Single Image Haze Removal Using Dark Channel Prior

Kaiming He, Jian Sun, and Xiaoou Tang, *Fellow*, *IEEE*

Abstract—In this paper, we propose a simple but effective image prior—dark channel prior to remove haze from a single input image. The dark channel prior is a kind of statistics of outdoor haze-free images. It is based on a key observation—most local patches in outdoor haze-free images contain some pixels whose intensity is very low in at least one color channel. Using this prior with the haze imaging model, we can directly estimate the thickness of the haze and recover a high-quality haze-free image. Results on a variety of hazy images demonstrate the power of the proposed prior. Moreover, a high-quality depth map can also be obtained as a byproduct of haze removal.

Index Terms—Dehaze, defog, image restoration, depth estimation.

1 Introduction

Images of outdoor scenes are usually degraded by the turbid medium (e.g., particles and water droplets) in the atmosphere. Haze, fog, and smoke are such phenomena due to atmospheric absorption and scattering. The irradiance received by the camera from the scene point is attenuated along the line of sight. Furthermore, the incoming light is blended with the *airlight* [1]—ambient light reflected into the line of sight by atmospheric particles. The degraded images lose contrast and color fidelity, as shown in Fig. 1a. Since the amount of scattering depends on the distance of the scene points from the camera, the degradation is spatially variant.

Haze removal[1] (or dehazing) is highly desired in consumer/computational photography and computer vision applications. First, removing haze can significantly increase the visibility of the scene and correct the color shift caused by the airlight. In general, the haze-free image is more visually pleasing. Second, most computer vision algorithms, from low-level image analysis to high-level object recognition, usually assume that the input image (after radiometric calibration) is the scene radiance. The performance of many vision algorithms (e.g., feature detection, filtering, and photometric analysis) will inevitably suffer from the biased and low-contrast scene radiance. Last, haze removal can provide depth information and benefit many vision algorithms and advanced image editing. Haze or fog can be a useful depth clue for scene understanding. A bad hazy image can be put to good use.

However, haze removal is a challenging problem because the haze is dependent on the unknown depth. The problem is underconstrained if the input is only a single hazy image. Therefore, many methods have been proposed by using multiple images or additional information. Polarization-based methods [3], [4] remove the haze effect through two or more images taken with different degrees of polarization. In [5], [6], [7], more constraints are obtained from multiple images of the same scene under different weather conditions. Depth-based methods [8], [9] require some depth information from user inputs or known 3D models.

Recently, single image haze removal has made significant progresses [10], [11]. The success of these methods lies on using stronger priors or assumptions. Tan [11] observes that a haze-free image must have higher contrast compared with the input hazy image and he removes haze by maximizing the local contrast of the restored image. The results are visually compelling but may not be physically valid. Fattal [10] estimates the albedo of the scene and the medium transmission under the assumption that the transmission and the surface shading are locally uncorrelated. This approach is physically sound and can produce impressive results. However, it cannot handle heavily hazy images well and may fail in the cases where the assumption is broken.

In this paper, we propose a novel prior—*dark channel prior*—for single image haze removal. The dark channel prior is based on the statistics of outdoor haze-free images. We find that, in most of the local regions which do not cover the sky, some pixels (called *dark pixels*) very often have very low intensity in at least one color (RGB) channel. In hazy images, the intensity of these dark pixels in that channel is mainly contributed by the airlight. Therefore, these dark pixels can directly provide an accurate estimation of the haze transmission. Combining a haze imaging model and a soft matting interpolation method, we can recover a high-quality haze-free image and produce a good depth map.

Our approach is physically valid and is able to handle distant objects in heavily hazy images. We do not rely on

1. Haze, fog, and smoke differ mainly in the material, size, shape, and concentration of the atmospheric particles. See [2] for more details. In this paper, we do not distinguish these similar phenomena and use the term *haze removal* for simplicity.

- *K. He and X. Tang are with the Department of Information Engineering, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong, China. E-mail: {hkm007, xtang}@ie.cuhk.edu.hk.*
- *J. Sun is with Microsoft Research Asia, F5, #49, Zhichun Road, Haldian District, Beijing, China. E-mail: jiansun@microsoft.com.*

Manuscript received 17 Dec. 2009; revised 24 June 2010; accepted 26 June 2010; published online 31 Aug. 2010.
Recommended for acceptance by S.B. Kang.
For information on obtaining reprints of this article, please send e-mail to: tpami@computer.org, and reference IEEECS Log Number TPAMISI-2009-12-0832.
Digital Object Identifier no. 10.1109/TPAMI.2010.168.

Fig. 1. Haze removal using a single image. (a) Input hazy image. (b) Image after haze removal by our approach. (c) Our recovered depth map.

significant variance of transmission or surface shading. The result contains few halo artifacts.

Like any approach using a strong assumption, our approach also has its own limitation. The dark channel prior may be invalid when the scene object is inherently similar to the airlight (e.g., snowy ground or a white wall) over a large local region and no shadow is cast on it. Although our approach works well for most outdoor hazy images, it may fail on some extreme cases. Fortunately, in such situations haze removal is not critical since haze is rarely visible. We believe that developing novel priors from different directions and combining them together will further advance the state of the art.

2 Background

In computer vision and computer graphics, the model widely used to describe the formation of a hazy image is [2], [5], [10], [11]:

$$\mathbf{I}(\mathbf{x}) = \mathbf{J}(\mathbf{x})t(\mathbf{x}) + \mathbf{A}(1 - t(\mathbf{x})), \tag{1}$$

where $\mathbf{I}$ is the observed intensity, $\mathbf{J}$ is the scene radiance, $\mathbf{A}$ is the global atmospheric light, and t is the medium transmission describing the portion of the light that is not scattered and reaches the camera. The goal of haze removal is to recover $\mathbf{J}$, $\mathbf{A}$, and t from $\mathbf{I}$. For an N-pixel color image $\mathbf{I}$, there are $3N$ constraints and $4N + 3$ unknowns. This makes the problem of haze removal inherently ambiguous.

In (1), the first term $\mathbf{J}(\mathbf{x})t(\mathbf{x})$ on the right-hand side is called *direct attenuation* [11], and the second term $\mathbf{A}(1 - t(\mathbf{x}))$ is called *airlight* [1], [11]. The direct attenuation describes the scene radiance and its decay in the medium, and the airlight results from previously scattered light and leads to the shift of the scene colors. While the direct attenuation is a *multiplicative* distortion of the scene radiance, the airlight is an *additive* one.

When the atmosphere is homogenous, the transmission t can be expressed as

$$t(\mathbf{x}) = e^{-\beta d(\mathbf{x})}, \tag{2}$$

where β is the scattering coefficient of the atmosphere and d is the scene depth. This equation indicates that the scene radiance is attenuated exponentially with the depth. If we can recover the transmission, we can also recover the depth up to an unknown scale.

Geometrically, the haze imaging equation (1) means that in RGB color space, the vectors $\mathbf{A}$, $\mathbf{I}(\mathbf{x})$, and $\mathbf{J}(\mathbf{x})$ are coplanar and their end points are collinear (see Fig. 2a). The transmission t is the ratio of two line segments:

$$t(\mathbf{x}) = \frac{\|\mathbf{A} - \mathbf{I}(\mathbf{x})\|}{\|\mathbf{A} - \mathbf{J}(\mathbf{x})\|} = \frac{A^c - I^c(\mathbf{x})}{A^c - J^c(\mathbf{x})}, \tag{3}$$

where $c \in \{r, g, b\}$ is the color channel index.

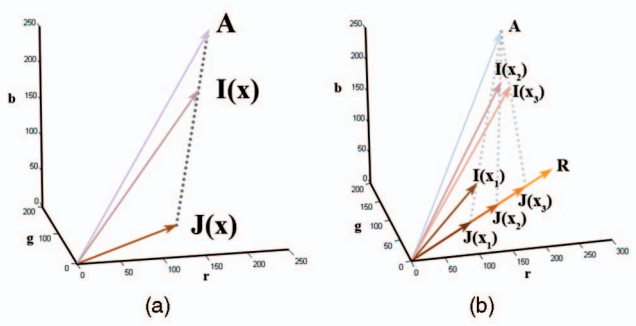

Fig. 2. (a) Haze imaging model. (b) Constant albedo model used in Fattal's work [10].

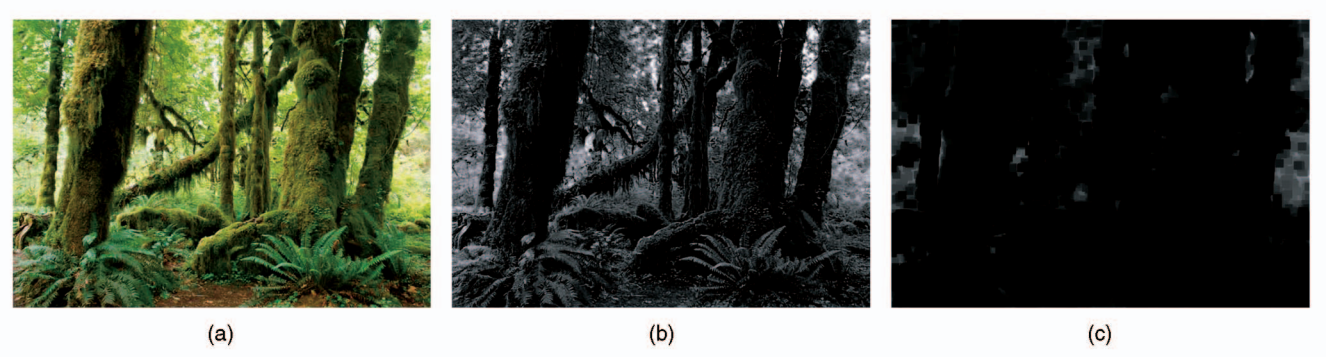

Fig. 3. Calculation of a dark channel. (a) An arbitrary image $\mathbf{J}$. (b) For each pixel, we calculate the minimum of its (r, g, b) values. (c) A minimum filter is performed on (b). This is the dark channel of $\mathbf{J}$. The image size is 800×551, and the patch size of Ω is 15×15.

Based on this model, Tan's method [11] focuses on enhancing the visibility of the image. For a patch with uniform transmission t, the visibility (sum of gradient) of the input image is reduced by the haze since $t < 1$:

$$\sum_{\mathbf{x}} \|\nabla \mathbf{I}(\mathbf{x})\| = t \sum_{\mathbf{x}} \|\nabla \mathbf{J}(\mathbf{x})\| < \sum_{\mathbf{x}} \|\nabla \mathbf{J}(\mathbf{x})\|. \quad (4)$$

The transmission t in a local patch is estimated by maximizing the visibility of the patch under a constraint that the intensity of $\mathbf{J}(\mathbf{x})$ is less than the intensity of $\mathbf{A}$. An MRF model is used to further regularize the result. This approach is able to greatly unveil details and structures from hazy images. However, the output images usually tend to have larger saturation values because this method focuses solely on the enhancement of the visibility and does not intend to physically recover the scene radiance. Besides, the result may contain halo effects near the depth discontinuities.

In [10], Fattal proposes an approach based on Independent Component Analysis (ICA). First, the albedo of a local patch is assumed to be a constant vector $\mathbf{R}$. Thus, all vectors $\mathbf{J}(\mathbf{x})$ in the patch have the same direction $\mathbf{R}$, as shown in Fig. 2b. Second, by assuming that the statistics of the surface shading $\|\mathbf{J}(\mathbf{x})\|$ and the transmission $t(\mathbf{x})$ are independent in the patch, the direction of $\mathbf{R}$ is estimated by ICA. Finally, an MRF model guided by the input color image is applied to extrapolate the solution to the whole image. This approach is physics-based and can produce a natural haze-free image together with a good depth map. But, due to the statistical independence assumption, this approach requires that the independent components vary significantly. Any lack of variation or low signal-to-noise ratio (often in dense haze region) will make the statistics unreliable. Moreover, as the statistic is based on color information, it is invalid for gray-scale images and it is difficult to handle dense haze that is colorless.

In the next section, we present a new prior—dark channel prior—to estimate the transmission directly from an outdoor hazy image.

3 Dark Channel Prior

The dark channel prior is based on the following observation on outdoor haze-free images: In most of the nonsky patches, at least one color channel has some pixels whose intensity are very low and close to zero. Equivalently, the minimum intensity in such a patch is close to zero.

To formally describe this observation, we first define the concept of a *dark channel*. For an arbitrary image $\mathbf{J}$, its dark channel J^{dark} is given by

$$J^{\text{dark}}(\mathbf{x}) = \min_{\mathbf{y} \in \Omega(\mathbf{x})} \Big(\min_{c \in \{r,g,b\}} J^c(\mathbf{y}) \Big), \quad (5)$$

where J^c is a color channel of $\mathbf{J}$ and $\Omega(\mathbf{x})$ is a local patch centered at $\mathbf{x}$. A dark channel is the outcome of two minimum operators: $\min_{c \in \{r,g,b\}}$ is performed on each pixel (Fig. 3b), and $\min_{\mathbf{y} \in \Omega(\mathbf{x})}$ is a minimum filter (Fig. 3c). The minimum operators are commutative.

Using the concept of a dark channel, our observation says that if $\mathbf{J}$ is an outdoor haze-free image, except for the sky region, the intensity of $\mathbf{J}$'s dark channel is low and tends to be zero:

$$J^{\text{dark}} \rightarrow 0. \quad (6)$$

We call this observation *dark channel prior*.

The low intensity in the dark channel is mainly due to three factors: a) shadows, e.g., the shadows of cars, buildings, and the inside of windows in cityscape images, or the shadows of leaves, trees, and rocks in landscape images; b) colorful objects or surfaces, e.g., any object with low reflectance in any color channel (for example, green grass/tree/plant, red or yellow flower/leaf, and blue water surface) will result in low values in the dark channel; c) dark objects or surfaces, e.g., dark tree trunks and stones. As the natural outdoor images are usually colorful and full of shadows, the dark channels of these images are really dark!

To verify how good the dark channel prior is, we collect an outdoor image set from Flickr.com and several other image search engines using 150 most popular tags annotated by the Flickr users. Since haze usually occurs in outdoor landscape and cityscape scenes, we manually pick out the haze-free landscape and cityscape ones from the data set. Besides, we only focus on daytime images. Among them, we randomly select 5,000 images and manually cut out the sky regions. The images are resized so that the maximum of width and height is 500 pixels and their dark channels are computed using a patch size 15×15. Fig. 4 shows several outdoor images and the corresponding dark channels.

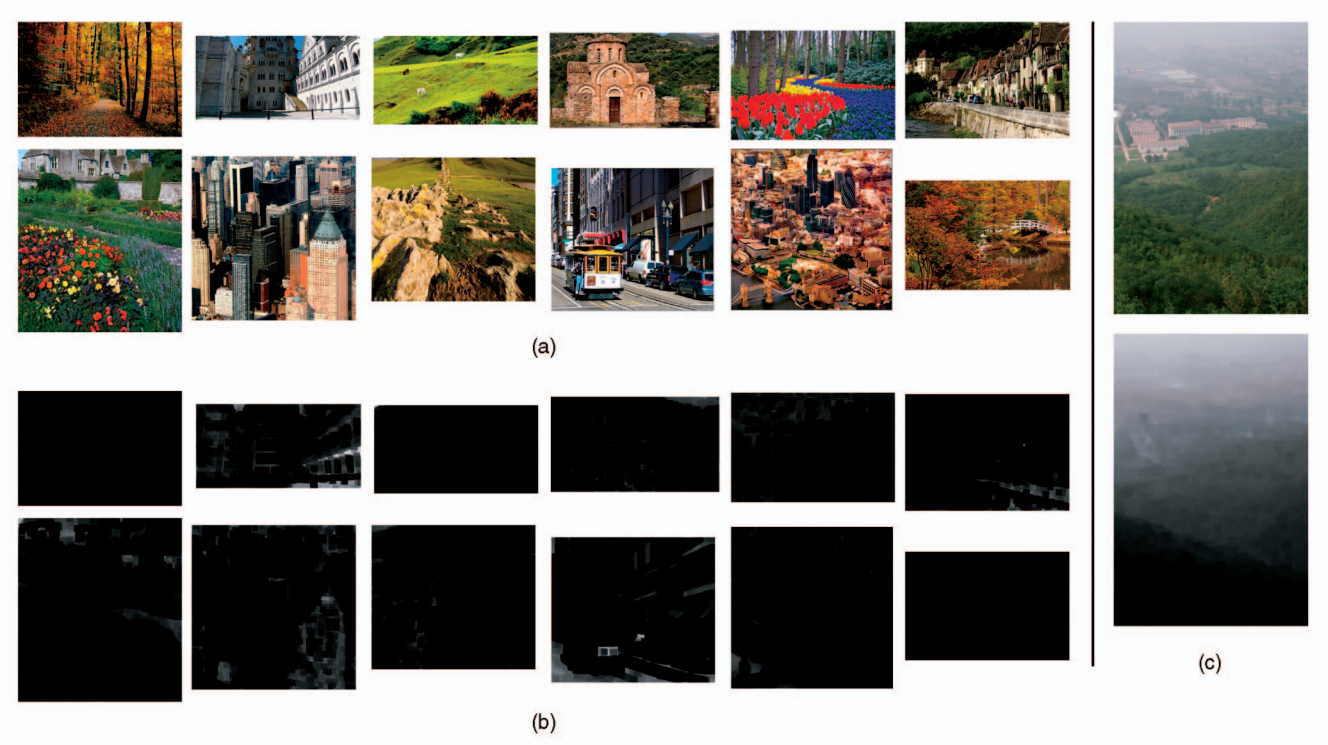

Fig. 4. (a) Example images in our haze-free image database. (b) The corresponding dark channels. (c) A hazy image and its dark channel.

Fig. 5a is the intensity histogram over all 5,000 dark channels and Fig. 5b is the corresponding cumulative distribution. We can see that about 75 percent of the pixels in the dark channels have zero values, and the intensity of 90 percent of the pixels is below 25. This statistic gives very strong support to our dark channel prior. We also compute the average intensity of each dark channel and plot the corresponding histogram in Fig. 5c. Again, most dark channels have very low average intensity, showing that only a small portion of outdoor haze-free images deviate from our prior.

Due to the additive airlight, a hazy image is brighter than its haze-free version where the transmission t is low. So, the dark channel of a hazy image will have higher intensity in regions with denser haze (see the right-hand side of Fig. 4). Visually, the intensity of the dark channel is a rough approximation of the thickness of the haze. In the next section, we will use this property to estimate the transmission and the atmospheric light.

Note that we neglect the sky regions because the dark channel of a haze-free image may have high intensity here. Fortunately, we can gracefully handle the sky regions by using the haze imaging model (1) and our prior together. It is not necessary to cut out the sky regions explicitly. We discuss this issue in Section 4.1.

Our dark channel prior is partially inspired by the well-known dark-object subtraction technique [12] widely used in multispectral remote sensing systems. In [12], spatially homogeneous haze is removed by subtracting a constant value corresponding to the darkest object in the scene. We generalize this idea and propose a novel prior for natural image dehazing.

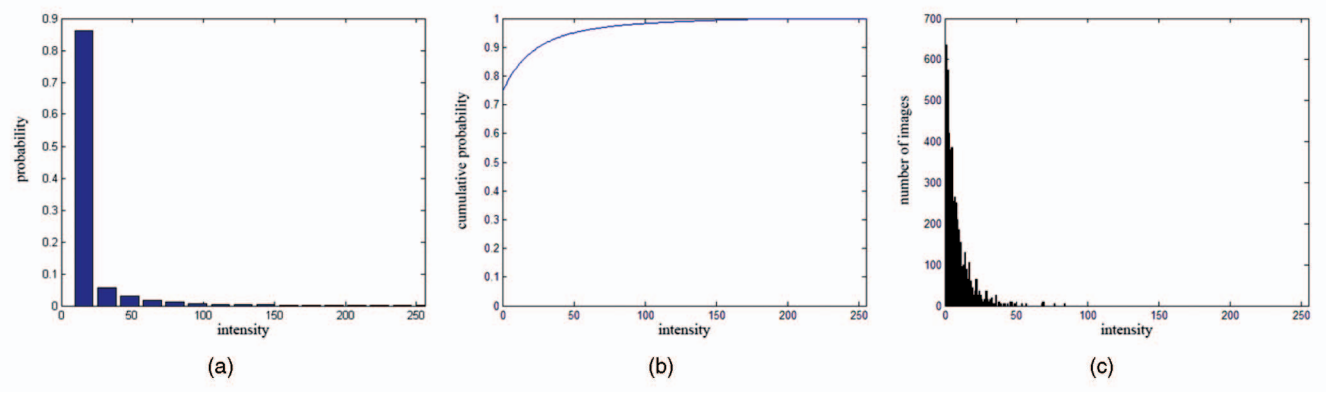

Fig. 5. Statistics of the dark channels. (a) Histogram of the intensity of the pixels in all of the 5,000 dark channels (each bin stands for 16 intensity levels). (b) Cumulative distribution. (c) Histogram of the average intensity of each dark channel.

4 Haze Removal Using Dark Channel Prior

4.1 Estimating the Transmission

We assume that the atmospheric light $\mathbf{A}$ is given. An automatic method to estimate $\mathbf{A}$ is proposed in Section 4.3. We first normalize the haze imaging equation (1) by A:

$$\frac{I^c(\mathbf{x})}{A^c} = t(\mathbf{x})\frac{J^c(\mathbf{x})}{A^c} + 1 - t(\mathbf{x}). \tag{7}$$

Note that we normalize each color channel independently.

We further assume that the transmission in a local patch $\Omega(\mathbf{x})$ is constant. We denote this transmission as $\tilde{t}(\mathbf{x})$. Then, we calculate the dark channel on both sides of (7). Equivalently, we put the minimum operators on both sides:

$$\min_{\mathbf{y}\in\Omega(\mathbf{x})}\left(\min_c \frac{I^c(\mathbf{y})}{A^c}\right) = \tilde{t}(\mathbf{x}) \min_{\mathbf{y}\in\Omega(\mathbf{x})}\left(\min_c \frac{J^c(\mathbf{y})}{A^c}\right) + 1 - \tilde{t}(\mathbf{x}). \tag{8}$$

Since $\tilde{t}(\mathbf{x})$ is a constant in the patch, it can be put on the outside of the min operators.

As the scene radiance $\mathbf{J}$ is a haze-free image, the dark channel of J is close to zero due to the dark channel prior:

$$J^{\text{dark}}(\mathbf{x}) = \min_{\mathbf{y}\in\Omega(\mathbf{x})}\left(\min_c J^c(\mathbf{y})\right) = 0. \tag{9}$$

As A^c is always positive, this leads to

$$\min_{\mathbf{y}\in\Omega(\mathbf{x})}\left(\min_c \frac{J^c(\mathbf{y})}{A^c}\right) = 0. \tag{10}$$

Putting (10) into (8), we can eliminate the multiplicative term and estimate the transmission $\tilde{t}$ simply by

$$\tilde{t}(\mathbf{x}) = 1 - \min_{\mathbf{y}\in\Omega(\mathbf{x})}\left(\min_c \frac{I^c(\mathbf{y})}{A^c}\right). \tag{11}$$

In fact, $\min_{\mathbf{y}\in\Omega(\mathbf{x})}(\min_c \frac{I^c(\mathbf{y})}{A^c})$ is the dark channel of the normalized hazy image $\frac{I^c(\mathbf{y})}{A^c}$. It directly provides the estimation of the transmission.

As we mentioned before, the dark channel prior is not a good prior for the sky regions. Fortunately, the color of the sky in a hazy image $\mathbf{I}$ is usually very similar to the atmospheric light $\mathbf{A}$. So, in the sky region, we have

$$\min_{\mathbf{y}\in\Omega(\mathbf{x})}\left(\min_c \frac{I^c(\mathbf{y})}{A^c}\right) \to 1,$$

and (11) gives $\tilde{t}(\mathbf{x}) \to 0$. Since the sky is infinitely distant and its transmission is indeed close to zero (see (2)), (11) gracefully handles both sky and nonsky regions. We do not need to separate the sky regions beforehand.

In practice, even on clear days the atmosphere is not absolutely free of any particle. So the haze still exists when we look at distant objects. Moreover, the presence of haze is a fundamental cue for human to perceive depth [13], [14]. This phenomenon is called *aerial perspective*. If we remove the haze thoroughly, the image may seem unnatural and we may lose the feeling of depth. So, we can optionally keep a very small amount of haze for the distant objects by introducing a constant parameter ω ($0 < \omega \leq 1$) into (11):

$$\tilde{t}(\mathbf{x}) = 1 - \omega \min_{\mathbf{y}\in\Omega(\mathbf{x})}\left(\min_c \frac{I^c(\mathbf{y})}{A^c}\right). \tag{12}$$

The nice property of this modification is that we adaptively keep more haze for the distant objects. The value of ω is application based. We fix it to 0.95 for all results reported in this paper.

In the derivation of (11), the dark channel prior is essential for eliminating the multiplicative term (direct transmission) in the haze imaging model (1). Only the additive term (airlight) is left. This strategy is completely different from previous single image haze removal methods [10], [11], which rely heavily on the multiplicative term. These methods are driven by the fact that the multiplicative term changes the image contrast [11] and the color variance [10]. On the contrary, we notice that the additive term changes the intensity of the local dark pixels. With the help of the dark channel prior, the multiplicative term is discarded and the additive term is sufficient to estimate the transmission. We can further generalize (1) by

$$\mathbf{I}(\mathbf{x}) = \mathbf{J}(\mathbf{x})t_1(\mathbf{x}) + \mathbf{A}(1 - t_2(\mathbf{x})), \tag{13}$$

where t_1 and t_2 are not necessarily the same. Using the method for deriving (11), we can estimate t_2 and thus separate the additive term. The problem is reduced to a multiplicative form ($\mathbf{J}(\mathbf{x})t_1$), and other constraints or priors can be used to further disentangle this term. In the literature of human vision research [15], the additive term is called a *veiling luminance*, and (13) can be used to describe a scene seen through a veil or glare of highlights.

Fig. 6b shows the estimated transmission maps using (12). Fig. 6d shows the corresponding recovered images. As we can see, the dark channel prior is effective on recovering the vivid colors and unveiling low contrast objects. The transmission maps are reasonable. The main problems are some halos and block artifacts. This is because the transmission is not always constant in a patch. In the next section, we propose a soft matting method to refine the transmission maps.

4.2 Soft Matting

We notice that the haze imaging equation (1) has a similar form as the image matting equation:

$$\mathbf{I} = \mathbf{F}\alpha + \mathbf{B}(1 - \alpha), \tag{14}$$

where $\mathbf{F}$ and $\mathbf{B}$ are foreground and background colors, respectively, and α is the foreground opacity. A transmission map in the haze imaging equation is exactly an alpha map. Therefore, we can apply a closed-form framework of matting [16] to refine the transmission.

Denote the refined transmission map by $t(\mathbf{x})$. Rewriting $t(\mathbf{x})$ and $\tilde{t}(\mathbf{x})$ in their vector forms as $\mathbf{t}$ and $\tilde{\mathbf{t}}$, we minimize the following cost function:

$$E(\mathbf{t}) = \mathbf{t}^{\mathrm{T}}\mathrm{L}\mathbf{t} + \lambda(\mathbf{t} - \tilde{\mathbf{t}})^{\mathrm{T}}(\mathbf{t} - \tilde{\mathbf{t}}). \tag{15}$$

Here, the first term is a smoothness term and the second term is a data term with a weight λ. The matrix L is called the matting Laplacian matrix [16]. Its (i, j) element is defined as

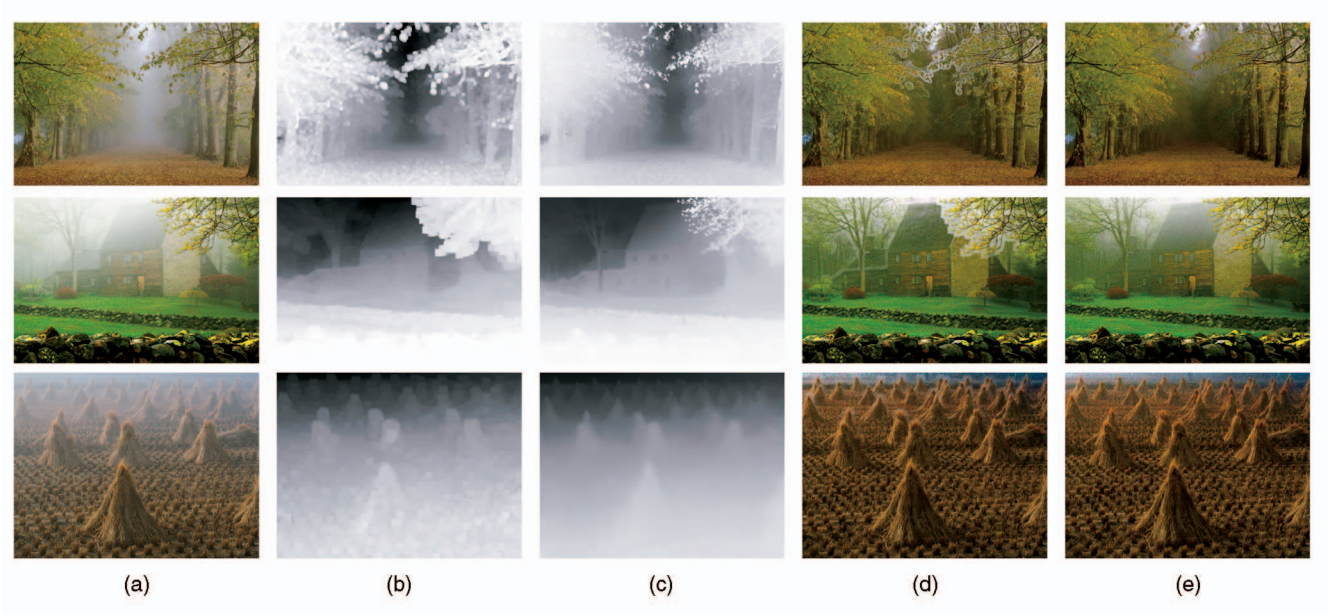

Fig. 6. Haze removal. (a) Input hazy images. (b) Estimated transmission maps before soft matting. (c) Refined transmission maps after soft matting. (d), (e) Recovered images using (b) and (c), respectively.

$$\sum_{k|(i,j)\in w_k}\left(\delta_{ij}-\frac{1}{|w_k|}\left(1+(\mathbf{I}_i-\mu_k)^T\left(\Sigma_k+\frac{\varepsilon}{|w_k|}\mathrm{U}_3\right)^{-1}(\mathbf{I}_j-\mu_k)\right)\right), \tag{16}$$

where $\mathbf{I}_i$ and $\mathbf{I}_j$ are the colors of the input image $\mathbf{I}$ at pixels i and j, δ_{ij} is the Kronecker delta, μ_k and Σ_k are the mean and covariance matrix of the colors in window w_k, U_3 is a 3×3 identity matrix, ε is a regularizing parameter, and $|w_k|$ is the number of pixels in the window w_k.

The optimal $\mathbf{t}$ can be obtained by solving the following sparse linear system:

$$(\mathrm{L}+\lambda\mathrm{U})\mathbf{t}=\lambda\tilde{\mathbf{t}}, \tag{17}$$

where U is an identity matrix of the same size as L. We set a small λ (10^{-4} in the experiments) so that $\mathbf{t}$ is softly constrained by $\tilde{\mathbf{t}}$.

The derivation of the matting Laplacian matrix in [16] is based on a *color line assumption*: The foreground/background colors in a small local patch lie on a single line in the RGB color space. The color line assumption is also valid in the problem of haze removal. First, the scene radiance $\mathbf{J}$ is a natural image. According to [16], [17], the color line model holds for natural images. Second, the atmospheric light $\mathbf{A}$ is a constant, which of course satisfies the assumption. Therefore, it is valid to use the matting Laplacian matrix as the smoothness term in the haze removal problem.

The closed-form matting framework has also been applied by Hsu et al. [18] to deal with the spatially variant white balance problem. In [16], [18], the constraint is known in some sparse regions, and this framework is mainly used to extrapolate the values into the unknown regions. In our application, the coarse constraint $\tilde{\mathbf{t}}$ has already filled the whole image. We consider this constraint as soft, and use the matting framework to refine the map.

After solving the linear system (17), we perform a bilateral filter [19] on t to smooth its small scale textures. Fig. 6c shows the refined results using Fig. 6b as the constraint. As we can see, the halos and block artifacts are suppressed. The refined transmission maps manage to capture the sharp edge discontinuities and outline the profile of the objects.

4.3 Estimating the Atmospheric Light

We have been assuming that the atmospheric light $\mathbf{A}$ is known. In this section, we propose a method to estimate $\mathbf{A}$. In the previous works, the color of the most haze-opaque region is used as $\mathbf{A}$ [11] or as $\mathbf{A}$'s initial guess [10]. However, little attention has been paid to the detection of the "most haze-opaque" region.

In Tan's work [11], the brightest pixels in the hazy image are considered to be the most haze-opaque. This is true only when the weather is overcast and the sunlight can be ignored. In this case, the atmospheric light is the only illumination source of the scene. So, the scene radiance of each color channel is given by

$$J(\mathbf{x})=R(\mathbf{x})A, \tag{18}$$

where $R\le 1$ is the reflectance of the scene points. The haze imaging equation (1) can be written as

$$I(\mathbf{x})=R(\mathbf{x})At(\mathbf{x})+(1-t(\mathbf{x}))A\le A. \tag{19}$$

When pixels at infinite distance ($t\approx 0$) exist in the image, the brightest I is the most haze-opaque and it approximately equals A. Unfortunately, in practice we can rarely ignore the sunlight. Considering the sunlight S, we modify (18) by

$$J(\mathbf{x})=R(\mathbf{x})(S+A), \tag{20}$$

and (19) by

$$I(\mathbf{x})=R(\mathbf{x})St(\mathbf{x})+R(\mathbf{x})At(\mathbf{x})+(1-t(\mathbf{x}))A. \tag{21}$$

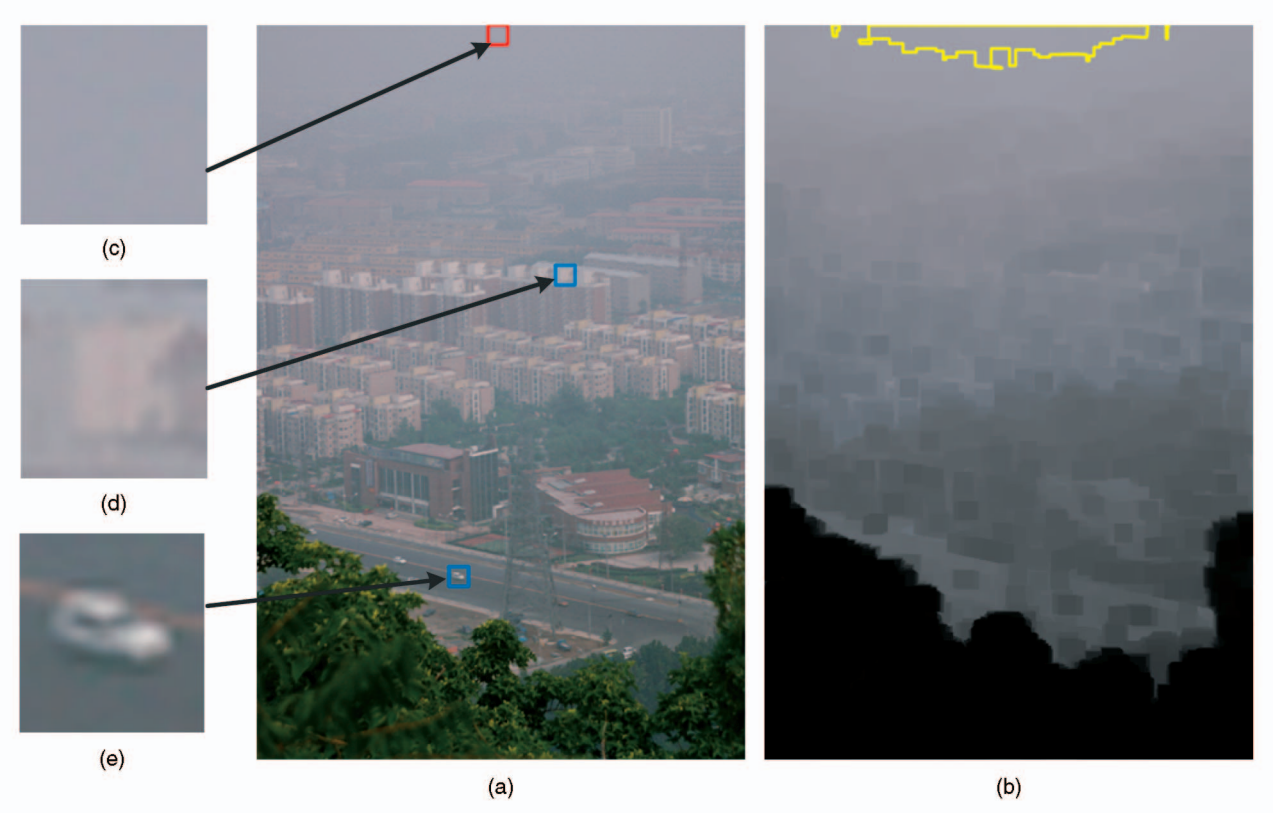

Fig. 7. Estimating the atmospheric light. (a) Input image. (b) Dark channel and the most haze-opaque region. (c) The patch from where our method automatically obtains the atmospheric light. (d), (e) Two patches that contain pixels brighter than the atmospheric light.

In this situation, the brightest pixel of the whole image can be brighter than the atmospheric light. They can be on a white car or a white building (Figs. 7d and 7e).

As discussed in Section 3, the dark channel of a hazy image approximates the haze denseness (see Fig. 7b). So we can use the dark channel to detect the most haze-opaque region and improve the atmospheric light estimation. We first pick the top 0.1 percent brightest pixels in the dark channel. These pixels are usually most haze-opaque (bounded by yellow lines in Fig. 7b). Among these pixels, the pixels with highest intensity in the *input* image $\mathbf{I}$ are selected as the atmospheric light. These pixels are in the red rectangle in Fig. 7a. Note that these pixels may not be brightest ones in the whole input image.

This method works well even when pixels at infinite distance do not exist in the image. In Fig. 8b, our method manages to detect the most haze-opaque regions. However, t is not close to zero here, so the colors of these regions may be different from $\mathbf{A}$. Fortunately, t is small in these most haze-opaque regions, so the influence of sunlight is weak (see (21)). Therefore, these regions can still provide a good approximation of $\mathbf{A}$. The haze removal result of this image is shown in Fig. 8c.

This simple method based on the dark channel prior is more robust than the "brightest pixel" method. We use it to automatically estimate the atmospheric lights for all images shown in this paper.

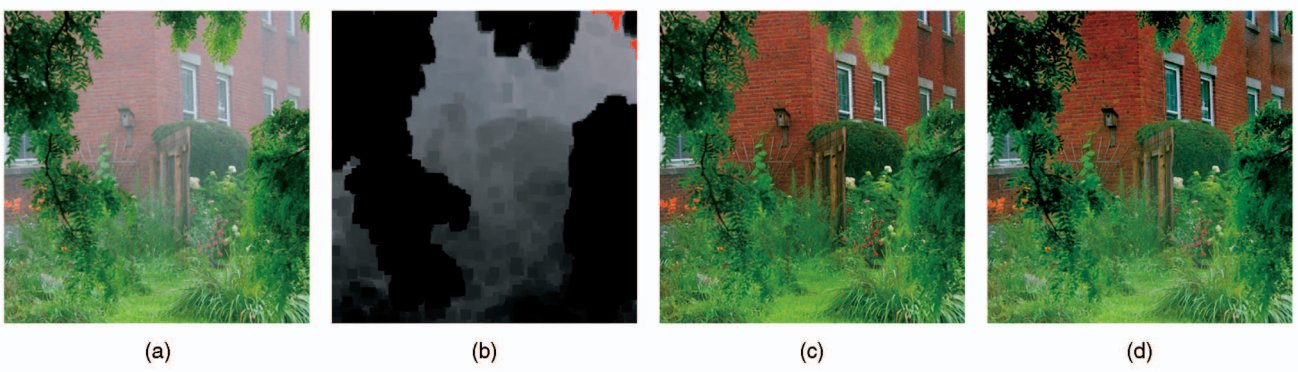

Fig. 8. (a) Input image. (b) Dark channel. The red pixels are the most haze-opaque regions detected by our method. (c) Our haze removal result. (d) Fattal's haze removal result [10].

Fig. 9. A haze-free image (600×400) and its dark channels using 3×3 and 15×15 patches, respectively.

4.4 Recovering the Scene Radiance

With the atmospheric light and the transmission map, we can recover the scene radiance according to (1). But the direct attenuation term $\mathbf{J}(\mathbf{x})t(\mathbf{x})$ can be very close to zero when the transmission $t(\mathbf{x})$ is close to zero. The directly recovered scene radiance $\mathbf{J}$ is prone to noise. Therefore, we restrict the transmission $t(\mathbf{x})$ by a lower bound t_0, i.e., we preserve a small amount of haze in very dense haze regions. The final scene radiance $\mathbf{J}(\mathbf{x})$ is recovered by

$$\mathbf{J}(\mathbf{x}) = \frac{\mathbf{I}(\mathbf{x}) - \mathbf{A}}{\max(t(\mathbf{x}), t_0)} + \mathbf{A}. \tag{22}$$

A typical value of t_0 is 0.1. Since the scene radiance is usually not as bright as the atmospheric light, the image after haze removal looks dim. So we increase the exposure of $\mathbf{J}(\mathbf{x})$ for display. Some final recovered images are shown in Fig. 6e.

4.5 Patch Size

A key parameter in our algorithm is the patch size in (11). On one hand, the dark channel prior becomes better for a larger patch size because the probability that a patch contains a dark pixel is increased. We can see this in Fig. 9: the larger the patch size, the darker the dark channel. Consequently, (9) is less accurate for a small patch, and the recovered scene radiance is oversaturated (Fig. 10b). On the other hand, the assumption that the transmission is constant in a patch becomes less appropriate. If the patch size is too large, halos near depth edges may become stronger (Fig. 10c).

Fig. 11 shows the haze removal results using different patch sizes. The image sizes are 600×400. In Fig. 11b, the patch size is 3×3. The colors of some grayish surfaces look oversaturated (see the buildings in the first row and the subimages in the second and the third rows). In Figs. 11c and 11d, the patch sizes are 15×15 and 30×30, respectively. The results appear more natural than those in Fig. 11b. This shows that our method works well for sufficiently large patch sizes. The soft matting technique is able to reduce the artifacts introduced by large patches. We also notice that the images in Fig. 11d seem to be slightly hazier than those in Fig. 11c (typically in distant regions), but the differences are small. In the remainder of this paper, we use a patch size of 15×15.

5 Experimental Results

In our experiments, we compute the minimum filter by van Herk's fast algorithm [20], whose complexity is linear to the image size. In the soft matting step, we use the Preconditioned Conjugate Gradient (PCG) algorithm as our solver. It takes about 10-20 seconds to process a 600×400 image on a PC with a 3.0 GHz Intel Pentium 4 Processor.

Figs. 1 and 12 show the recovered images and the depth maps. The depth maps are computed according to (2) and are up to an unknown scaling parameter β. The atmospheric lights in these images are automatically estimated using the method described in Section 4.3 (indicated by the red rectangles in Fig. 12). As can be seen, our approach can unveil the details and recover vivid colors even in heavily

Fig. 10. (a) A 600×400 hazy image and the recovered scene radiance using (b) 3×3 and (c) 15×15 patches, respectively (without soft matting). The recovered scene radiance is oversaturated for a small patch size, while it contains apparent halos for a large patch size.

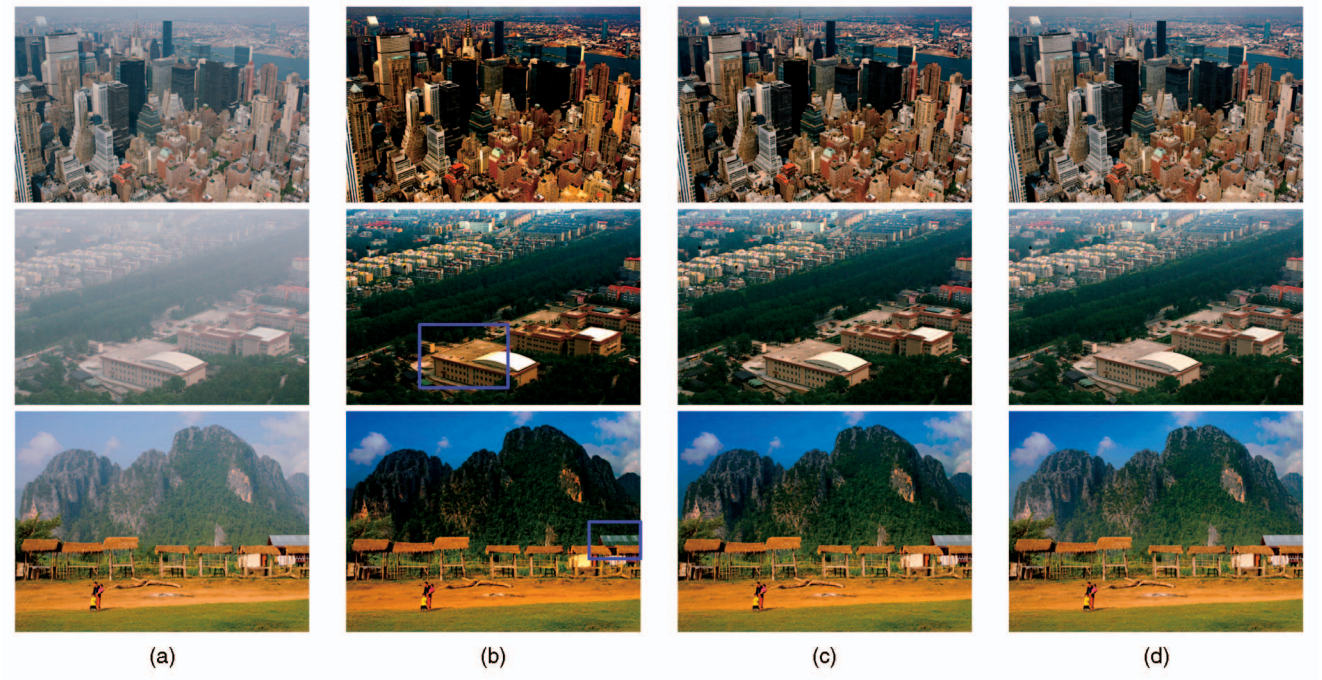

Fig. 11. Recovering images using different patch sizes (after soft matting). (a) Input hazy images. (b) Using 3×3 patches. (c) Using 15×15 patches. (d) Using 30×30 patches.

hazy regions. The estimated depth maps are sharp near depth edges and consistent with the input images.

Our approach also works for gray-scale images if there are enough shadows. Cityscape images usually satisfy this condition. In this case, we omit the operator $\min_c$ in the whole derivation. Fig. 13 shows an example.

In Fig. 14, we compare our approach with Tan's work [11]. The result of this method has oversaturated colors because maximizing the contrast tends to overestimate the haze layer. Our method recovers the structures without sacrificing the fidelity of the colors (e.g., swan). The halo artifacts are also significantly small in our result.

Next, we compare our approach with Fattal's work. In Fig. 8, our result is comparable to Fattal's result. In Fig. 15, we show that our approach outperforms Fattal's for dense haze. His method is based on local statistics and requires sufficient color information and variance. When the haze is dense, the color is faint and the variance is not high enough

Fig. 12. Haze removal results. (a) Input hazy images. (b) Restored haze-free images. (c) Depth maps. The red rectangles in the top row indicate where our method automatically obtains the atmospheric light.

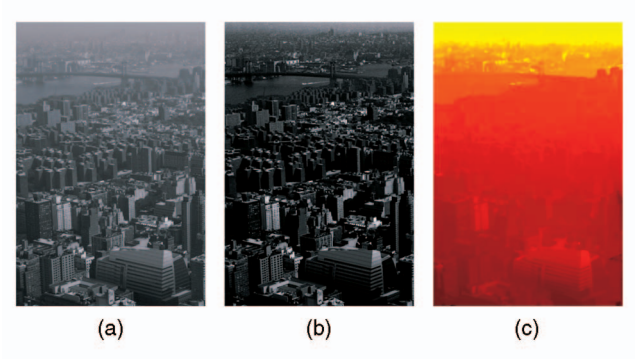

Fig. 13. A gray-scale image example. (a) Input image. (b) Our result. (c) Our recovered depth map.

for his method to reliably estimate the transmission. Figs. 15b and 15c show his results before and after the MRF extrapolation. Since only parts of transmission can be reliably recovered, even after extrapolation some regions are too dark (mountains) and some hazes are not removed (distant part of the cityscape). Our approach gives natural results in these regions (Fig. 15d).

We also compare our method with Kopf et al.'s work [8] in Fig. 16. To remove the haze, they utilize the 3D models and texture maps of the scene. Such additional information may come from Google Earth and satellite images. Our technique can generate comparable results from a single image without any geometric information.

In Fig. 17, we further compare our method with more previous methods. Fig. 17e shows the result of Photoshop's auto curve function. It is equivalent to the dark-object subtraction method [12] applied for each color channel

Fig. 14. Comparison with Tan's work [11]. (a) Input image. (b) Tan's result. (c) Our result.

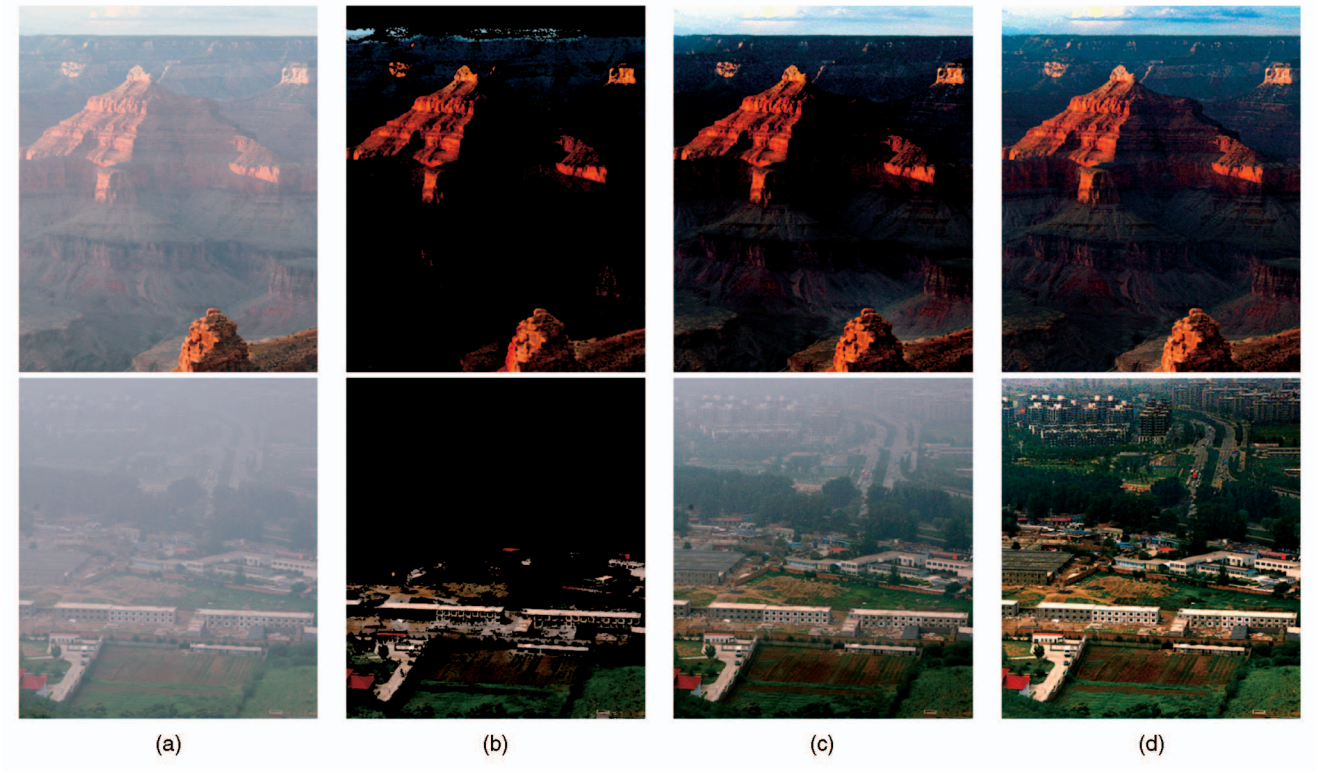

Fig. 15. More comparisons with Fattal's work [10]. (a) Input images. (b) Results before extrapolation, using Fattal's method. The transmission is not estimated in the black regions. (c) Fattal's results after extrapolation. (d) Our results.

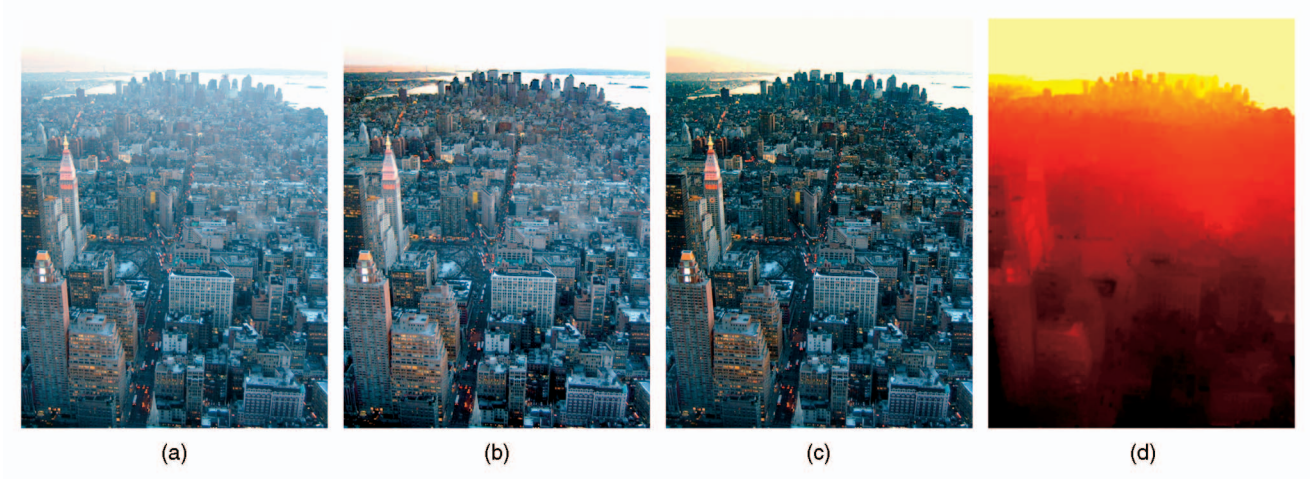

Fig. 16. Comparison with Kopf et al.'s work [8]. (a)-(d) Input image, Kopf et al.'s result, our result, and our recovered depth map, respectively.

independently. As this method is designed to remove spatially invariant haze, when the depth of the image is not constant it can only remove a thin haze layer corresponding to the nearest objects. Figs. 17f and 17g are Photoshop's Unsharp Mask and the classical histogram equalization, which are two widely used contrast enhancement techniques. However, the enhancement should be spatially variant if the depth is not uniform. These techniques cannot determine to what extent the contrast should be enhanced, so the distant objects are not well recovered. Besides, as they are not based on the haze imaging model, they cannot give a depth map.

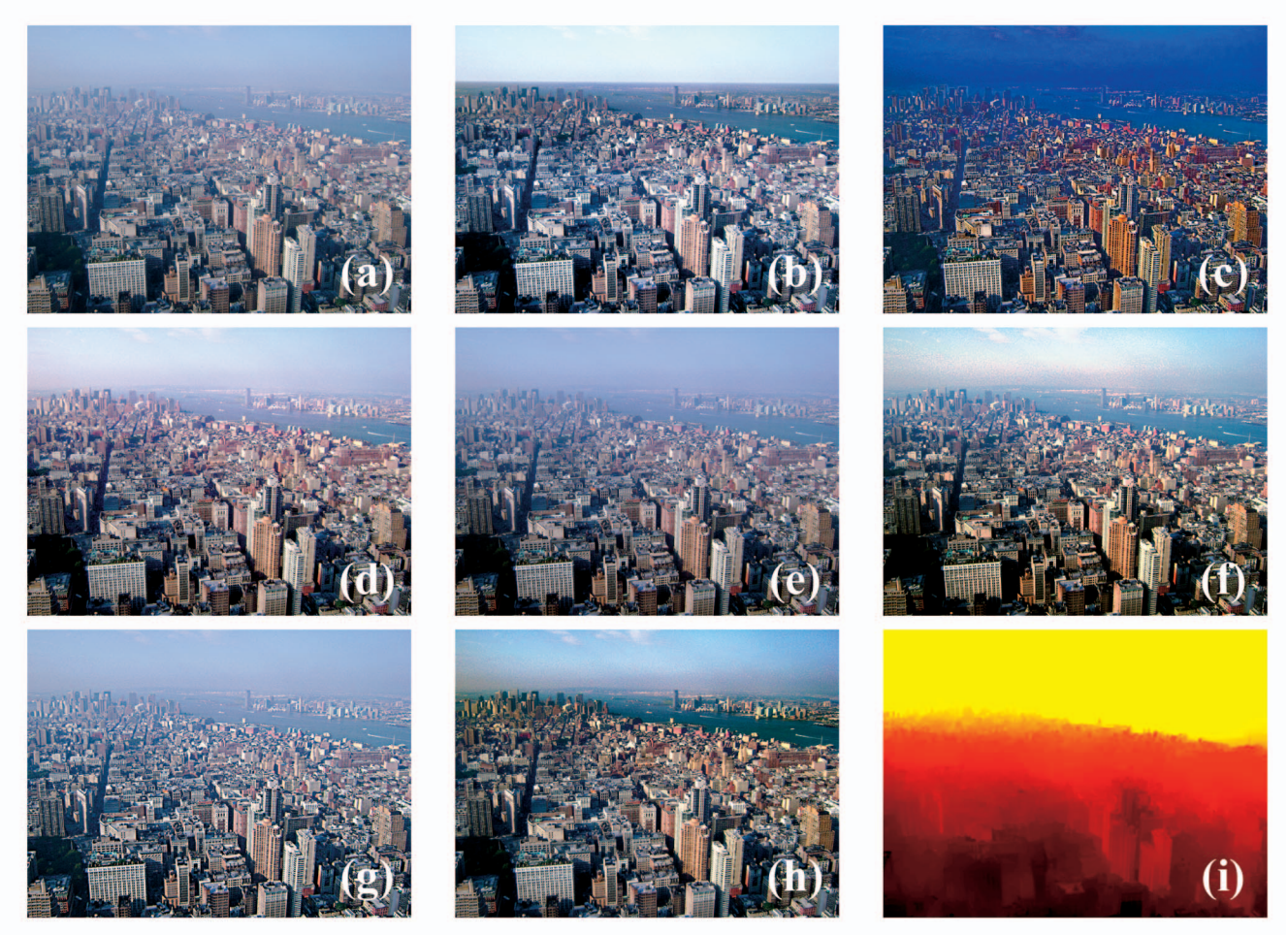

Fig. 17. Comparison with other methods. (a) Input image. (b) Kopf et al.'s result [8]. (c) Tan's result [11]. (d) Fattal's result [10]. (e) Photoshop's auto curve. (f) Photoshop's Unsharp Mask. (g) Histogram equalization. (h) Our result. (i) Our recovered depth map.

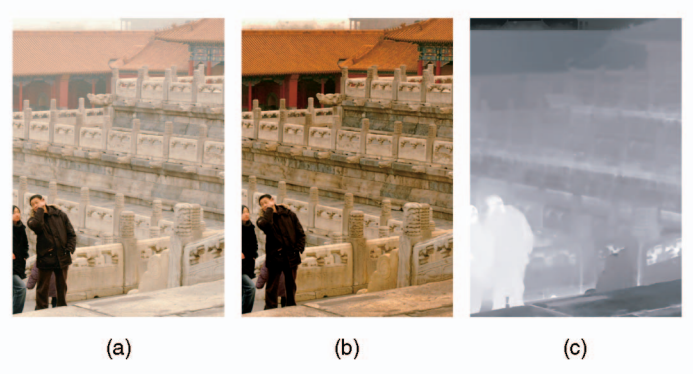

Fig. 18. Failure of the dark channel prior. (a) Input image. (b) Our result. (c) Our transmission map. The transmission of the marble is underestimated.

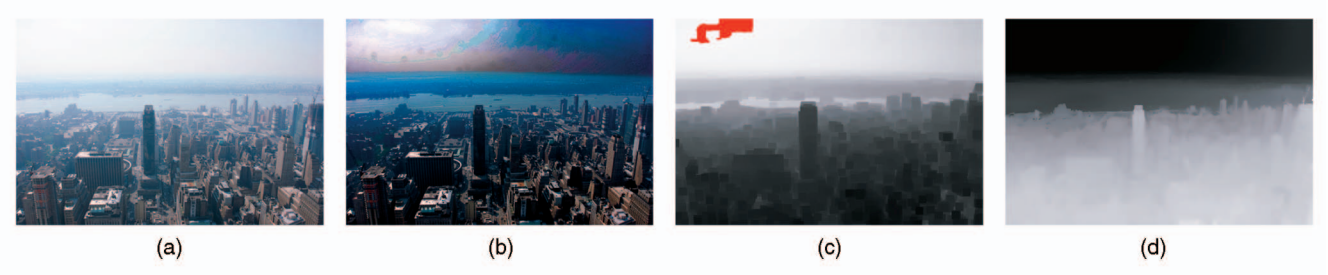

Fig. 19. Failure of the haze imaging model. (a) Input image. (b) Our result. (c) Dark channel. Red pixels indicate where the atmospheric light is estimated. (d) Estimated transmission map.

6 Discussion

In this paper, we have proposed a very simple but powerful prior, called the dark channel prior, for single image haze removal. The dark channel prior is based on the statistics of outdoor haze-free images. Combining the prior with the haze imaging model, single image haze removal becomes simpler and more effective.

Since the dark channel prior is a kind of statistics, it may not work for some particular images. When the scene objects are inherently similar to the atmospheric light and no shadow is cast on them (such as the white marble in Fig. 18), the dark channel prior is invalid. The dark channel of the scene radiance has bright values near such objects. As a result, our method will underestimate the transmission of these objects and overestimate the haze layer.

Moreover, as our method depends on the haze imaging model (1), it may fail when this model is physically invalid. First, the constant-airlight assumption may be unsuitable when the sunlight is very influential. In Fig. 19a, the atmospheric light is bright on the left and dim on the right. Our automatically estimated $\mathbf{A}$ (Fig. 19c) is not the real $\mathbf{A}$ in the other regions, so the recovered sky region on the right is darker than it should be (Fig. 19b). More advanced models [14] can be used to describe this complicated case. Second, the transmission t is wavelength dependent if the particles in the atmosphere are small (i.e., the haze is thin) and the objects are kilometers away [7]. In this situation, the transmission is different among color channels. This is why the objects near the horizon appear bluish (Fig. 19a). As the haze imaging model (1) assumes common transmission for all color channels, our method may fail to recover the true scene radiance of the distant objects and they remain bluish. We leave this problem for future research.

References

[1] H. Koschmieder, "Theorie der Horizontalen Sichtweite," *Beitr. Phys. Freien Atm.*, vol. 12, pp. 171-181, 1924.

[2] S.G. Narasimhan and S.K. Nayar, "Vision and the Atmosphere," *Int'l J. Computer Vision*, vol. 48, pp. 233-254, 2002.

[3] Y.Y. Schechner, S.G. Narasimhan, and S.K. Nayar, "Instant Dehazing of Images Using Polarization," *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, vol. 1, pp. 325-332, 2001.

[4] S. Shwartz, E. Namer, and Y.Y. Schechner, "Blind Haze Separation," *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, vol. 2, pp. 1984-1991, 2006.

[5] S.G. Narasimhan and S.K. Nayar, "Chromatic Framework for Vision in Bad Weather," *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, vol. 1, pp. 598-605, June 2000.

[6] S.K. Nayar and S.G. Narasimhan, "Vision in Bad Weather," *Proc. Seventh IEEE Int'l Conf. Computer Vision*, vol. 2, pp. 820-827, 1999.

[7] S.G. Narasimhan and S.K. Nayar, "Contrast Restoration of Weather Degraded Images," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 25, no. 6 pp. 713-724, June 2003.

[8] 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. Graphics*, vol. 27, no. 5, pp. 116:1-116:10, 2008.

[9] S.G. Narasimhan and S.K. Nayar, "Interactive Deweathering of an Image Using Physical Models," *Proc. IEEE Workshop Color and Photometric Methods in Computer Vision, in Conjunction with IEEE Int'l Conf. Computer Vision*, Oct. 2003.

[10] R. Fattal, "Single Image Dehazing," *Proc. ACM SIGGRAPH '08*, 2008.

[11] R. Tan, "Visibility in Bad Weather from a Single Image," *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, June 2008.

[12] P. Chavez, "An Improved Dark-Object Subtraction Technique for Atmospheric Scattering Correction of Multispectral Data," *Remote Sensing of Environment*, vol. 24, pp. 450-479, 1988.

[13] E.B. Goldstein, *Sensation and Perception.* Cengage Learning 1980.

[14] A.J. Preetham, P. Shirley, and B. Smits, "A Practical Analytic Model for Daylight," *Proc. ACM SIGGRAPH '99*, 1999.

[15] A.L. Gilchrist and A. Jacobsen, "Lightness Constancy through a Veiling Luminance," *J. Experimental Psychology: Human Perception and Performance*, vol. 9, no. 6, pp. 936-944, 1983.

[16] A. Levin, D. Lischinski, and Y. Weiss, "A Closed Form Solution to Natural Image Matting," *Proc. IEEE Conf. Computer Vision and Pattern Recognition,* vol. 1, pp. 61-68, 2006.
[17] I. Omer and M. Werman, "Color Lines: Image Specific Color Representation," *Proc. IEEE Conf. Computer Vision and Pattern Recognition,* vol. 2, pp. 946-953, June 2004.
[18] E. Hsu, T. Mertens, S. Paris, S. Avidan, and F. Durand, "Light Mixture Estimation for Spatially Varying White Balance," *Proc. ACM SIGGRAPH '08,* 2008.
[19] C. Tomasi and R. Manduchi, "Bilateral Filtering for Gray and Color Images," *Proc. Sixth IEEE Int'l Conf. Computer Vision,* p. 839, 1998.
[20] M. van Herk, "A Fast Algorithm for Local Minimum and Maximum Filters on Rectangular and Octagonal Kernels," *Pattern Recognition Letters,* vol. 13, pp. 517-521, 1992.

Kaiming He received the BS degree from the Academic Talent Program, Physics Department, Tsinghua University, Beijing, in 2007. He is currently working toward the PhD degree in the Multimedia Laboratory, Department of Information Engineering, The Chinese University of Hong Kong. His research interests include computer vision and computer graphics. His personal home page is http://personal.ie.cuhk.edu.hk/~hkm007/.

Jian Sun received the BS, MS, and PhD degrees from Xian Jiaotong University in 1997, 2000, and 2003, respectively. He joined Microsoft Research Asia in July 2003. His current two major research interests include interactive computer vision (user interface + vision) and Internet computer vision (large image collection + vision). He is also interested in stereo matching and computational photography.

Xiaoou Tang received the BS degree from the University of Science and Technology of China, Hefei, in 1990, and the MS degree from the University of Rochester, Rochester, New York, in 1991. He received the PhD degree from the Massachusetts Institute of Technology, Cambridge, in 1996. He is a professor in the Department of Information Engineering and Associate Dean (Research) of the Faculty of Engineering of the Chinese University of Hong Kong. He worked as the group manager of the Visual Computing Group at Microsoft Research Asia from 2005 to 2008. His research interests include computer vision, pattern recognition, and video processing. Dr. Tang received the Best Paper Award at the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2009. He was a program chair of the IEEE International Conference on Computer Vision (ICCV) 2009 and is an associate editor of the I*EEE Transactions on Pattern Analysis and Machine Intelligence* (*TPAMI*) and the *International Journal of Computer Vision* (*IJCV*). He is a fellow of the IEEE.