This paper has been accepted by IEEE ICASSP 2011.

SINGLE-FRAME-BASED RAIN REMOVAL VIA IMAGE DECOMPOSITION+

Yu-Hsiang Fu[1], *Li-Wei Kang*[2], *Chia-Wen Lin*[1*], *and Chiou-Ting Hsu*[3]

[1]Department of Electrical Engineering, National Tsing Hua University, Hsinchu, Taiwan
[2]Institute of Information Science, Academia Sinica, Taipei, Taiwan
[3]Department of Computer Science, National Tsing Hua University, Hsinchu, Taiwan

ABSTRACT

Rain removal from a video is a challenging problem and has been recently investigated extensively. Nevertheless, the problem of rain removal from a single image has been rarely studied in the literature, where no temporal information among successive images can be exploited, making it more challenging. In this paper, to the best of our knowledge, we are among the first to propose a single-frame-based rain removal framework via properly formulating rain removal as an image decomposition problem based on morphological component analysis (MCA). Instead of directly applying conventional image decomposition technique, we first decompose an image into the low-frequency and high-frequency parts using a bilateral filter. The high-frequency part is then decomposed into "rain component" and "non-rain component" via performing dictionary learning and sparse coding. As a result, the rain component can be successfully removed from the image while preserving most original image details. Experimental results demonstrate the efficacy of the proposed algorithm.

***Index Terms*—** Rain removal, sparse coding, dictionary learning, image decomposition, morphological component analysis (MCA).

1. 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 [1]–[4]. Such effects may significantly degrade the performances of outdoor vision systems relying on image/video feature extraction, such as object detection, tracking, recognition, indexing, and retrieval. Removal of rain streaks has recently received much attention. The first work for detecting and removing rain from videos 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. Then, the same authors [2] also showed that the 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. Moreover, an improved video rain removal algorithm incorporating both temporal and chromatic properties was proposed in [3]. On the other hand, a model of the shape and appearance of a single rain or snow streak in the image space was developed in [4], which can be fit to a video and used to detect rain or snow streaks. Then, the amount of rain or snow in the video can be reduced or increased.

So far, the research works on rain removal found in the literature have been mainly focused on video-based approaches that exploit information in multiple successive frames. Nevertheless, when only a single image is available, such as an image captured from a digital camera or downloaded from the Internet, a single-frame based rain removal approach is required, which was rarely investigated before. In this paper, we propose a single-frame-based rain removal framework via formulating rain removal as an image decomposition problem based on morphological component analysis (MCA) [5]-[6]. The major contribution of this paper is three-fold: (i) our scheme is among the first to formulate the single image rain removal problem as an image decomposition problem; (ii) the learning of the dictionary for decomposing rain steaks is self-contained, where no extra training samples are required in the dictionary learning stage; and (iii) no temporal or motion information among successive images is required.

The rest of this paper is organized as follows. In Sec. 2, we briefly review the concepts of MCA-based image decomposition, sparse coding, and dictionary learning, respectively. Sec. 3 presents the proposed single-image-based rain removal framework. In Sec. 4, simulation results are demonstrated. Finally, Sec. 5 concludes this paper.

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

2.1. MCA-based Image Decomposition

Suppose that an image I of N pixels is a linear combination 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 geometric or textured component of I. To decompose the image I into $\{I_s\}_{s=1}^{S}$, the MCA algorithms [5], [6] iteratively minimize the energy function defined as:

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

where θ_s denotes the sparse coefficients corresponding to I_s with respect to the dictionary D_s described later, μ is a

+This work was supported in part by the National Science Council, Taiwan, under Grants NSC99-2218-E-001-010 and NSC99-2811-E-001-006.
*Corresponding author: cwlin@ee.nthu.edu.tw.

regularization parameter, and E_s is the energy defined according to the type of D_s (global or local). For a global dictionary $D_s \in \mathrm{R}^{N\times M_s}$, $N \le M_s$, $\theta_s \in \mathrm{R}^{M_s}$ is the sparse coefficients of I_s with respect to D_s. The energy function E_s for global dictionary is defined in [6] as

$$E_s\left(I_s,\theta_s\right)=\frac{1}{2}\left\|I_s-D_s\theta_s\right\|_2^2+\lambda\left\|\theta_s\right\|_1, \tag{2}$$

where λ is a regularization parameter. Usually, to decompose an image into its geometric and textured components, traditional basis, such as wavelet or curvelet, is used as the global dictionary for representing the geometric component of the image.

For a local dictionary $D_s \in \mathrm{R}^{n\times m_s}$, $n \le m_s$, $\theta_s^k \in \mathrm{R}^{m_s}$ is the sparse coefficients of the patch $b_s^k \in \mathrm{R}^n$, k = 1, 2, …, N, extracted from I_s, with respect to D_s. Each patch b_s^k can be extracted centralized with a pixel of I_s and overlapped with adjacent patches. The energy function E_s for the local dictionary can be defined as [6]

$$E_s\left(I_s,\theta_s\right)=\frac{1}{n}\sum_{k=1}^{N}\left(\frac{1}{2}\left\|b_s^k-D_s\theta_s^k\right\|_2^2+\lambda\left\|\theta_s^k\right\|_1\right), \tag{3}$$

where the weight $1/n$ compensates for the redundancy factor introduced by the overlap between patches b_s^k. Usually, the local dictionary for representing the textured component of an image is constructed from the dictionary learning procedure described in Sec. 2.2.

To decompose the image I into $\left\{I_s\right\}_{s=1}^{S}$, the MCA algorithm [6] solves (1) by iteratively performing the two steps for each component I_s, as follows: (i) update of the sparse coefficients: this step performs sparse coding via solving a convex non-smooth optimization to solve θ_s or $\left\{\theta_s^k\right\}_{k=1}^{N}$ to minimize $E_s\left(I_s,\theta_s\right)$ while fixing I_s; and (ii) update of the components: this step updates I_s or $\left\{b_s^k\right\}_{k=1}^{N}$ while fixing θ_s or $\left\{\theta_s^k\right\}_{k=1}^{N}$.

2.2. Sparse Coding and Dictionary Learning

Sparse coding is the 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 [5]-[7]. Recall from Sec. 2.1, it is required to construct a dictionary D_s containing the local structures of textures for sparsely representing each patch b_s^k extracted from the textured component I_s of an image I. In some applications, we may use a set of available training exemplars (similar to the patches extracted from the component we want to extract) $y^k \in \mathrm{R}^n$, k = 1, 2, …, p, to learn a dictionary D_s sparsifying y^k via solving the following optimization problem [7]:

$$\min_{D_s\in\mathrm{R}^{n\times m_s},\,\theta^k\in\mathrm{R}^{m_s}}\sum_{k=1}^{p}\left(\frac{1}{2}\left\|y^k-D_s\theta^k\right\|_2^2+\lambda\left\|\theta^k\right\|_1\right), \tag{4}$$

where θ^k denotes the sparse coefficients of y^k with respect to D_s and λ is a regularization parameter, which can be efficiently solved via performing the K-SVD dictionary learning algorithm [7]. Finally, the image decomposition is achieved via iteratively performing the MCA algorithm to solve I_s (while fixing D_s) and the dictionary learning algorithm to learn D_s (while fixing I_s) until convergence [6].

3. PROPOSED RAIN REMOVAL FRAMEWORK

As shown in Fig. 1, in the proposed framework rain removal is formulated as an image decomposition problem. It can be observed from Fig. 2(b) that directly applying the MCA-based image decomposition algorithm described in Sec. 2 [6] by treating rain streaks as the textured component in an image will seriously blur the image even if the rain streaks can be removed. To prevent original image details from being removed together with rain streaks, we propose to first roughly decompose an image into the low-frequency (LF) part and the high-frequency (HF) part, as illustrated in Figs. 2(c) and 2(d), respectively. Obviously, the most basic information will be included in the LF part while the rain streaks and the other edge/texture information will be included in the HF part of an image. Then, the HF part is further decomposed into the "rain component" and "non-rain component," as illustrated in Figs. 2(e) and 2(f), respectively, where the training exemplars are extracted from the image itself in the dictionary learning stage. The details and the problem formulation of the proposed scheme are elaborated in the following subsections.

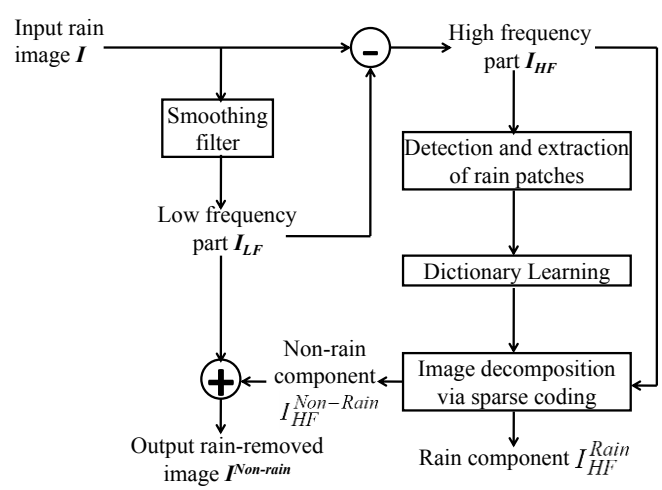

Fig. 1. Proposed rain removal framework.

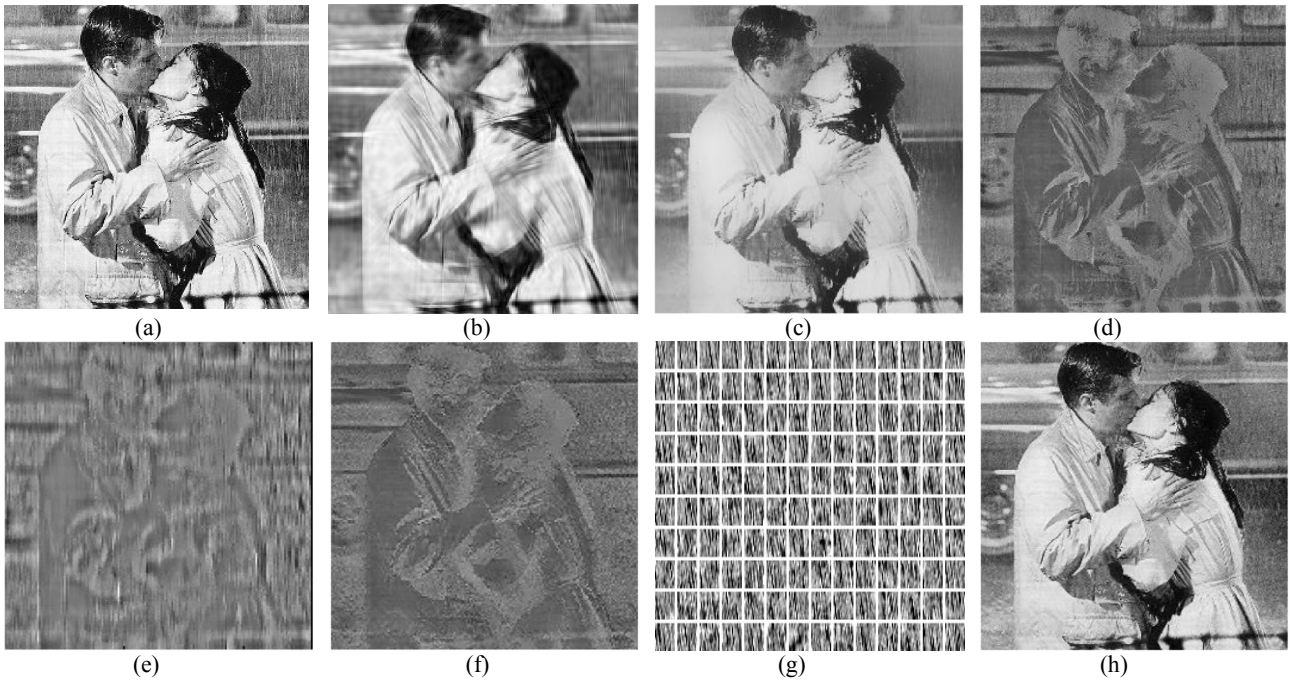

Fig. 2. Illustrations of rain removal for a single image: **(a)** the original rain image; **(b)** the rain-removed version of (a) with the image decomposition via MCA [6]; **(c)** the low-frequency part (separated by the bilateral filter [8]) of (a); **(d)** the high-frequency part of (a); **(e)** the rain component of (d); **(f)** the non-rain component of (d); **(g)** the learned dictionary for decomposing (d); and **(h)** the rain-removed version of (a) with the proposed scheme.

3.1. Preprocessing and Problem Formulation

For an input rain image I, in the preprocessing stage, we apply an edge-preserving smoothing filter, called bilateral filter [8] to obtain the LF part I_{LF} of I and roughly decompose I into the LF part (I_{LF}) and HF part (I_{HF}), *i.e.*, $I = I_{LF} + I_{HF}$. Similar to the MCA algorithm proposed in [6], we select the curvelet basis $D_{Curvelet}$ as the global dictionary for representing the geometric component of I_{HF}. For representing the rain component of I_{HF}, we use a dictionary D_{Rain} learned from the training exemplars of rain streaks extracted from I_{HF} (described in Sec. 3.2). Hence, we formulate the problem of rain removal for image I of N pixels as an image decomposition problem to minimize the energy function defined as

$$E(I_{HF}, \theta_{HF}) = \sum_{k=1}^{N} \left(\frac{1}{2} \left\| b_{HF}^{k} - D_{HF}\theta_{HF}^{k} \right\|_2^2 + \lambda \left\| \theta_{HF}^{k} \right\|_1 \right), \quad (5)$$

where θ_{HF} denotes the sparse coefficients of I_{HF} with respect to the dictionary $D_{HF} = [D_{Curvelet} | D_{Rain}] \in \mathrm{R}^{n \times m}$, $n \leq m$, $b_{HF}^{k} \in \mathrm{R}^{n}$ represents the k-th patch extracted from I_{HF}, $k = 1, 2, \ldots, N$, $\theta_{HF}^{k} \in \mathrm{R}^{m}$ is the sparse coefficients of b_{HF}^{k} with respect to D_{HF}, and λ is a regularization parameter.

3.2. Detection of Rain Streaks and Dictionary Learning

In the dictionary learning stage, different from the MCA algorithm, where several training patches are usually collected in advance, in the proposed scheme, we extract the rain patches from I_{HF}, to be the training exemplars. First, we detect and extract all the patches containing rain streaks from I_{HF}. Based on the fact that the edge directions of rain streaks in a patch should be almost consistent, rain patches can be well distinguishable from other texture patches. To detect the rain streaks in a patch, we derive the two criteria: (i) the intensity of each pixel in a rain streak is larger than those of the neighboring pixels outside the streak; and (ii) the pixels in a rain streak can be projected onto a single position if the projection direction is correct.

After extracting a set of rain patches $y^{k} \in \mathrm{R}^{n}$, $k = 1, 2, \ldots, p$, from I_{HF} as training exemplars for learning the dictionary D_{Rain}, we formulate the dictionary learning problem as:

$$\min \sum_{k=1}^{p} \left(\frac{1}{2} \| y^{k} - D_{Rain}\theta^{k} \|_2^2 + \lambda \| \theta^{k} \|_1 \right), \quad (6)$$

where θ^{k} denotes the sparse coefficients of y^{k} with respect to D_{rain}. Similar to [6], we also apply the K-SVD algorithm [7] to solve (6) to obtain the optimized D_{Rain} that consists of the atoms of rain streaks, as illustrated in Fig. 2(g).

3.3. Removal of Rain Streaks

Based on the dictionary D_{HF}, we can perform sparse coding via applying the OMP (orthogonal matching pursuit) algorithm [9] (also applied in [6]) for each patch b_{HF}^{k} extracted from I_{HF} via minimizing (5) to find its sparse coefficients $\tilde{\theta}$. Different from the MCA algorithm [6], where the sparse coding and dictionary learning will be iteratively performed, we perform sparse coding only once for each patch with respect to D_{HF}. Then, each reconstructed

patch b_{HF}^{k} can be used to recover either geometric component $I_{HF}^{Non-Rain}$ or rain component I_{HF}^{Rain} of I_{HF} based on the sparse coefficients $\tilde{\theta}_{HF}^{k}$ as follows. We let the coefficients corresponding to $D_{Curvelet}$ in $\tilde{\theta}_{HF}^{k}$ to zeros to obtain $\tilde{\theta}_{HF\_Rain}^{k}$, while the coefficients corresponding to D_{Rain} in $\tilde{\theta}_{HF}^{k}$ to zeros to obtain $\tilde{\theta}_{HF\_Non-Rain}^{k}$. Therefore, each patch b_{HF}^{k} can be re-expressed as either $\tilde{b}_{HF\_Non-Rain}^{k} = D_{Curvelet} \times \tilde{\theta}_{HF\_Non-Rain}^{k}$ or $\tilde{b}_{HF\_Rain}^{k} = D_{Rain} \times \tilde{\theta}_{HF\_Rain}^{k}$, which can be used to recover $I_{HF}^{Non-Rain}$ or I_{HF}^{Rain}, respectively, by averaging the pixel values in overlapping regions. Finally, the rain-removed version of I can be obtained via $I^{Non-Rain} = I_{LF} + I_{HF}^{Non-Rain}$, as illustrated in Fig. 2(h).

4. SIMULATION RESULTS

To evaluate the performance of the proposed rain removal algorithm, we compared the proposed scheme with the bilateral filter proposed in [8] and the MCA-based image decomposition scheme [6]. The rain removal results obtained from the three evaluated schemes for the three test images are shown in Figs. 2-4, respectively, which demonstrate that the proposed scheme significantly outperforms the other two schemes.

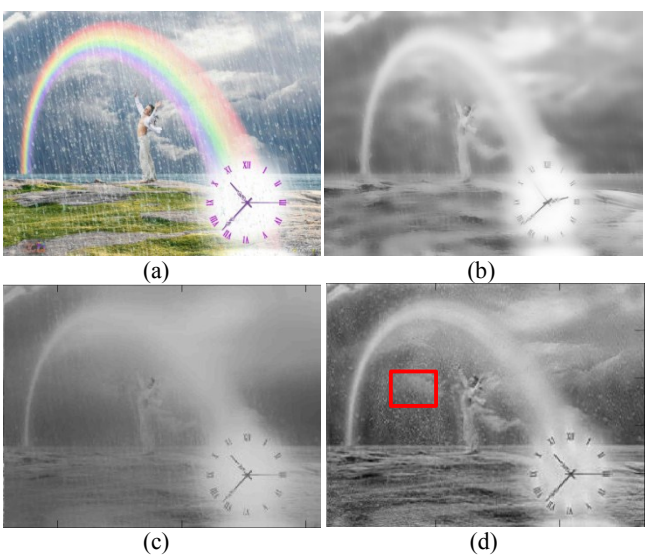

Fig. 3. Comparison of rain removal results: **(a)** the original rain image; the rain-removed versions via: **(b)** the bilateral filter [8]; **(c)** the MCA-based image decomposition [6]; and **(d)** the proposed scheme.

As illustrated in Fig. 2(c), although the bilateral filter [8] can remove most rain streaks, it simultaneously removes other image detail as well. With the MCA-based image decomposition scheme [6], most image details, together with rain streaks, will be always filtered out in all the three test cases. The proposed scheme successfully removes most rain streaks while preserving most original image details (*e.g.*, the bubbles in the red box in Fig. 3(d)). More test results can be found in [10].

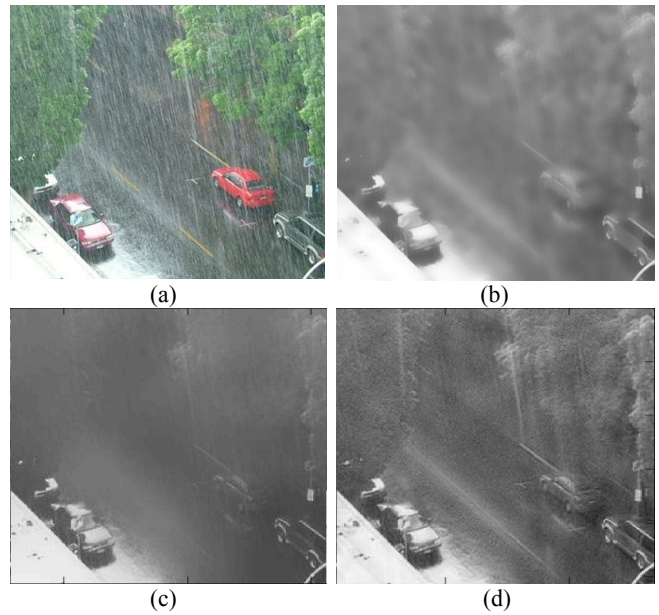

Fig. 4. Comparison of rain removal results: (a) the original image; the rain-removed versions with: **(b)** the bilateral filter [8]; **(c)** the MCA-based image decomposition [6]; **(d)** the proposed scheme.

5. CONCLUSION

In this paper, we have proposed a single-frame-based rain removal framework via formulating rain removal as an image decomposition problem solved by performing sparse coding and dictionary learning algorithms. Our experimental results show that the proposed scheme can effectively remove rain steaks without significantly blurring the original image details. For future work, the performance may be further improved by enhancing the sparse coding and dictionary learning steps. Moreover, the proposed scheme may be extended to remove other kinds of repeated textures.

REFERENCES

[1] K. Garg and S. K. Nayar, "Detection and removal of rain from videos," *Proc. CVPR*, June 2004, vol. 1, pp. 528-535.
[2] K. Garg and S. K. Nayar, "When does a camera see rain?" *Proc. ICCV*, Oct. 2005, vol. 2, pp. 1067-1074.
[3] X. Zhang, et al., "Rain removal in video by combining temporal and chromatic properties," *Proc. ICME*, July 2006, pp. 461-464.
[4] 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.
[5] J. M. Fadili, J. L. Starck, J. Bobin, and Y. Moudden, "Image decomposition and separation using sparse representations: an overview," *Proc. of the IEEE*, vol. 98, no. 6, pp. 983-994, June 2010.
[6] G. Peyré, J. Fadili, and J. L. Starck, "Learning adapted dictionaries for geometry and texture separation," *Proc. SPIE*, vol. 6701, 2007.
[7] M. Aharon, M. Elad, and A. M. Bruckstein, "The K-SVD: An algorithm for designing of overcomplete dictionaries for sparse representation," *IEEE Trans. Signal Process.*, vol. 54, 2006.
[8] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," *Proc. ICCV*, Bombay, India, Jan. 1998, pp. 839-846.
[9] S. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Trans. Signal Process.*, vol. 41, no. 12, Dec. 1993.
[10] NTHU Rain Removal project. [Online]. Available: http://www.ee.nthu.edu.tw/cwlin/Rain_Removal/Rain_Removal.htm.