TEMPORALLY COHERENT REAL-TIME VIDEO DEHAZING

Jin-Hwan Kim[1], *Won-Dong Jang*[1], *Yongsup Park*[2], *Dong-Hahk Lee*[2], *Jae-Young Sim*[3], *and Chang-Su Kim*[1]

[1] School of EE, Korea University, Korea
[2] Intelligent Video Tech. Lab., SK Telecom, Korea
[3] School of ECE, Ulsan National Institute of Science and Technology, Korea
E-mails: {arite, dotol1216, changsukim}@korea.ac.kr,
{yongsup.park, adhlee}@sk.com, jysim@unist.ac.kr

ABSTRACT

A real-time video dehazing algorithm, which reduces flickering artifacts and yields high quality output videos, is proposed in this work. Assuming that a scene point yields highly correlated transmission values between adjacent image frames, we develop the temporal coherence cost. Then, we add the temporal coherence cost to the contrast cost and the truncation loss cost to define the overall cost function. By minimizing the overall cost function, we obtain the optimal transmission. Moreover, to reduce the computational complexity and facilitate real-time applications, we approximate the conventional edge preserving filter by the overlapped block filter. Experimental results demonstrate that the proposed algorithm is sufficiently fast for real-time applications and effectively removes haze and flickering artifacts.

Index Terms— Video enhancement, video dehazing, contrast enhancement, and temporal coherence.

1. INTRODUCTION

Haze in the atmosphere attenuates the radiance of a scene point according to its depth, usually degrading image contrast. In order to enhance low contrast hazy images, we can employ various contrast enhancement techniques, such as histogram equalization or unsharp masking [1]. These techniques, however, stretch the histogram distribution of an image without considering the haze thickness. Recently, several algorithms have been proposed to remove the haze of an image adaptively [2, 3, 4]. Since haze thickness is related to scene depth, these dehazing algorithms first estimate scene depth using additional information. For example, multiple images with different weather conditions [2] or differently polarized images [3] are employed for the depth estimation. Alternatively, predetermined geometry information is used in [4]. Even though these algorithms provide acceptable dehazing results, their applicability is limited due to the requirement of additional information.

To overcome this drawback, single image dehazing algorithms have been proposed [5, 6, 7, 8, 9]. They assume that pixels in a local block have the same amount of haze, or equivalently the same depth. Therefore, blocking artifacts occur, and the edge preserving filtering [5] or the Markov random field (MRF) modeling is employed to refine depth values at the pixel level. However, the pixel-level processing increases the computational complexity and thus may not be suitable for enhancing a sequence of hazy images, *i.e.*, a hazy video. Moreover, since these algorithms focus on static image dehazing, they may yield flickering artifacts due to the lack of temporal coherence, when applied to video dehazing. Zhang *et al*. proposed a temporal MRF model to alleviate flickering artifacts [10]. However, their algorithm requires high complexity and is not applicable to real-time video dehazing.

In this work, we propose a fast and efficient video dehazing algorithm. We assume that a hazy video is temporally coherent and thus the transmission values of an object are similar between adjacent image frames. Based on the assumption, we design the temporal coherence cost, and add it to the contrast and loss costs to define the overall cost. By minimizing the overall cost, we obtain the optimal transmission value. Moreover, we approximate the edge preserving filter with the overlapped block filter to reduce the computational complexity. Experimental results demonstrate that the proposed algorithm achieves real-time video processing, while providing high quality dehazing results.

This paper is organized as follows. Section 2 describes the static image dehazing algorithm, and Section 3 proposes the video dehazing algorithm. Section 4 presents experimental results. Finally, Section 5 concludes the paper.

2. STATIC IMAGE DEHAZING

The optical model of haze is formulated as

$$I_p = t_p \cdot J_p + (1 - t_p) \cdot A, \tag{1}$$

where J_p and I_p denote the original scene radiance and the observed radiance at pixel p, respectively [2, 6]. A is the ambient light in the atmosphere, referred to as the airlight. Also t_p is the transmission at p, determined by the distance between the scene point and the camera. Note that the observed radiance I_p is the weighted sum of the original scene radiance J_p attenuated by t_p and the airlight A weighted by $(1 - t_p)$.

We improve our previous work in [9] to dehaze the first frame in a hazy video. Assuming that a restored image should yield high contrast, we find the optimal transmission for each block by minimizing a cost function. More specifically, we maximize the contrast of a hazy image while minimizing the truncation loss. Our cost function E has two terms, the contrast cost E_C and the loss cost E_L, which is given by

$$E = E_C + \lambda_L E_L, \tag{2}$$

This research was supported partly by Basic Science Research Program through the NRF of Korea funded by the Ministry of Education, Science and Technology (No. 2012-0000916), partly by a grant of the Emerging Technology R&D Center, SK Telecom, and partly by Basic Science Research Program through the NRF of Korea funded by the MEST (No. 2010-0006595).

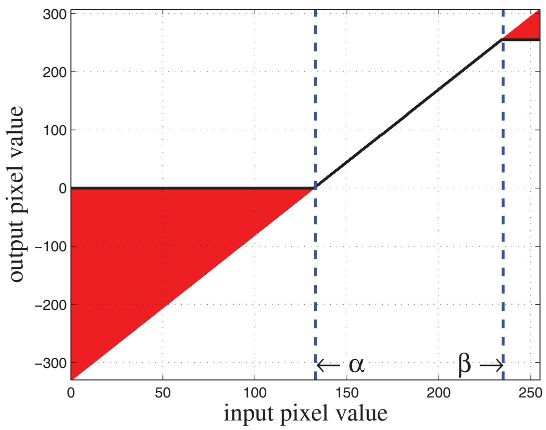

Fig. 1. Computation of the truncation loss cost E_L. The black line is a transformation function of dehazing algorithm, and the red regions depict the truncation errors.

where λ_L is a weighting parameter that controls the relative importance of the two terms.

We define the contrast cost E_C as the variance of restored pixel values multiplied by -1. Based on the haze model in (1), E_C can be simplified to

$$E_C(t) = -\sum_{I=0}^{255} \frac{\left(I - \bar{I}\right)^2 h(I)}{Nt^2}, \tag{3}$$

where $h(I)$ denotes the histogram value of pixel intensity I, $\bar{I}$ is the average pixel intensity, and N is the number of pixels in a block. Notice that E_C is a monotonic decreasing function of transmission t. Therefore, if we minimize only the contrast cost, the optimal transmission becomes 0.

In addition to the contrast cost, by modifying the histogram uniformness in [9], we define the truncation loss cost E_L as

$$\begin{aligned} E_L(t) =& \frac{1}{N}\sum_{I=0}^{\alpha}\left(0 - \left(\frac{I-A}{t} + A\right)\right)^2 h(I) \\ &+ \frac{1}{N}\sum_{I=\beta}^{255}\left(\left(\frac{I-A}{t} + A\right) - 255\right)^2 h(I), \end{aligned} \tag{4}$$

where α and β are the truncation points, as illustrated in Fig. 1. Note that, to maintain the valid range of pixel intensity, the transformed pixel values from the domains of $[0, \alpha]$ and $[\beta, 255]$ are truncated to 0 and 255, respectively. Therefore, the red areas in Fig. 1 correspond to truncation errors. In (4), we multiply the squared truncation errors with the corresponding histogram values to calculate the loss cost.

We select the brightest pixel value in an opaque region of an input image as the airlight A, determine the optimal transmission that minimizes the cost function (2), and then apply a filter to convert the block-based transmission values into the pixel-based ones. Finally, given the transmission of each pixel, we can recover the original scene radiance from (1).

3. VIDEO DEHAZING

The dehazing algorithm in Section 2 provides good results on static images. However, when applied to each frame of a hazy video independently, it may break the temporal coherence and yield a restored video with severe flicking artifacts. Moreover, its high computational complexity prohibits real-time applications. In this section, we propose a fast and efficient video dehazing algorithm.

3.1. Temporal Coherence

Let us first consider the relationship between the transmission values of consecutive image frames. The transmission values change due to object motions. As an object approaches the camera, the observed radiance gets closer to the original scene radiance. On the contrary, when an object moves away from the camera, the observed radiance becomes more similar to the airlight. Thus, we should modify the transmission of a scene point adaptively according to its brightness change.

Let J_p^f and I_p^f be the scene radiance and the observed radiance, respectively, at pixel p of the f-th image frame. It is reasonable to assume that the scene radiance of a scene point is the same over all image frames. In particular,

$$J_p^{f-1} = J_p^f. \tag{5}$$

Then, from (1), we can easily obtain the relationship between the transmission t_p^f of the current frame and the transmission t_p^{f-1} of the previous frame, given by

$$t_p^f = \kappa_p^f t_p^{f-1} \tag{6}$$

where κ_p^f is

$$\kappa_p^f = \frac{I_p^f - A}{I_p^{f-1} - A}. \tag{7}$$

Thus, κ_p^f is the temporal coherence coefficient, which corrects the transmission according to the brightness change.

In (6), we compare two pixels at the same position in the f-th frame and the $(f-1)$-th frame. However, an object may move and the same scene point may be captured at different pixel positions. To address this issue, we employ a probability model, based on the differential image between the two frames, which is given by

$$w_p^f = \exp\left(-\frac{(I_p^f - I_p^{f-1})^2}{\sigma^2}\right) \tag{8}$$

where σ is empirically selected as 10. Note that w_p^f gets larger as I_p^f becomes more similar to I_p^{f-1}. Thus, w_p^f represents the probability that the two pixels are the matching ones. Then, we define the temporal coherence coefficient $\bar{\kappa}^f$ for block B as

$$\bar{\kappa}^f = \frac{\sum_{p\in B} w_p^f \kappa_p^f}{\sum_{p\in B} w_p^f}. \tag{9}$$

In other words, the pixel-based coefficient κ_p^f is weighted by a larger coefficient in the block-based computation, when I_p^f and I_p^{f-1} are more likely to come from the same scene point.

3.2. Cost Function Optimization

We define the temporal coherence cost for each block as

$$E_T = (t^f - \bar{\kappa}^f t^{f-1})^2 \tag{10}$$

which is the squared difference of the transmission t^f of the block in the current frame from its estimation $\bar{\kappa}^f t^{f-1}$ using the previous frame. Then, we reformulate the overall cost in (2) by adding E_T,

$$E = E_C + \lambda_L E_L + \lambda_T E_T \tag{11}$$

(a) Input hazy image

(b) Initial transmission

(c) Sliding window filtering

(d) Overlapped block filtering

Fig. 2. Blocking artifact reduction using the overlapped block filtering.

where λ_T is a weighting parameter. As λ_T gets larger, we emphasize the temporal coherence cost more strongly and hence can alleviate flickering artifacts more effectively. However, in such a case, the optimal transmission value tends to be fixed over all frames, causing blurring artifacts and degrading the quality of each individual image.

We find the optimal transmission t^0 of each block in the first frame by minimizing the cost function in (2), since there is no previous frame. On the other hand, for subsequent frames, we obtain the optimal transmission t^f of the block by minimizing the overall cost function in (11).

3.3. Complexity Reduction

The edge preserving filtering in [5], which assigns a transmission value to each pixel, requires very high computational complexity in general. We attempt to reduce the complexity by approximating the edge preserving filtering. Based on the guided filter [11], we approximate the transmission value by an affine function of the pixel intensity, given by

$$\hat{t}_p = aI_p + b \tag{12}$$

where $\hat{t}_p$ is the approximated transmission. Then, the affine parameters a and b are obtained by minimizing the approximation errors over a sliding window W. Specifically, the optimal parameters a^* and b^* are obtained using the least squares method,

$$(a^*, b^*) = \arg\min_{a,b} \sum_{p \in W} (\hat{t_p} - t_p)^2. \tag{13}$$

If this operation is performed for each pixel p, it requires high complexity. We can significantly reduce the complexity by sampling the evaluation points for (13), but blocking artifacts occur near the boundaries of sliding window. To reduce such artifacts, we adopt the notion of the overlapped sub-block scheme [12]. Fig. 2 shows that the proposed algorithm removes most blocking artifacts in the sliding window filtering by employing the overlapped block filtering. Moreover, the proposed algorithm reduces the size of an input image when computing the transmission to further reduce the complexity.

Table 1. Comparison of the processing speeds in terms of frames per second (fps).

Video size	[9]	Proposed algorithm w/o parallel proc.	Proposed algorithm w parallel proc.
640×360	5.03	27.28	59.41
640×480	4.10	19.89	48.54
720×480	3.66	19.48	41.07

4. EXPERIMENTAL RESULTS

We first evaluate the computational complexity of the proposed algorithm on test videos with different sizes. Table 1 compares the processing speed of the proposed algorithm with that of the previous image dehazing method [9]. A personal computer with an Intel 3.3GHz quad core processor is employed in this test. The method in [9] employs the pixel-based guided filter and takes $0.2 \sim 0.3$ seconds to dehaze an image frame. On the other hand, the proposed algorithm approximates the filter and increases the speed by a factor of about 5. Moreover, when we implement parallel processing routines using the openMP library and the SIMD technique, we achieve the processing speed of about 50 frames per second (fps) on 640×480 video sequences.

Next, we compare the qualities of restored videos. Fig. 3 and Fig. 4 show the results on the test videos ”Intersection” and ”Road View,” respectively. We fix the parameters λ_L and λ_T in (11) to 1.0 and 1.0. It is observed that the previous method [9] yields severe flickering artifacts, for example, on the road region in Fig. 3(b). Moreover, although it is not easily observed from juxtaposed still images, the proposed algorithm without temporal coherence cost also yields flickering artifacts in Fig. 3(d) and Fig. 4(d). However, with the temporal coherence cost, the proposed algorithm effectively removes the flickering artifacts as well as the haze in Fig. 3(e) and Fig. 4(e).

In Fig. 3 and Fig. 4, we also provide the result images of the Tarel and Hautiere's algorithm [8], which yields the fastest processing speed among the conventional image dehazing methods, about 0.15 seconds to dehaze an image frame. However, as shown in Fig. 3(c) and Fig. 4(c), [8] cannot remove the haze sufficiently. The resultant video clips are available in [13].

5. CONCLUSIONS

In this work, we proposed a real-time video dehazing algorithm. We introduced the notion of temporal coherence into the cost function to yield a dehazed video without flickering artifacts. Moreover, we developed the overlapped block filter to replace the edge preserving filter and reduce the computational complexity significantly. Experimental results demonstrated that the proposed algorithm can efficiently restore a hazy video at low computational complexity. However, the overlapped block filter sometimes blurs the transmission near edges, which should be addressed as a future research issue. Furthermore, we will simplify the cost function to a closed-form to obtain the optimal transmission more efficiently.

6. REFERENCES

[1] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 3rd ed. Prentice-Hall, 2007.

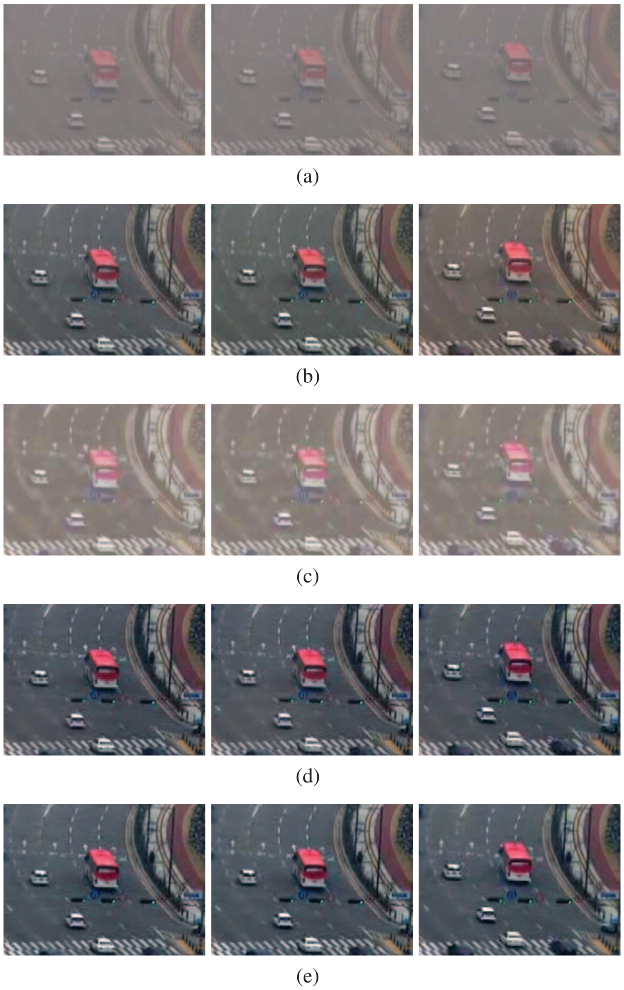

Fig. 3. Comparison of the dehazed video "Intersection." The frame numbers of the left, middle and right columns are 7, 8 and 14, respectively. (a) Input hazy video. The dehazing results of (b) the previous method [9], (c) the Tarel and Hautiere's algorithm [8], (d) the proposed algorithm without the temporal coherence cost and (e) the proposed algorithm with the temporal coherence cost.

Fig. 4. Comparison of the dehazed video "Road View." The frame numbers of the left, middle and right columns are 7, 8 and 12, respectively. (a) Input hazy video. The dehazing results of (b) the previous method [9], (c) the Tarel and Hautiere's algorithm [8], (d) the proposed algorithm without the temporal coherence cost and (e) the proposed algorithm with the temporal coherence cost.

[2] S. G. Narasimhan and S. K. Nayar, "Vision and the atmosphere," *Int. J. Comput. Vis.*, vol. 48, no. 3, pp. 233–254, Jul. 2002.

[3] S. Shwartz, E. Namer, and Y. Y. Schechner, "Blind haze separation," in *Proc. IEEE CVPR*, Jun. 2006, pp. 1984–1991.

[4] J. Kopf, B. Neubert, B. Chen, M. Cohen, D. Cohen-Or, O. Deussen, M. Uyttendaele, and D. Lischinski, "Deep photo: Model-based photograph enhancement and viewing," *ACM Trans. Graph.*, vol. 27, no. 5, pp. 1–10, Dec. 2008.

[5] K. He, J. Sun, and X. Tang, "Single image haze removal using dark channel prior," in *Proc. IEEE CVPR*, Jun. 2009, pp. 1956–1963.

[6] R. Fattal, "Single image dehazing," *ACM Trans. Graph.*, vol. 27, no. 3, pp. 1–9, Aug. 2008.

[7] R. T. Tan, "Visibility in bad weather from a single image," in *Proc. IEEE CVPR*, Jun. 2008, pp. 1–8.

[8] J. P. Tarel and N. Hautiere, "Fast visibility restoration from a single color or gray level image," in *Proc. IEEE ICCV*, Sep. 2009, pp. 2201–2208.

[9] J.-H. Kim, J.-Y. Sim, and C.-S. Kim, "Single image dehazing based on contrast enhancement," in *Proc. IEEE ICASSP*, May 2010, pp. 1273–1276.

[10] J. Zhang, L. Li, Y. Zhang, G. Yang, X. Cao, and J. Sun, "Video dehazing with spatial and temporal coherence," *The Visual Computer*, vol. 27, no. 6, pp. 749–757, Jun. 2011.

[11] K. He, J. Sun, and X. Tang, "Guided image filtering," in *Proc. ECCV*, Sep. 2010, pp. 1–14.

[12] J.-Y. Kim, L.-S. Kim, and S.-H. Hwang, "An advanced contrast enhancement using partially overlapped sub-block histogram equalization," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 11, no. 4, pp. 475–484, Apr. 2001.

[13] Supplementary video clips, Sep. 2012. [Online]. Available: http://mcl.korea.ac.kr/~rokkim/dehazing