cross-resolution face recognition with pose variations via multilayer locality-constrained structural orthogonal procrustes regression. *Information Sciences* 2020; 506: 19–36.

25. Khan S, Nawaz M, Guoxia X, Yan H. Image Correspondence with CUR Decomposition based Graph Completion and Matching. *IEEE Transactions on Circuits and Systems for Video Technology* 2019.

26. Yu YF, Xu G, Jiang M, Zhu H, Dai DQ, Yan H. Joint Transformation Learning via the L2, 1-Norm Metric for Robust Graph Matching. *IEEE transactions on cybernetics* 2019.

27. Li C, Ma Y, Huang J, Mei X, Ma J. Hyperspectral image denoising using the robust low-rank tensor recovery. *JOSA A* 2015; 32(9): 1604–1612.

28. Khan S, Xu G, Chan R, Yan H. An online spatio-temporal tensor learning model for visual tracking and its applications to facial expression recognition. *Expert Systems with Applications* 2017; 90: 427–438.

29. Jiang TX, Huang TZ, Zhao XL, Deng LJ, Wang Y. A novel tensor-based video rain streaks removal approach via utilizing discriminatively intrinsic priors. In: ; 2017: 4057–4066.

30. Hore A, Ziou D. Image quality metrics: PSNR vs. SSIM. In: IEEE. ; 2010: 2366–2369.

31. Zhang L, Zhang L, Mou X, Zhang D. FSIM: A feature similarity index for image quality assessment. *IEEE transactions on Image Processing* 2011; 20(8): 2378–2386.

32. He K, Sun J, Tang X. Single image haze removal using dark channel prior. *IEEE transactions on pattern analysis and machine intelligence* 2010; 33(12): 2341–2353.

33. Cai B, Xu X, Jia K, Qing C, Tao D. Dehazenet: An end-to-end system for single image haze removal. *IEEE Transactions on Image Processing* 2016; 25(11): 5187–5198.

34. Berman D, Avidan S, others . Non-local image dehazing. In: ; 2016: 1674–1682.

35. Salazar-Colores S, Cruz-Aceves I, Ramos-Arreguin JM. Single image dehazing using a multilayer perceptron. *Journal of Electronic Imaging* 2018; 27(4): 043022.