

# Self-Learning-based Rain Streak Removal for Image/Video

Li-Wei Kang  
Institute of Information Science  
Academia Sinica  
Taipei, Taiwan  
lwkang@iis.sinica.edu.tw

Chia-Wen Lin\*  
Department of Electrical Engineering  
National Tsing Hua University  
Hsinchu, Taiwan  
cwwlin@ee.nthu.edu.tw

Che-Tsung Lin<sup>1</sup> and Yu-Chen Lin<sup>2</sup>  
Safety Sensing and Control Department  
Intelligent Mobility Division, Mechanical  
and Systems Laboratories  
Industrial Technology Research Institute  
Hsinchu, Taiwan  
<sup>1</sup>AlexLin@itri.org.tw, <sup>2</sup>DavidLin@itri.org.tw

**Abstract**—Rain removal from an image/video is a challenging problem and has been recently investigated extensively. In our previous work, we have proposed the first single-image-based rain streak removal framework via properly formulating it as an image decomposition problem based on morphological component analysis (MCA) solved by performing dictionary learning and sparse coding. However, in this previous work, the dictionary learning process cannot be fully automatic, where the two dictionaries used for rain removal were selected heuristically or by human intervention. In this paper, we extend our previous work to propose an automatic self-learning-based rain streak removal framework for single image. We propose to automatically self-learn the two dictionaries used for rain removal without additional information or any assumption. We then extend our single-image-based method to video-based rain removal in a static scene by exploiting the temporal information of successive frames and reusing the dictionaries learned by the former frame(s) in a video while maintaining the temporal consistency of the video. As a result, the rain component can be successfully removed from the image/video while preserving most original details. Experimental results demonstrate the efficacy of the proposed algorithm.

## I. INTRODUCTION

Different weather conditions such as rain, snow, haze, or fog will cause complex visual effects of spatial or temporal domains in images or videos. Such effects may significantly degrade the performances of outdoor vision systems relying on image/video feature extraction. Removal of rain streaks has recently received much attention [1]–[5]. A pioneering work on detecting and removing rain streaks in a video was proposed in [1], where the authors developed a correlation model capturing the dynamics of rain and a physics-based motion blur model characterizing the photometry of rain. It was subsequently shown in [2] that some camera parameters, such as exposure time and depth of field can be selected to mitigate the effects of rain without altering the appearance of the scene. Furthermore, a model of the shape and appearance of a single rain or snow streak in the image space was developed in [3] to detect rain or snow streaks. Then, the amount of rain or snow in the video can be reduced.

So far, the research works on rain streak removal found in the literature have been mainly focused on video-based approaches that exploit temporal correlation in multiple

successive frames [1]–[4]. However, when only a single image is available, a single-image based rain streak removal approach is required. In our previous work [5], we proposed a single-image-based rain streak removal framework by formulating it as an image decomposition problem based on morphological component analysis (MCA) [6], [7]. However, in [5], we just followed the concept of the current MCA-based image decomposition approach [6], [7] to heuristically select the curvelet/wavelet dictionary for sparsely representing the non-rain component, while assuming the rain patches can be available for learning a rain dictionary for sparsely representing the rain component in a rain image.

In this paper, we extend our work presented in [5] to propose an automatic self-learning-based rain streak removal framework for single image. We then extend our single-image-based method to video-based rain removal in a static scene by exploiting the temporal information of successive frames and reusing the dictionaries learned by the former frame(s) in a video while maintaining the temporal consistency of the video. The major contribution of this paper is that the learning of the dictionaries used for removing rain streaks from an image/video is fully automatic and self-contained without any prior knowledge, where no extra training samples are required in the dictionary learning stage.

## II. MCA-BASED IMAGE DECOMPOSITION, SPARSE CODING, AND DICTIONARY LEARNING

### A. MCA-based Image Decomposition

Suppose that an image  $I$  of  $N$  pixels is a superposition of  $S$  layers (called morphological components), denoted by  $I = \sum_{s=1}^S I_s$ , where  $I_s$  denotes the  $s$ -th component, such as the geometric or textural component of  $I$ . To decompose the image  $I$  into  $\{I_s\}_{s=1}^S$ , the MCA algorithms [6], [7] iteratively minimize the following energy function:

$$E(\{I_s\}_{s=1}^S, \{\theta_s\}_{s=1}^S) = \frac{1}{2} \left\| I - \sum_{s=1}^S I_s \right\|_2^2 + \tau \sum_{s=1}^S E_s(I_s, \theta_s), \quad (1)$$

where  $\theta_s$  denotes the sparse coefficients corresponding to  $I_s$  with respect to dictionary  $D_s$ ,  $\tau$  is a regularization parameter, and  $E_s$  is the energy defined according to the type of  $D_s$  (global or local dictionary).

This work was supported in part by the National Science Council, Taiwan, under Grants NSC98-2221-E-007-080-MY3, NSC100-2218-E-001-007-MY3, and NSC100-2811-E-001-005.

\*Corresponding author

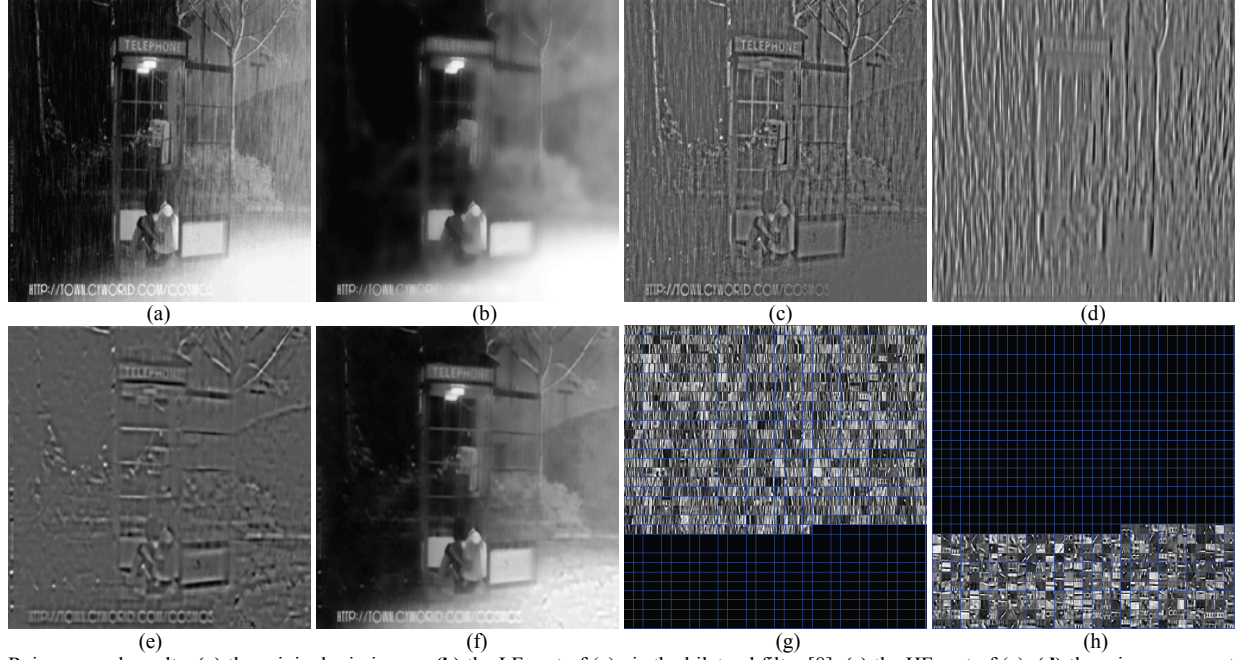


Fig. 1. Rain removal results: (a) the original rain image; (b) the LF part of (a) via the bilateral filter [8]; (c) the HF part of (a); (d) the rain component of (a); (e) the non-rain component of (a); (f) the rain-removed version of (a) via the proposed single-image-based method; (g) the rain sub-dictionary for (c); and (h) the non-rain sub-dictionary for (c).

The MCA algorithms solve (1) by iteratively performing for each component  $I_s$  the following two steps: (i) update of the sparse coefficients: this step performs sparse coding to solve  $\theta_s$  or  $\{\theta_s^k\}_{k=1}^N$ , where  $\theta_s^k$  represents the sparse coefficients of patch  $b_s^k$  extracted from  $I_s$ , to minimize  $E_s(I_s, \theta_s)$  while fixing  $I_s$ ; and (ii) update of the components: this step updates  $I_s$  or  $\{b_s^k\}_{k=1}^N$  while fixing  $\theta_s$  or  $\{\theta_s^k\}_{k=1}^N$ .

### B. Sparse Coding and Dictionary Learning

Sparse coding [9] is a technique of finding a sparse representation for a signal with a small number of nonzero or significant coefficients corresponding to the atoms in a dictionary. To construct a dictionary  $D_s$  for sparsely representing each patch  $b_s^k$  extracted from the component  $I_s$  of the image  $I$ , we may use a set of available training exemplars  $y^k, k = 1, 2, \dots, P$ , to learn a dictionary  $D_s$  sparsifying  $y^k$  by solving the following optimization problem:

$$\min_{D_s, \theta^k} \sum_{k=1}^P \left( \frac{1}{2} \|y^k - D_s \theta^k\|_2^2 + \lambda \|\theta^k\|_1 \right), \quad (2)$$

where  $\theta^k$  denotes the sparse coefficients of  $y^k$  with respect to  $D_s$  and  $\lambda$  is a regularization parameter. Equation (2) can be efficiently solved by performing a dictionary learning algorithm, such as the online dictionary learning algorithm, where the sparse coding step is usually achieved via OMP (orthogonal matching pursuit) [9]. Finally, the image decomposition is achieved by iteratively performing the MCA algorithm to solve  $I_s$  (while fixing  $D_s$ ) described in Sec. II-A and the dictionary learning algorithm to learn  $D_s$  (while fixing  $I_s$ ) until convergence.

## III. PROPOSED RAIN STREAK REMOVAL FRAMEWORK

Similar to [5], in our method, an input rain image is first roughly decomposed into the low-frequency (LF) part and the

high-frequency (HF) part using the bilateral filter [8], where the most basic information will be retained in the LF part while the rain streaks and the other edge/texture information will be included in the HF part of the image as illustrated in Figs. 1(b) and 1(c). Then, we perform the proposed MCA-based image decomposition to the HF part that can be further decomposed into the rain component [see Fig. 1(d)] and the non-rain component [see Fig. 1(e)]. Different from [5], in the image decomposition step, a dictionary learned from the training exemplars extracted from the HF part of the image itself can be divided into two sub-dictionaries by performing HOG (histograms of oriented gradients) [10] feature-based dictionary atom clustering. Then, we perform sparse coding [9] based on the two sub-dictionaries to achieve MCA-based image decomposition, where the non-rain component in the HF part can be obtained, followed by integrating with the LF part of the image to obtain the rain-removed version of this image as illustrated in Fig. 1(f). The detailed method shall be elaborated below.

### A. Preprocessing and Problem Formulation

For an input rain image  $I$ , in the preprocessing step, we apply a bilateral filter [8] to decompose  $I$  into the LF part ( $I_{LF}$ ) and HF part ( $I_{HF}$ ), i.e.,  $I = I_{LF} + I_{HF}$ . Then, our method learns a dictionary  $D_{HF}$  based on the training exemplar patches extracted from  $I_{HF}$  to further decompose  $I_{HF}$ , where  $D_{HF}$  can be further divided into two sub-dictionaries,  $D_{HF,G}$  and  $D_{HF,R}$  ( $D_{HF} = [D_{HF,G} | D_{HF,R}]$ ), for representing the non-rain and rain components of  $I_{HF}$ , respectively. As a result, we formulate the problem of rain streak removal for image  $I$  as a sparse coding-based image decomposition problem:

$$\min_{\theta_{HF}^k} \|b_{HF}^k - D_{HF} \theta_{HF}^k\|_2^2 \text{ s.t. } \|\theta_{HF}^k\|_0 \leq L, \quad (3)$$

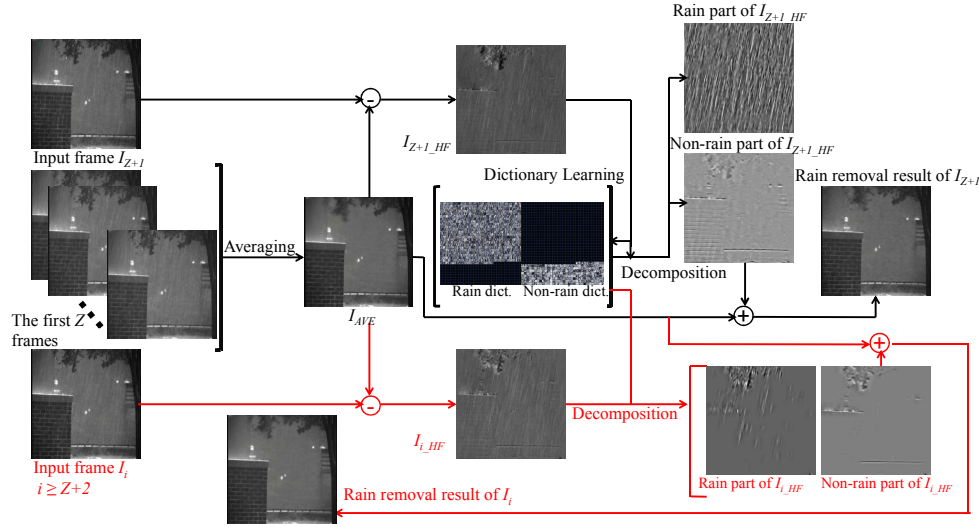


Fig. 2. Block diagram of the proposed video-based rain streak removal method.

where  $b_{HF}^k$  represents the  $k$ -th patch extracted from  $I_{HF}$ .  $\theta_{HF}^k$  are the sparse coefficients of  $b_{HF}^k$  with respect to  $D_{HF}$ , and  $L$  denotes the sparsity or maximum number of nonzero coefficients of  $\theta_{HF}^k$ .

### B. Dictionary Learning and Partition

In this step, we extract from  $I_{HF}$  a set of overlapping patches as the training exemplars  $y^k$  for learning dictionary  $D_{HF}$  by the dictionary learning technique described in Sec. II-B using the online dictionary learning algorithm [9] to obtain  $D_{HF}$ . We find that the atoms constituting  $D_{HF}$  can be roughly divided into two clusters (sub-dictionaries) for representing the non-rain and rain components of  $I_{HF}$ , respectively. Intuitively, the most significant feature for a rain atom can be extracted via “image gradient.” In this work, we utilize the HOG descriptor [10] to describe each atom in  $D_{HF}$ . After extracting the HOG feature for each atom in  $D_{HF}$ , we then apply the  $K$ -means algorithm to classify all of the atoms in  $D_{HF}$  into two clusters  $D_1$  and  $D_2$  based on their HOG feature descriptors. Then, we calculate the variance of gradient direction for each atom  $d_{ij}$  in cluster  $D_i$ , as  $VG_{ij}$ ,  $i = 1, 2$ . Then, we calculate the mean of  $VG_{ij}$  for each cluster  $D_i$  as  $MVG_i$ . Based on the fact that the edge directions of rain streaks in an atom are usually consistent, *i.e.*, the variance of gradient direction for a rain atom should be small, we identify the cluster with the smaller  $MVG_i$  as rain sub-dictionary  $D_{HF\_R}$ , and the other one as non-rain sub-dictionary  $D_{HF\_G}$ , as depicted in Figs. 1(g) and 1(h).

### C. Removal of Rain Streaks

Based on the two dictionaries  $D_{HF\_R}$  and  $D_{HF\_G}$ , we perform sparse coding by applying the OMP algorithm for each patch  $b_{HF}^k$  extracted from  $I_{HF}$  via minimization of (3), where  $D_{HF} = [D_{HF\_G} | D_{HF\_R}]$ , to find its sparse coefficients  $\tilde{\theta}_{HF}^k$ . Then, each reconstructed patch  $b_{HF}^k$  can be used to recover either non-rain component  $I_{HF}^G$  or rain component  $I_{HF}^R$  of  $I_{HF}$  based on the sparse coefficients  $\tilde{\theta}_{HF}^k$  as follows. We set the coefficients corresponding to  $D_{HF\_G}$  in  $\tilde{\theta}_{HF}^k$  to zeros to obtain  $\tilde{\theta}_{HF\_R}^k$ , while the coefficients corresponding to  $D_{HF\_R}$  in

$\tilde{\theta}_{HF}^k$  to zeros to obtain  $\tilde{\theta}_{HF\_G}^k$ . Therefore, each patch  $b_{HF}^k$  can be re-expressed as either  $\tilde{b}_{HF\_G}^k = D_{HF\_G} \times \tilde{\theta}_{HF\_G}^k$  or  $\tilde{b}_{HF\_R}^k = D_{HF\_R} \times \tilde{\theta}_{HF\_R}^k$ , which can be used to recover  $I_{HF}^G$  or  $I_{HF}^R$ , respectively. Finally, the rain-removed version of the image  $I$  can be obtained via  $I^{Non\_Rain} = I_{LF} + I_{HF}^G$ . More details of the proposed method for single image can be found in [11].

### D. Extension to Video-based Rain Streak Removal

The most straightforward way to extend our single-image-based method to video-based rain removal is to individually apply the single-image-based method to each frame in a video. Without exploiting the temporal information in a video, there are major two drawbacks of this strategy: (i) the temporal consistency of the video cannot be maintained; and (ii) the overall computational complexity is expensive. In this work, we consider a rain video of static scene without significant moving objects. Fig. 2 shows the proposed video-based rain streak removal framework, where we find that by averaging a number of successive frames in a static scene, the rain streaks can be eliminated in this “average frame,” which can be used to replace the bilateral-filtered image (LF part) in our single-image-based method.

For an input rain video of  $V$  frames  $I_i$ ,  $i = 1, 2, \dots, V$ , we average the first  $Z$  successive frames  $I_i$ ,  $i = 1, 2, \dots, Z$ , to generate the common LF part  $I_{AVE}$  for all of the remaining frames  $I_i$ ,  $i = Z + 1, Z + 2, \dots, V$ , in the video. We then apply our single-image-based method to  $I_{Z+1}$  with LF part being set to  $I_{AVE}$  to obtain the rain and non-rain dictionaries,  $D_{(Z+1)\_HF\_R}$  and  $D_{(Z+1)\_HF\_G}$ , to decompose the HF part  $I_{(Z+1)\_HF}$  ( $= I_{Z+1} - I_{AVE}$ ) of  $I_{Z+1}$  into the rain and non-rain components,  $I_{(Z+1)\_HF}^R$  and  $I_{(Z+1)\_HF}^G$ , respectively. Then, the rain-removed version of  $I_{Z+1}$  can be obtained via  $I_{(Z+1)}^{Non\_Rain} = I_{AVE} + I_{(Z+1)\_HF}^G$ . For removing rain streaks from  $I_i$ ,  $i = Z + 2, Z + 3, \dots, V$ , in the video, we use the same LF part  $I_{AVE}$  to obtain the HF part  $I_{i\_HF}$  ( $= I_i - I_{AVE}$ ). We then directly perform MCA decomposition to  $I_{i\_HF}$ ,  $i = Z + 2, Z + 3, \dots, V$ , using the same two dictionaries,  $D_{(Z+1)\_HF\_R}$  and  $D_{(Z+1)\_HF\_G}$ , learned from  $I_{(Z+1)\_HF}$  to obtain the rain and non-rain



components,  $I_{LHF}^R$  and  $I_{LHF}^G$ , respectively. Finally, the rain-removed version of  $I_i$ ,  $i = Z + 2, Z + 3, \dots, V$ , can be obtained via  $I_i^{Non\_Rain} = I_{AVE} + I_{LHF}^G$ . The main advantage of the proposed video-based method is two-fold: (i) for all of the frames in a video, using the same LF part and the same two dictionaries for rain removal can maintain the temporal consistency of the video; and (ii) the dictionary learning process, which induces the major computational burden, is only performed once, which can significantly save the overall computational complexity.

#### IV. EXPERIMENTS AND DISCUSSION

To evaluate the performance of the proposed single-image-based method, we compare the proposed method with the bilateral filter [8] and our previous semi-automatic method [5]. Moreover, to evaluate the performance of the proposed video-based method, we compare the proposed method with the video-based rain removal method proposed in [1] and our single-image-based method. The parameter settings of the proposed methods are described as follows. For each test gray-scale image of size  $256 \times 256$ , the patch size, dictionary size, and the number of training iterations are set to  $16 \times 16$ , 1024, and 100, respectively. The number of frames used to generate the LF part for a static video is set to  $Z = 50$ . The rain removal results obtained from our previous method [5], the method proposed in [1], our single-image-based method, and our video-based method are shown in Figs. 1, 3–4. It can be observed that our methods can remove most rain streaks while preserving most non-rain image/video details, thereby improving the subjective visual quality significantly. More experimental results and tested videos are available in [12].

#### V. CONCLUSIONS AND FUTURE WORK

In this paper, we have proposed an automatic self-learning-based rain streak removal framework for single image by formulating it as an MCA-based image decomposition problem solved by performing sparse coding and dictionary learning algorithms. We have also extended it to video-based rain removal in a static scene. Our experimental results show that the proposed methods achieve comparable/better performance with/than our previous work in [5] and the video-based rain removal method proposed in [1]. For future work, the proposed methods may be extended to rain removal for video of dynamic scenes.

#### REFERENCES

- [1] K. Garg and S. K. Nayar, "Detection and removal of rain from videos," in *Proc. CVPR*, June 2004, vol. 1, pp. 528–535.
- [2] K. Garg and S. K. Nayar, "When does a camera see rain?" in *Proc. of IEEE Int. Conf. Comput. Vis.*, Oct. 2005, vol. 2, pp. 1067–1074.
- [3] P. C. Barnum, S. Narasimhan, and T. Kanade, "Analysis of rain and snow in frequency space," *IJCV*, vol. 86, no. 2–3, pp. 256–274, 2010.
- [4] J. Bossu, N. Hautière, and J. P. Tarel, "Rain or snow detection in image sequences through use of a histogram of orientation of streaks," *Int. J. Comput. Vis.*, vol. 93, no. 3, pp. 348–367, July 2011.
- [5] Y.-H. Fu, L.-W. Kang, C.-W. Lin, and C.-T. Hsu, "Single-frame-based rain removal via image decomposition," in *Proc. ICASSP*, May 2011.
- [6] J. M. Fadili, J. L. Starck, J. Bobin, and Y. Moudden, "Image decomposition and separation using sparse representations: an overview," *Proc. IEEE*, vol. 98, no. 6, pp. 983–994, June 2010.
- [7] G. Peyré, J. Fadili, and J. L. Starck, "Learning adapted dictionaries for geometry and texture separation," in *Proc. SPIE*, vol. 6701, 2007.
- [8] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *Proc. IEEE Int. Conf. Comput. Vis.*, Jan. 1998, pp. 839–846.
- [9] J. Mairal, F. Bach, J. Ponce, and G. Sapiro, "Online learning for matrix factorization and sparse coding," *J. Mach. Learn. Res.*, vol. 11, 2010.
- [10] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proc. CVPR*, San Diego, CA, USA, June 2005.
- [11] L.-W. Kang, C.-W. Lin, and Y.-H. Fu, "Automatic single-image-based rain streaks removal via image decomposition," *IEEE Trans. Image Process.* (in press).
- [12] NTHU Rain Removal project. [Online]. Available: [http://www.ee.nthu.edu.tw/cwlin/Rain\\_Removal/Rain\\_Removal.htm](http://www.ee.nthu.edu.tw/cwlin/Rain_Removal/Rain_Removal.htm).

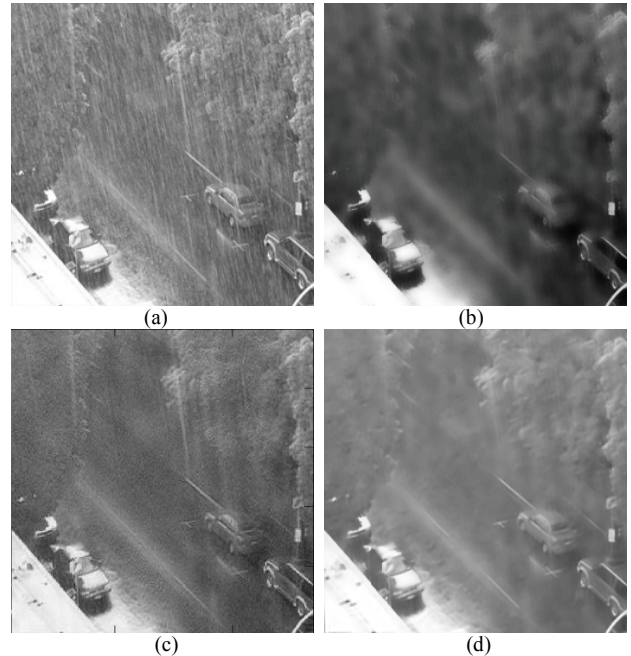


Fig. 3. Rain removal results: (a) the original rain image; and the rain-removed versions of (a) via: (b) the bilateral filter [8]; (c) our previous semi-automatic method [5]; and (d) the proposed single-image-based method.

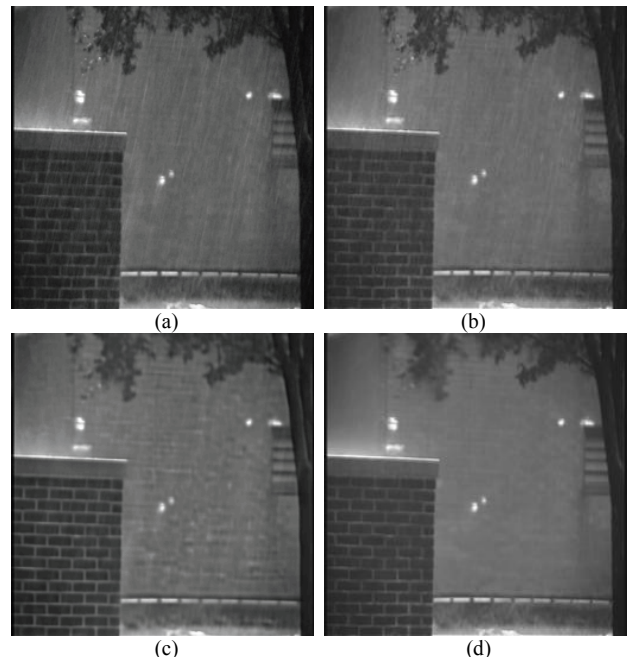


Fig. 4. Rain removal results: (a) the original rain video frame; and the rain-removed versions of (a) via: (b) the method proposed in [1]; (c) the proposed single-image-based method; and (d) the proposed video-based method.