

# Single Image Dehazing Using Haze-Lines

Dana Berman, Tali Treibitz, and Shai Avidan

## Abstract

Haze often limits visibility and reduces contrast in outdoor images. The degradation varies spatially since it depends on the objects' distances from the camera. This dependency is expressed in the transmission coefficients, which control the attenuation. Restoring the scene radiance from a single image is a highly ill-posed problem, and thus requires using an image prior. Contrary to methods that use patch-based image priors, we propose an algorithm based on a non-local prior. The algorithm relies on the assumption that colors of a haze-free image are well approximated by a few hundred distinct colors, which form tight clusters in RGB space. Our key observation is that pixels in a given cluster are often non-local, i.e., spread over the entire image plane and located at different distances from the camera. In the presence of haze these varying distances translate to different transmission coefficients. Therefore, each color cluster in the clear image becomes a line in RGB space, that we term a haze-line. Using these haze-lines, our algorithm recovers the atmospheric light, the distance map and the haze-free image. The algorithm has linear complexity, requires no training, and performs well on a wide variety of images compared to other state-of-the-art methods.

## Index Terms



- 
- *D. Berman and S. Avidan are with the School of of Electrical Engineering, Tel Aviv University, Israel.  
E-mail: danamena@post.tau.ac.il*
  - *T. Treibitz is with the School of Marine Sciences, University of Haifa, Israel.*

# Single Image Dehazing Using Haze-Lines

## 1 INTRODUCTION

OUTDOOR images taken in bad weather conditions, such as fog and haze, often suffer from low contrast and limited visibility. These phenomena greatly affect both computer vision and consumers applications. The poor visibility limits the performance of high level recognition algorithms, as well as low level feature detection and matching. In addition, clear images are often more visually pleasing than washed out scenes. "Defogging" or "dehazing" methods address this problem by trying to invert the medium effects and produce a clear image.

Image degradation is caused by suspended particles in the air that absorb and scatter the light. Such particles include dust (haze) and water droplets (fog), and they attenuate the object radiance along the line of sight [1]. Furthermore, ambient light is reflected into the line of sight, causing the scene to look faded. This veiling layer is also referred to as atmospheric light. Thus, hazy images can be modeled as a per-pixel convex combination of a haze-free image and the global atmospheric light [2], where the convex combination coefficient is called the transmission. The transmission is the fraction of the object radiance that reaches the camera. The farther light has to travel from the object to the camera, the greater it will be affected by the scattering medium and the transmission will be smaller. Therefore, this spatially-variant process cannot be globally corrected.

Our goal is to recover the RGB values of the haze-free image and the transmission for each pixel. This is an ill-posed problem that has an under-determined system of three equations and at least four unknowns per pixel, with inherent ambiguity between haze and object radiance. To handle this ambiguity, early works use additional information such as taking two images with different polarizer states [3] or using multiple images taken in different weather conditions [4]. This line of work impose constraints on the imaging conditions, which are often impossible to achieve, e.g. in dynamic scenes.

More recent works utilize image priors to solve the problem from a single image, for the most part by assuming that the transmission is constant within a small patch, e.g. [5], [6], [7]. Patch-based methods take great care to avoid artifacts by overlapping patches [5], using connections between distant pixels for regularization [6], and using multiple patch sizes [8]. See Sec. 2 for a more detailed survey.

As opposed to recent state-of-the-art methods, the proposed method is global and does not divide the image to patches. We use the observation that colors of a haze-free image can be well approximated by a few hundred distinct colors [9]. This implies that pixels in a hazy image can be modeled by lines in RGB space that pass through the atmospheric light coordinates, i.e., *haze-lines*. This model was proposed in [10], which relied on previous methods for atmospheric light estimation, and was later shown to be effective for atmospheric light estimation as well in [11]. The major advantage in this model is that the pixels that form the

haze-lines are spread across the *entire* image and therefore capture a *global* phenomena that is not limited to small image patches. This prior is more robust and significantly more efficient in run-time than patch-based methods.

This paper presents an end-to-end dehazing method based on the haze-line model, which was found to be one of the top methods in an independent evaluation [12]. Furthermore, we propose a more robust optimization scheme that improves upon previous results. The implementation is publicly available: <https://github.com/danaberman/non-local-dehazing>.

## 2 RELATED WORK

Haze or fog removal from a single image is an ill-posed problem. Therefore, several approaches require additional information in order to improve the visibility. For example, Narasimhan and Nayar [4] use as input at least two images taken under different weather conditions, while Schechner et al. [3] take two images with orthogonal polarizing filter positions. The rest of this survey focuses on single image dehazing techniques for day-time images. The interested reader is referred to an extensive survey that was recently publish by Li et al. [12] for a more comprehensive overview.

Since haze reduces image contrast, Tan [13] maximizes the contrast per patch while maintaining a global coherent image. This often leads to halos at the boundaries between objects at different distances from the camera.

Some methods use a prior on the depth of the image. Tarel and Hutière [14], assume the atmospheric light is smooth except for depth discontinuities. Nishino et al. [15] assume the scene albedo and depth are statistically independent and jointly estimate them using priors on both. The prior on the albedo assumes the distribution of gradients in images of natural scenes exhibits a heavy-tail distribution, and it is approximated as a generalized normal distribution. The depth prior is scene-dependent, and is chosen manually, either as piece-wise constant for urban scenes or smoothly varying for non-urban landscapes.

Several methods assume that transmission and radiance are piece-wise constant, and employ a prior on a patch basis [5], [6], [7], [16]. The dark channel prior [5] assumes that within small image patches there will be at least one pixel with a dark color channel and use this minimal value as an estimate of the present haze. This prior works very well, except in bright areas of the scene where the prior does not hold. In addition, it assumes that the atmospheric light is visible in the image, i.e. that a region with no object in the line of sight exists. This assumption does not always hold. Meng et al. [17] estimate an initial transmission similar to the dark channel prior, but propose a weighted  $L_1$ -norm based contextual regularization. Gibson and Nguyen [16] fit a color ellipsoid per-patch in RGB space. In order to cope with depth discontinuities a median operation is used to estimate the ellipsoid parameters, however it cannot reliably estimate the transmission of fine structures such as leaves.



In [6], [18], color lines are fitted in RGB space per-patch, looking for small patches with a constant transmission. This prior is based on the observation by Omer and Werman [19] that pixels in a haze-free image form color lines in RGB space. These lines pass through the origin and stem from shading variations within objects. As shown by Fattal [6], the color lines in hazy images do not pass through the origin anymore, due to the additive haze component. In small patches that contain pixels in a uniform distance from the camera, these lines are shifted from the origin by the atmospheric light at that distance. By fitting a line to each such patch, the transmission in the patch is estimated using the line's shift from the origin. Unfortunately, not all patches conform to this assumption, both due to uniform lighting and due to depth discontinuities. In practice only a small subset of the patches are used to estimate the transmission and information is propagated to the rest of the image via a Markov random field. In order to calculate the atmospheric light, pairs of patches that satisfy the color-line prior should be found, which is a restricting condition.

The variety of priors has led Tang et al. [8] to examine different patch features in a learning framework. They take into account priors such as the Dark Channel and contrast in their training, and use a Random Forest to estimate the transmission. The training data is synthesized in patches using a constant white atmospheric light and a random uniform transmission, which does not represent true data.

Bahat and Irani [7] use the patch recurrence prior in order to estimate the distance map of the scene. According to this prior, small image patches tend to repeat inside a natural image, both within the same scale as well as across different scales. They utilize the deviations from the ideal patch recurrence, which is diminished in a scattering medium, in order to identify the unknown haze parameters and to reconstruct the haze-free image. The search for pairs of patches that satisfy the recurrence property is prohibitive and the computation time is long.

Ancuti and Ancuti [20] take an image-processing approach to dehazing, and apply local white balance and contrast enhancement procedures. This approach yields visually-pleasing results. Nonetheless, it hasn't been shown to be physically valid.

Choi et al. [21] develop a perceptual fog density prediction model based on natural image statistics. Using this model, they propose an image dehazing method, which works well except for dark areas which are over-corrected.

Recently Convolutional Neural Networks (CNNs) have been proposed to solve the single image dehazing problem, using synthetic data due to the lack of ground truth dataset for this problem. Cai et al. [22] suggest DehazeNet, a CNN that extract features from the full resolution image using different filter sizes and applies a non-linear function to predict the transmission map. They train DehazeNet on  $16 \times 16$  patches of natural images using random transmission, since the transmission is independent of the image content. Ren et al. [23] present a Multi-Scale CNN that has two steps: the first network uses relatively large filter sizes and predicts a coarse transmission map, and this map is concatenated as input to a second CNN that predicts the fine details of the transmission, using smaller filter sizes. Both CNNs alternate pooling and upsampling in order to maintain the original

resolution. These networks are trained based on synthetic hazy images, generated using the NYU RGB-D dataset [24]. Both of these methods require that the atmospheric light is visible in the image, and estimate it based on the pixels with the smallest transmission. The clear image is restored based on the estimated transmission map and atmospheric light value. In contrast, Li et al. [25] posit that better reconstruction can be achieved by an All-in-One Dhezing Network (AOD-Net) that directly outputs a dehazed clear image. This architecture fuses varied size filters to form multi-scale features, and predict a term which encompasses both the atmospheric light and the transmission. They too train on synthetic hazy images, generated using the NYU RGB-D dataset.

Here, we suggest the haze-lines model, which is based on the observation that colors tend to repeat in natural scenes and are often located at different distances from the camera. Due to the different distances, the paths to the camera are different and so is the transmission. The transmission is estimated by analyzing colors in RGB space and specifically the distance to the atmospheric light. This model has been proposed in [10] and later shown to be effective for atmospheric light estimation [11] as well. Unfortunately, the refinement step used in [10] defines an optimization function and uses a weighted least squares solver, which does not always adhere to a lower bound constraint. As a result, some of the restored images have visual artifacts, in addition to the inaccurate transmission estimation. Here, we propose a different optimization and show it has superior performance.

While the haze-lines might seem similar to [6], they are inherently different. In [6], lines are defined by the pixels of small patches in the image plane assuming constant transmission, with often small intensity differences caused by shading. This is a *local* phenomena that does not always hold and indeed, in [6] care is taken to ensure only patches where the assumption holds are considered. We, on the other hand, look at lines that are formed by individual pixels that are scattered over the entire image. These pixels usually have large intensity differences that are caused by changes in transmission rather than local shading effects.

### 3 THE HAZE-LINES MODEL

We first present the physical image formation model in a scattering medium, and then describe the haze-lines model which is the basis for the proposed dehazing method.

#### 3.1 Image Formation Model

The common image formation model [2] in haze and fog is a linear combination of the attenuated scene radiance and the ambient light that is scattered into the line of sight:

$$\mathbf{I}(\mathbf{x}) = t(\mathbf{x}) \cdot \mathbf{J}(\mathbf{x}) + [1 - t(\mathbf{x})] \cdot \mathbf{A}, \quad (1)$$

where  $\mathbf{x}$  is the pixel coordinates,  $\mathbf{I}$  is the observed hazy image (the linear sensor response), and  $\mathbf{J}$  is the true radiance of the scene point imaged at  $\mathbf{x}$ . The atmospheric light  $\mathbf{A}$  is a single color representing the ambient light in image areas where  $t = 0$ . Bold symbols represent vectors, i.e.  $\mathbf{I}, \mathbf{J}, \mathbf{A}$  have three components (three color channels  $R, G, B$ ) while  $\mathbf{x}$  has two components (coordinates in the image plane).

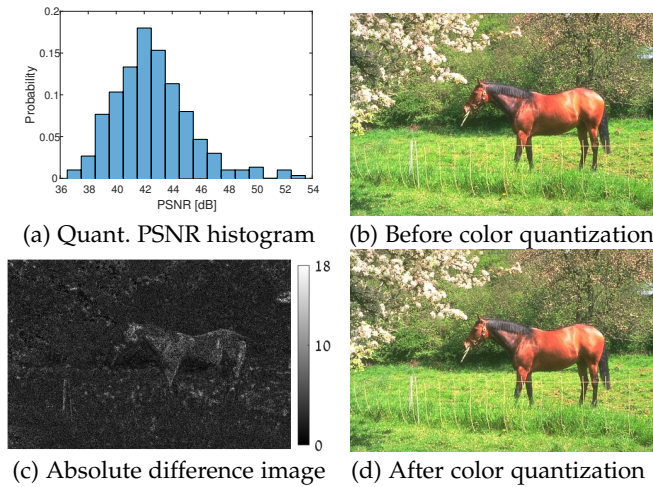


Fig. 1. **Prior validation:** (a) A PSNR histogram of the quantization errors on the Berkeley Segmentation Dataset (BSDS300): The RGB values of each image were clustered using K-means to 500 clusters and replaced by the cluster center. The histogram shows the PSNRs measured on the entire dataset. (b,d) The image that had the worst PSNR, 36.64dB, before (b) and after (d) color quantization. (c) Absolute difference image to color-quantized version (the contrast was stretched for display, note that the maximal difference was 18 out of 256).

This model applies only to day-time images and assumes a single atmospheric light. During night time, or in thick fog, the atmospheric light often varies spatially and a more elaborate image formation model should be used. [26], [27].

Note that Eq. (1) holds for the scene radiance, whereas most camera-processed images (commonly in JPEG format) do not have a linear relationship with the radiance [28]. Therefore the dehazing process should be applied to a radiometrically corrected [29] image to obtain the best results.

The scene transmission  $t(x)$  acts as a matting coefficient between the scene radiance  $J$  and the atmospheric light  $A$ , and is distance-dependent:

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

where  $\beta$  is the attenuation coefficient of the atmosphere and  $d(x)$  is the distance of the scene at pixel  $x$ . Generally,  $\beta$  is wavelength dependent and therefore  $t$  is different per color channel [3], [4]. This dependency has been assumed negligible in previous single image dehazing methods to reduce the number of unknowns and we follow this assumption.

Thus, per-pixel  $x$ , Eq. (1) has three observations  $I(x)$  and four unknowns:  $J(x)$  and  $t(x)$ , resulting in an under-determined estimation problem.

### 3.2 The Prior

Our method is based on the observation that the number of distinct colors in an image is orders of magnitude smaller than the number of pixels [9]. This assumption has been used extensively in the past and is used for saving and transmitting color images using indexed colormaps. We validate and quantify it on the Berkeley Segmentation Dataset (BSDS300). This is a diverse dataset of clear outdoor natural images and thus represents the type of scenes that might be degraded by haze. We clustered the RGB pixel values of each image using K-means to a maximum of 500 clusters, and replaced every pixel in the image with its respective

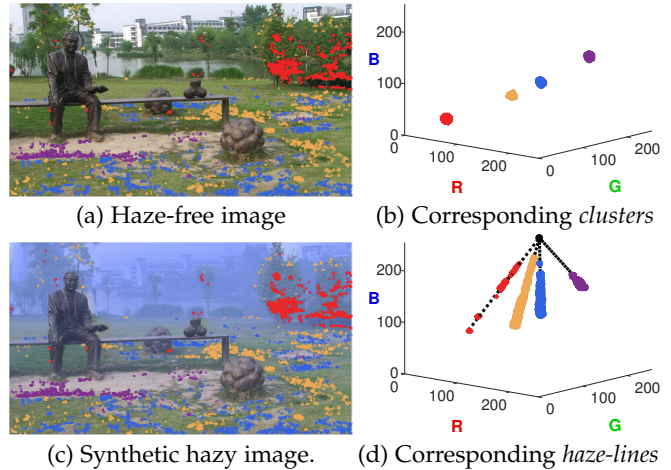
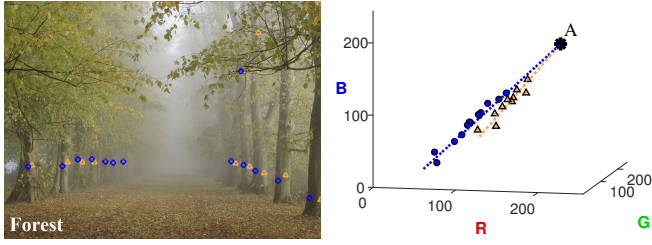


Fig. 2. **Haze-lines demonstration (synthetic image).** (a) Pixels of a haze free color image are clustered using K-means. Pixels belonging to four of the clusters are marked. Note that the pixels are non-local and are spread all over the image plane. (b) The four color clusters are depicted in RGB space. Colors of the clusters correspond to the highlighted pixels in (a). (c) Synthetic haze is added to (a). The same clustered pixels are marked, but their observed colors are affected by different amounts of haze. (d) The hazy pixels depicted in RGB color space. They are distributed along lines, termed *haze-lines*, passing through the atmospheric light, marked in black.

cluster center. The result is an image with 500 different RGB values at most (two orders of magnitude smaller than image size). The PSNR values of the images generated with the reduced color set, compared to the original ones, were high and ranged from 36.6dB to 52.6dB. A histogram of the obtained PSNR values is shown in Fig. 1, as well as the image that had the *worst* PSNR, before and after color quantization. As seen by the PSNR distribution, the color-quantized images are nearly indistinguishable from the original images.

The observation regarding a small number of distinct colors holds for haze-free images. In the presence of haze, object points that belong to the same color cluster end up with different acquired colors, since they are located in disparate image areas and thus have different distances from the camera. As a result, pixels that are clustered together in a haze-free image form a line in RGB space in a hazy image. Based on Eq. (1), the two end points of the line are the original color  $J$  and the atmospheric light  $A$ . These are the *haze-lines*.

This prior is demonstrated in Fig. 2. A haze-free image is clustered using K-means to 500 clusters. The pixels belonging to four of these clusters are marked by different color markers in Fig. 2a and their RGB coordinates are plotted in Fig. 2b, demonstrating tight clusters. Note that the clusters include pixels distributed over the entire image that come from objects with different distances from the camera. A synthetic hazy image was generated from the clear image (Fig. 2c) by [6]. The same pixels as in Fig. 2a are marked. However, now, colors of pixels that belonged to the same color cluster are no longer similar. This is depicted in RGB space in Fig. 2d, where the color coordinates of these pixels are distributed along a haze-line spanned by the original color and the atmospheric light. In each haze-line, each pixel is a convex combination of the original color



(a) Input hazy image: Forest. Note pixels marked in blue circles and orange triangles.

(b) Each group of marked pixels forms a **haze-line** in RGB color space.

Fig. 3. **Haze-lines demonstration (natural image).** (a) An input hazy image, with two types of pixels marked by a blue circle and an orange triangle. Those pixels show objects with similar radiance (shaded and lit tree trunks, respectively) at different distances from the camera. (b) The color coordinates of the pixels depicted in (a) are shown in RGB color space, with a corresponding color coding. They are distributed over two different **haze-lines**, as identified by our method. The lines pass through the atmospheric light, marked in black. The other end of the line is the haze-free color of these pixels. As expected, the pixels along the haze-line marked in blue, which correspond to the shaded side of the trees, have a darker haze-free color than the haze-line marked in orange.

and the atmospheric light. The pixels marked by purple circles (originating from the sand patch) are located in similar distances, so their distribution along the haze-line is rather tight. However, the pixels marked by orange triangles (grassy areas) are found at different locations in the real world, so they are distributed along the haze-line.

Fig. 3 demonstrates the haze-lines prior on a hazy outdoor image. Two different haze-lines are shown both in the image plane and in RGB space. The pixels marked by a blue circle correspond to shaded tree trunks and branches, and are likely to have similar radiance  $J$ . Similarly, the pixels marked by an orange triangle show lit branches and tree trunks. However, their observed intensity  $I$  is quite different. Fig. 3b shows that these pixels form two haze-lines in RGB space that pass through the atmospheric light.

## 4 DEHAZING USING HAZE-LINES

### 4.1 Atmospheric Light Estimation

We wish to model pixel intensities by an intersection point (i.e., the atmospheric light) and a collection of lines (i.e., the haze-lines) emanating from it. An atmospheric light in the correct RGB location will fit the model to the data better than an atmospheric light in a wrong location.

We use a Hough transform to estimate the atmospheric light, since it is an efficient technique to detect unknown parameters of a model given noisy data. In this case, the Hough voting procedure is carried out in a parameter space consisting of candidate atmospheric light values in RGB space. In particular, we uniformly sample a fixed set of line angles  $\{\theta_k, \phi_k\}_{k=1}^K$ . Given this set, we consider a discrete set of possible atmospheric light values. The distance between a pixel  $I(x)$  and the line defined by the atmospheric light  $A$  and a pair of angles  $(\theta, \phi)$  is:

$$d(I(x), (A, \phi, \theta)) = \|(A - I(x)) \times (\cos(\theta), \sin(\phi))\|. \quad (3)$$

A pixel votes for a candidate  $A$  only if the distance to one of the lines is smaller than a threshold  $\tau$ . This threshold is adaptive and depends on the distance between  $A$  and

### Algorithm 1: Atmospheric Light Estimation

**Input:** hazy image  $I(x)$

**Output:**  $\hat{A}$

- 1: Cluster the pixels' colors and generate an indexed image  $\hat{I}(x)$  whose values are  $n \in \{1, \dots, N\}$ , a colormap  $\{I_n\}_{n=1}^N$ , and cluster sizes  $\{w_n\}_{n=1}^N$
- 2: **for all** pairs of color channels  $(c_1, c_2) \in \{RG, GB, RB\}$  **do**
- 3:   Initialize  $\text{accum}_{c_1, c_2}$  to zero
- 4:   **for**  $A = (m \cdot \Delta A, l \cdot \Delta A)$ ,  $m, l \in \{0, \dots, M\}$  **do**
- 5:     **for**  $\theta_k = \frac{\pi}{K}$ ,  $k \in \{1, \dots, K\}$  **do**
- 6:       **for**  $n \in \{1, \dots, N\}$  **do**
- 7:           $d = \|(A - I_n(c_1, c_2)) \times (\cos(\theta_k), \sin(\theta_k))\|$
- 8:          **if**  $(d < \tau) \wedge (m \cdot \Delta A > I_n(c_1)) \wedge (l \cdot \Delta A > I_n(c_2))$  **then**
- 9:              $\text{accum}_{c_1, c_2}(k, m, l) += w_n \cdot f(\|A - I_n\|)$
- 10:          **end if**
- 11:       **end for**
- 12:     **end for**
- 13:   **end for**
- 14: **end for**
- 15:  $\hat{A} = \arg \max \{\text{accum}_{RG} \otimes \text{accum}_{GB} \otimes \text{accum}_{RB}\}$ , where  $\otimes$  is an outer product
- 16: **return**  $\hat{A}$

$I(x)$  to allow for small intensity variations. I.e., instead of working with cylinders (lines with a fixed threshold) we work with cones (lines with a variable threshold). Formally:

$$\tau = \tau_0 \cdot \left(1 + \frac{\|I(x) - A\|}{\sqrt{3}}\right). \quad (4)$$

In addition, we allow a pixel to vote only for an atmospheric light that is brighter than itself. This is due to the fact that bright objects are quite rare, as shown empirically to justify the dark channel prior [5], and usually do not contain information about the haze (e.g., a bright building close to the camera).

Our method can be described as finding the best representation of the pixels' values of a hazy image with atmospheric light  $A$  and fixed line directions  $\{\theta_k, \phi_k\}_{k=1}^K$ . This can be formulated as follows:

$$\arg \max_A \sum_x \sum_k \mathbb{1}[d(I(x), (A, \phi_k, \theta_k)) < \tau] \cdot \mathbb{1}[A > I(x)], \quad (5)$$

where  $\mathbb{1}[\cdot]$  is an indicator function that equals 1 if true and 0 otherwise. The term  $\mathbb{1}[A > I(x)]$  equals 1 if all elements of  $A$  are greater than the corresponding elements of  $I(x)$ .

A huge value of  $A \gg 1$  might be chosen as the solution, since it maximizes Eq. (5) with all of the pixels in the same large cone. To prevent this, we give a larger weight to values of  $A$  that are close to the pixels' values. Formally, we optimize:

$$\arg \max_A \sum_x \sum_k f(\|I(x) - A\|) \cdot \mathbb{1}[d(I(x), (A, \phi_k, \theta_k)) < \tau] \cdot \mathbb{1}[A > I(x)], \quad (6)$$

where  $f(y) = 1 + 4 \cdot e^{-y}$  is a fast decaying weight that gives preference to values of  $A$  in the vicinity of the pixels' distribution.



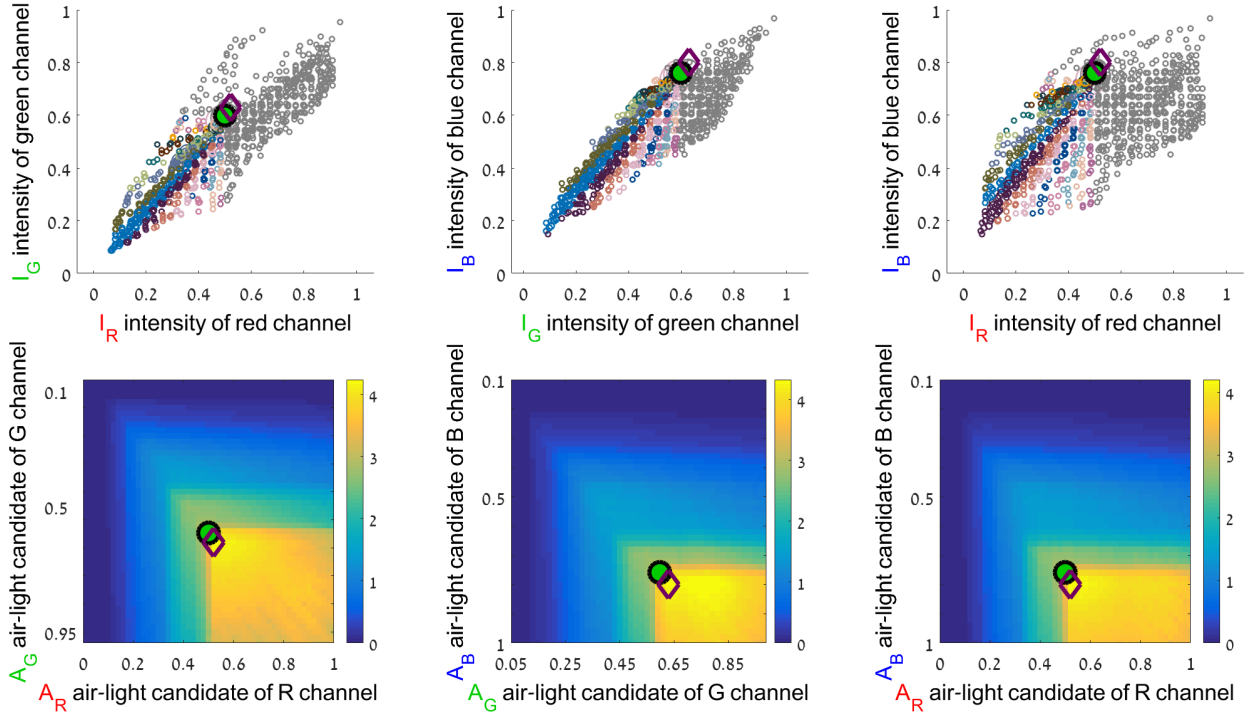


Fig. 4. **Hough votes.** [Top] The color clusters  $\{I_n\}_{n=1}^N$  are projected onto 3 different 2D planes. Each cluster  $n$  is marked by a circle with a size proportional to  $w_n$ . The ground-truth (GT) atmospheric light (extracted from a visible patch in the image) is marked by a green circle while our estimate is marked by a purple diamond. Each colored cluster votes for the GT value, where different colors indicate different haze-lines. The gray colored clusters do not vote for the GT since the following holds:  $\mathbb{1}[A > I_n] = 0$ . [Bottom] The three voting arrays,  $\text{accum}_{c_1, c_2}$ ,  $(c_1, c_2) \in RG, GB, RB$ . Best viewed in color.

#### 4.1.1 Optimizing Computational Efficiency

The proposed scheme, which includes collecting votes from all pixels for all angles and atmospheric light candidates in the 3D RGB space, is computationally expensive. Therefore, we propose the following approximations, which significantly accelerate the computation while maintaining accuracy. The first, clustering the colors in the image and using the cluster centers instead of all the pixels. The second, performing the voting scheme in two dimensions. The voting is repeated three times, with only two of the  $(R, G, B)$  color channels being used each time.

**Color clusters.** Before we start the Hough voting we quantize the image into  $N$  clusters. We do this by converting the RGB image into an indexed image with a unique color palette of length  $N$ . This gives us a set of  $N$  typical color values,  $\{I_n\}_{n=1}^N$ , where  $N$  is much smaller than the number of pixels in the image. In addition, we have  $\{w_n\}_{n=1}^N$ , the number of pixels in the image belonging to each cluster. During the Hough voting procedure, each representative color value  $I_n$  votes based on its distance to the candidate atmospheric light, and the vote has a relative strength  $w_n$ . Therefore, the final optimization function is:

$$\arg \max_A \sum_n \sum_k w_n \cdot f(\|I_n - A\|) \cdot \mathbb{1}[d(I_n, (A, \phi_k, \theta_k)) < \tau] \cdot \mathbb{1}[A > I_n] \quad (7)$$

**Two-dimensional vote.** Calculating the full 3D accumulator for all possible atmospheric light values is computationally expensive. Therefore, we perform this calculation in a lower dimension. The accumulator can be seen as the joint probability distribution of the atmospheric light in all three

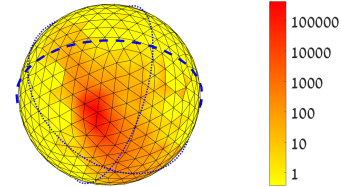


Fig. 5. **Atmospheric light centered spherical representation.** The sphere was sampled uniformly using 500 points. The color at each point  $[\phi, \theta]$  indicates the number of pixels  $x$  with these angles when writing  $I_A(x)$  in spherical coordinates (image size  $768 \times 1024$ ).

color channels, where the final selected value is the one with the maximal probability. By performing the accumulation two color channels at a time, we calculate three marginal probabilities, where each time the summation is performed on a different color channel. Finally, we look for a candidate atmospheric light that will maximize the 3D volume created by the outer product of the marginal accumulators.

The proposed method is summarized in Alg. 1. Denoting the number of pixels as  $P$  and assuming that the number of iterations for color clustering is bounded, the atmospheric light estimation run-time is  $O(N \cdot P + N \cdot M^2 \cdot K + M^3)$ , which is linear in the number of pixels.

## 4.2 Transmission Estimation

### 4.2.1 Clustering pixels to Haze-Lines

Once  $A$  is estimated, we can define  $I_A$  as:

$$I_A(x) = I(x) - A, \quad (8)$$

i.e., the 3D RGB coordinate system is translated such that the atmospheric light is at the origin. Following Eq. (1),

$$I_A(x) = t(x) \cdot [J(x) - A] \quad (9)$$

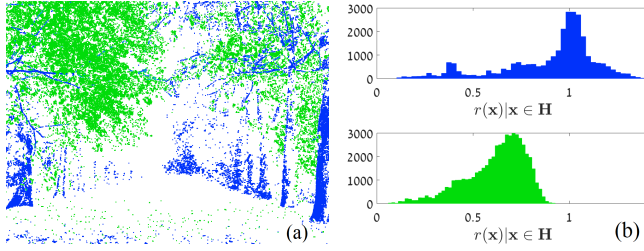


Fig. 6. **Distance distribution per haze-line:** (a) Pixels belonging to two different haze-lines are depicted in green and blue, respectively. (b) A histogram of  $r(\mathbf{x})$  within each cluster. The horizontal axis is limited to the range  $[0, \|\mathbf{A}\|]$ , as no pixel can have a radius outside that range in this particular image.

We express  $\mathbf{I}_A(\mathbf{x})$  in spherical coordinates:

$$\mathbf{I}_A(\mathbf{x}) = [r(\mathbf{x}), \theta(\mathbf{x}), \phi(\mathbf{x})] \quad (10)$$

Here  $r$  is the distance to the origin (i.e.,  $\|\mathbf{I} - \mathbf{A}\|$ ), and  $\theta$  and  $\phi$  are the longitude and latitude, respectively.

Let us look at Eq. (9). For given values of  $\mathbf{J}$  and  $\mathbf{A}$ , scene points at different distances from the camera differ only in the value of  $t$ . In the spherical coordinate system we defined, changes in  $t$  affect only  $r(\mathbf{x})$  without changing either  $\phi(\mathbf{x})$  or  $\theta(\mathbf{x})$ . In other words, pixels  $\mathbf{x}$  and  $\mathbf{y}$  have similar RGB values in the underlying haze-free image if their  $[\phi, \theta]$  are similar:

$$\mathbf{J}(\mathbf{x}) \approx \mathbf{J}(\mathbf{y}) \Rightarrow \{\phi(\mathbf{x}) \approx \phi(\mathbf{y}), \theta(\mathbf{x}) \approx \theta(\mathbf{y})\}, \forall t. \quad (11)$$

Therefore, pixels belong to the same haze-line if their  $[\phi(\mathbf{x}), \theta(\mathbf{x})]$  values are similar. Each point on the sphere in Fig. 5 represents a haze-line, in which all the pixels have approximately the same angles  $[\phi(\mathbf{x}), \theta(\mathbf{x})]$ . The pixels on each haze-line have similar values in the non-hazy image  $\mathbf{J}$  with high probability.

Note that there is inherent ambiguity between color and haze when colors are collinear with the atmospheric light:

$$\mathbf{J}_1 - \mathbf{A} = \alpha(\mathbf{J}_2 - \mathbf{A}) \Rightarrow \mathbf{J}_1 = (1 - \alpha)\mathbf{A} + \alpha\mathbf{J}_2, \quad (12)$$

where  $\alpha$  is a scale factor. In this case all single image dehazing methods will correct  $\mathbf{J}_1$  and  $\mathbf{J}_2$  to the same color. This is the only case in our method when two color clusters will be mapped to the same haze-line.

In order to determine which pixels are on the same haze-line, pixels should be grouped according to their angles  $[\phi, \theta]$ . A 2-D histogram binning of  $\theta$  and  $\phi$  with uniform edges in the range  $[0, 2\pi] \times [0, \pi]$  will not generate a uniform sampling of a sphere. Instead, the samples will be denser near the poles [30], since the distance on the sphere is relative to  $\sin(\theta)$ . Therefore, we sample the unit sphere uniformly using  $Q$  points, as shown in Fig. 5, where each vertex represents one of the  $Q$  cluster centers. Each vertex corresponds to a haze-line. For clarity of display, the number of samples in Fig. 5 is smaller than the actual  $Q$  we use. We group the pixels based on their  $[\phi(\mathbf{x}), \theta(\mathbf{x})]$  values, according to the closest sample point on the surface. This can be implemented efficiently by building a KD-Tree from the pre-defined tessellation and querying the tree for each pixel. This is much faster than running a clustering algorithm such as k-means.

The colors of the pixels are now represented in a spherical coordinate system around the atmospheric light. Fig. 5

shows the histogram of the Forest image (Fig. 3a) projected onto a sphere. The color represents the number of pixels pointing at each direction. The equator ( $\phi = 0$ ) is marked by a bold dashed blue line, while the longitudes  $\theta = 0, \frac{\pi}{2}$  are marked by dotted blue lines. The color-mapping is logarithmic for illustration purposes. The histogram indicates that the pixels are highly concentrated in terms of their longitude and latitude.

Based on the analysis described in section 3.2, several hundreds of haze-lines represent an image with a good approximation. Fig. 6a depicts the layout of two different haze-lines in the image plane for the Forest image.

#### 4.2.2 Estimating Initial Transmission

For a given haze-line defined by  $\mathbf{J}$  and  $\mathbf{A}$ ,  $r(\mathbf{x})$  depends on object distance:

$$r(\mathbf{x}) = t(\mathbf{x})\|\mathbf{J}(\mathbf{x}) - \mathbf{A}\|, \quad 0 \leq t(\mathbf{x}) \leq 1. \quad (13)$$

Thus,  $t = 1$  corresponds to the largest radial coordinate:

$$r_{\max} \stackrel{\text{def}}{=} \|\mathbf{J} - \mathbf{A}\|. \quad (14)$$

Combining Eqs. (13,14) results in an expression for the transmission based on radii in the haze-line:

$$t(\mathbf{x}) = r(\mathbf{x})/r_{\max}. \quad (15)$$

Now, the question is how to find an estimate  $\hat{r}_{\max}$  for the maximal radius? If a haze-line  $H$  contains a haze-free pixel, then  $\hat{r}_{\max}$  is the maximal radius of that haze-line:

$$\hat{r}_{\max}(\mathbf{x}) = \max_{\mathbf{x} \in H} \{r(\mathbf{x})\}, \quad (16)$$

where the estimation is done per haze-line  $H$ . Fig. 6b displays the radii histograms of the two clusters shown in Fig. 6a. We assume that the farthest pixel from the atmospheric light is haze free, and that such a pixel exists for every haze-line. This assumption does not hold for all of the haze-lines in an image, however the regularization step partially compensates for it. Combining Eqs. (15,16) results in a per-pixel estimation of the transmission:

$$\tilde{t}(\mathbf{x}) = r(\mathbf{x})/\hat{r}_{\max}(\mathbf{x}). \quad (17)$$

#### 4.2.3 Regularization

Since the radiance  $\mathbf{J}$  should always be positive (i.e.,  $\mathbf{J} \geq 0$ ), Eq. (1) yields a lower bound on the transmission:

$$t_{LB}(\mathbf{x}) = 1 - \min_{c \in \{R, G, B\}} \{I_c(\mathbf{x})/A_c\}. \quad (18)$$

In [5], the transmission estimate is based on an eroded version of  $t_{LB}$ . We impose this bound on the estimated transmission, per-pixel:

$$\tilde{t}_{LB}(\mathbf{x}) = \max\{\tilde{t}(\mathbf{x}), t_{LB}(\mathbf{x})\}. \quad (19)$$

The estimation in Eq. (17) is performed per-pixel, without imposing spatial coherency. This estimation can be inaccurate if a small amount of pixels were mapped to a particular haze-line, or in very hazy areas, where  $r(\mathbf{x})$  is very small and noise can affect the angles significantly. The transmission map should be smooth, except for depth discontinuities [6], [13], [14], [15]. We seek a transmission map  $\hat{t}(\mathbf{x})$  that is similar to  $\tilde{t}_{LB}(\mathbf{x})$  and is smooth when the

### Algorithm 2: Haze Removal

**Input:**  $I(x)$ ,  $A$   
**Output:**  $\hat{J}(x)$ ,  $\hat{t}(x)$

- 1: Estimate  $A$  using Alg. 1
- 2:  $I_A(x) = I(x) - A$
- 3: Convert  $I_A$  to spherical coordinates to obtain  $[r(x), \phi(x), \theta(x)]$
- 4: Cluster the pixels according to  $[\phi(x), \theta(x)]$ .  
Each cluster  $H$  is a *haze-line*.
- 5: **for** each cluster  $H$  **do**
- 6:   Estimate max. radius:  $\hat{r}_{\max}(x) = \max_{x \in H} \{r(x)\}$
- 7: **end for**
- 8: **for** all pixels  $x$  **do**
- 9:   Estimate transmission:  $\tilde{t}(x) = \frac{r(x)}{\hat{r}_{\max}}$
- 10: **end for**
- 11: Perform regularization by calculating  $\hat{t}(x)$  that minimizes Eq. (20)
- 12: Calculate the dehazed image using Eq. (21)

input image is smooth. Mathematically, we minimize the following function w.r.t.  $\hat{t}(x)$ :

$$\sum_x f(\sigma(x)) \cdot [\hat{t}(x) - \tilde{t}_{LB}(x)]^2 + \lambda \sum_{x \in N_x} \frac{[\hat{t}(x) - \hat{t}(y)]^2}{\|\tilde{I}(x) - \tilde{I}(y)\|^2}, \quad (20)$$

where  $\lambda$  is a parameter that controls trade-off between the data and the smoothness terms,  $N_x$  denotes the four nearest neighbors of  $x$  in the image plane,  $\tilde{I}$  is the image without radiometric correction, and  $\sigma(x)$  is the standard deviation of  $r(x)$ , which is calculated per haze-line. We use Weighted Least Squares (WLS) [31] to solve Eq. (20).

$\sigma(x)$  plays a significant role since it allows us to apply our estimate only to pixels where the assumptions hold. When the radii distribution in a given haze-line is small, our haze-line assumption does not hold since we do not observe pixels with different amounts of haze. Mathematically, when the standard deviation of the radii is low, the estimation of the maximal radius (Eq. 16) is less reliable. The standard deviation is merely an indicator to the validity of our assumptions, and beyond a certain value there is no additional benefit in having a large radii distribution. Therefore,  $f(\sigma(x))$  is a soft step function of  $\sigma(x)$ . Our experiments show that the algorithm is insensitive to the exact parameters of the step function, and we use the following:  $f(\sigma(x)) = \min\{1, 3 \cdot \max\{0.001, \frac{\sigma(x)}{\max_x \sigma(x)} - 0.1\}\}$ .

### 4.3 Haze Removal

Once  $\hat{t}(x)$  is calculated as the minimum of Eq. (20), the dehazed image is calculated using Eq. (1):

$$\hat{J}(x) = \{I(x) - [1 - \hat{t}(x)] A\} / \hat{t}(x) \quad (21)$$

The method is summarized in Alg. 2 and demonstrated in Fig. 7. Fig. 7a shows the input hazy image. The final, dehazed image is shown in Fig. 7b. Fig. 7c shows the distance in RGB space of every pixel in the hazy image to the atmospheric light. Note that this distance decreases as haze increases. Fig. 7d shows the maximum radii  $\hat{r}_{\max}(x)$  per haze-line. Observe that Fig. 7d is much brighter than Fig. 7c. Since larger values are represented by brighter

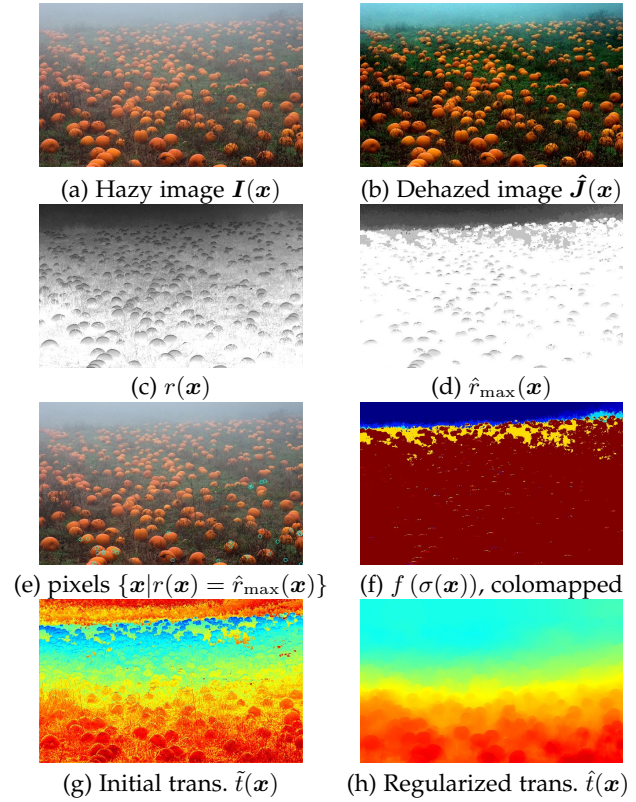


Fig. 7. **Intermediate and final results:** (a) An input hazy image; (b) The output image; (c) The distance  $r(x)$  of every pixel of the hazy image to the atmospheric light; (d) the estimated radii  $\hat{r}_{\max}(x)$  calculated according to Eq. (16); (e) The input image is shown, with the pixels  $x$  for which  $r(x) = \hat{r}_{\max}(x)$  marked by cyan circles; (f) The data term confidence in Eq. (20) colormapped (warm colors show the larger values); (g) The estimated transmission map  $\tilde{t}(x)$  before the regularization; (h) The final transmission map  $\hat{t}(x)$  after regularization. (g) and (h) are colormapped.

colors, this indicates that the distance to the atmospheric light is increased. The pixels with the maximum radius in their haze-line are marked on the hazy image in Fig. 7e. Note that these pixels are mostly at the foreground, where indeed there is a minimal amount of haze. We filtered out pixels that had a maximum radius in the haze line, yet had a  $f(\sigma) < 0.5$ , since the model assumptions do not hold for these haze lines. The aforementioned pixels are found in the sky, since the distance to the atmospheric light in RGB space is very short. Therefore, clustering them according to their angles is not reliable due to noise. In the regularization step this fact is taken into consideration through the data-term weight, which is shown in Fig. 7f (warm colors depict high values). The ratio of Figs. 7c and 7d yields the initial transmission  $\tilde{t}(x)$  that is shown in Fig. 7g. The transmission map after regularization is shown in Fig. 7h. While  $\tilde{t}(x)$  contains fine details even in grass areas that are at the same distance from the camera,  $\hat{t}(x)$  does not exhibit this behavior. This indicates the regularization is necessary.

## 5 EXPERIMENTS

### 5.1 Experimental Setup

We use a fixed set of parameters in all of our experiments. For atmospheric light estimation: the number of different color clusters per image  $N = 1000$ ; the number of angles, i.e., haze-lines, in each plane  $K = 40$ ; the number of different atmospheric light candidates is  $M = \frac{1}{\Delta A}$ , where



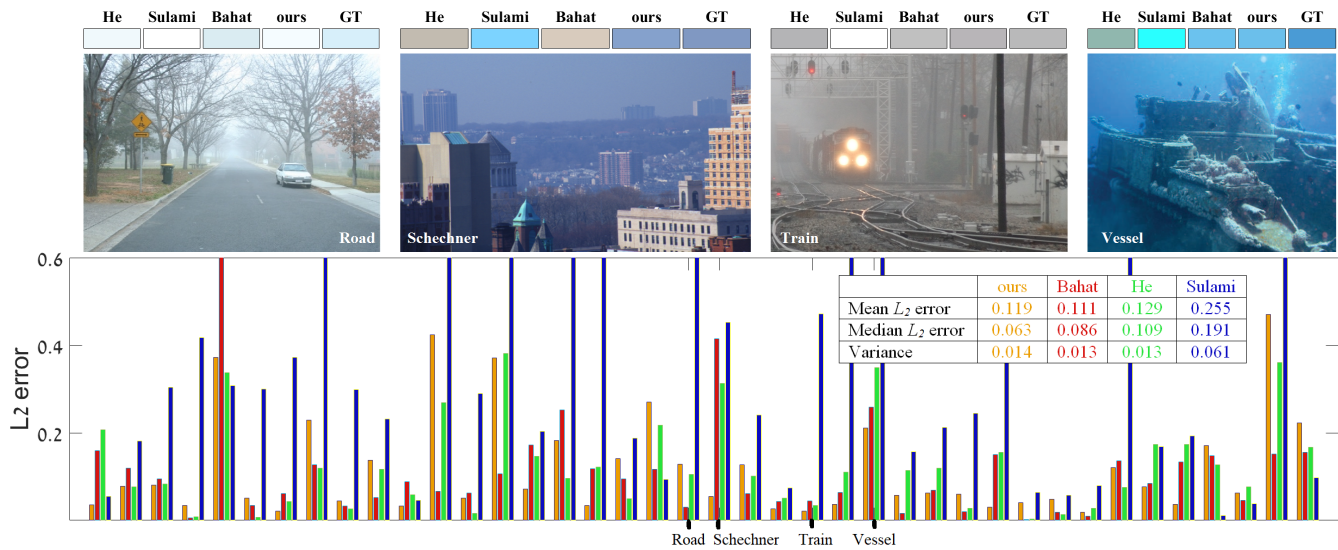


Fig. 8. **Evaluating the accuracy of the estimated atmospheric light on natural images.** Top: Examples of hazy images, along with their manually extracted ground-truth atmospheric light (GT), and the results of Sulami et al. [18], He et al. [5], Bahat and Irani [7], and ours. Bottom:  $L_2$  errors calculated on 40 hazy images (for which the ground truth could be manually reliably extracted).

$\Delta A = 0.02$  since the intensity range is normalized to  $[0, 1]$ ; and the threshold  $\tau_0 = 0.02$  determines whether a pixel  $I_n$  supports a certain haze-line. For transmission estimation: the number of different haze-lines is  $Q = 1000$ . The regularization in Eq. (20) is performed using  $\lambda = 0.1$  and we solve the equation using a closed-form solution to the weighted least squares problem, whose complexity is linear in the image size and is extremely efficient. The code is available at: <https://github.com/danaberman/non-local-dehazing>.

## 5.2 Atmospheric light Estimation

We start by evaluating the first step of the algorithm - the atmospheric light estimation. Bahat and Irani [7] compiled a diverse set of 40 images to quantitatively evaluate the accuracy of the estimated atmospheric light. This set consists of 35 images that contain points at an infinite distance from the camera, whose color is the atmospheric light  $A$ . Five additional images were generated by cropping, so that the sky is no longer visible in the image, yet the atmospheric light is known. This procedure verifies the algorithms' robustness in cases where the atmospheric light is not visible (as is often the case, for example in aerial photos). The authors manually marked the distant regions to extract the atmospheric light color that is used as ground-truth. Even though they graciously sent us their data, we also manually marked regions of extremely distant scene points, in order to evaluate the accuracy of the ground-truth, as well as the accuracy expected from an automatic algorithm. The median  $L_2$  difference between the two manual selections was 0.02, which we consider a lower bound to the accuracy of atmospheric light estimation methods.

The complete breakdown of this experiment is depicted in Fig. 8[bottom], where for each image we show the error of our method and the errors of [5], [7], [18]. We also present a table summarizing the errors. Generally, our method and [7] outperform [5] and [18]. Compared to [7], our method results in a lower median error, with slightly higher mean and variance. Interestingly, the performance comparison of both methods is not always consistent, i.e., on some images

we perform better than [7], and vice versa. The performance depends on the extent the image adheres to the prior used by each method.

A few example photos are shown in Fig. 8[Top] together with the ground-truth atmospheric light color and the ones estimated by the four methods. The error bars corresponding to them in Fig. 8[Bottom] are labeled. In *Schechner* our method outperforms all methods. The Hough votes for this image are depicted in Fig. 4. In *Vessel* all methods yield relatively high errors. This is probably because the atmospheric light is not uniform across the scene.

## 5.3 Natural Images

We present end-to-end dehazing results on natural images along with the estimated transmission maps, and compare them to state-of-the-art single image dehazing methods in Figs. 9 and 10. The comparison includes the seminal Dark Channel Prior [5] and more recent works published since 2014 [6], [7], [8], [22], [23]. Fig. 11 further compares methods for which the transmission is not available, either it is not published [8] or it is not estimated in the first place [25].

As previously noted by [5], the image after haze removal might look dim, since the scene radiance is usually not as bright as the atmospheric light. For display, we perform a global linear contrast stretch on the output, clipping 0.5% of the pixel values both in the shadows and in the highlights. Compared to recent methods [7], [22], [23], the proposed method is able to significantly reduce the amount of haze, while maintaining a realistic look. The green rectangles highlight areas where the proposed method is able to remove haze better. For example, in Tiananmen (Fig. 9 top) the proposed method shows the tree and the farthest building is great detail, while maintaining a realistic color of the sky. While the result of [7] shows fine detail on the farthest building, it comes at a cost of red tint in the sky and inconsistent dehazing of the tree. In New York (Fig. 9) the proposed method shows much more details of the land in the background than any of the other methods. In Fig. 10 top, the sign marked by the green rectangle is more easily readable in our result.



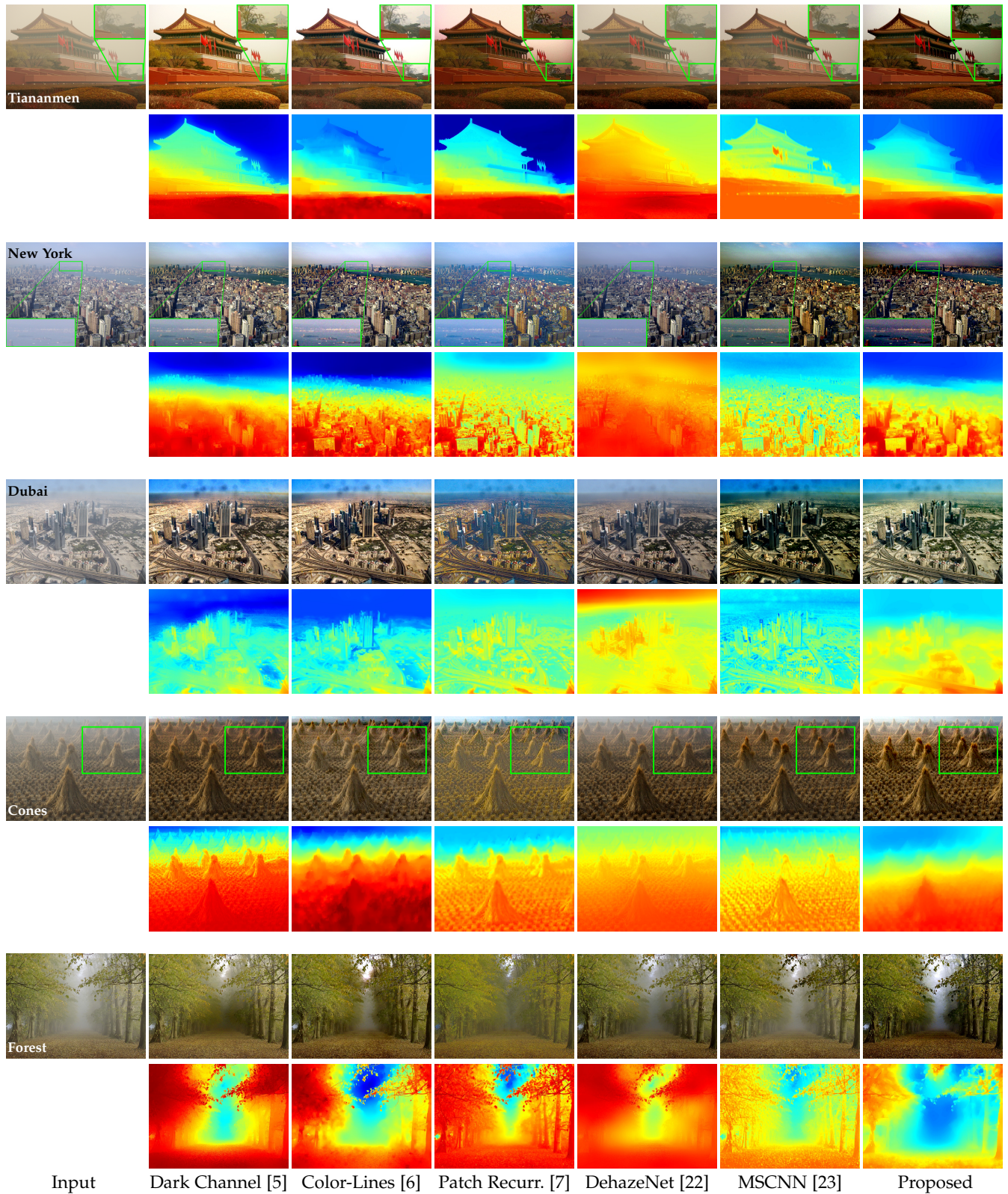


Fig. 9. **Natural images:** A comparison of several end-to-end single image dehazing methods. The leftmost column is the input image and the rest of the columns show the output of different methods. The transmission maps are shown below each output image.



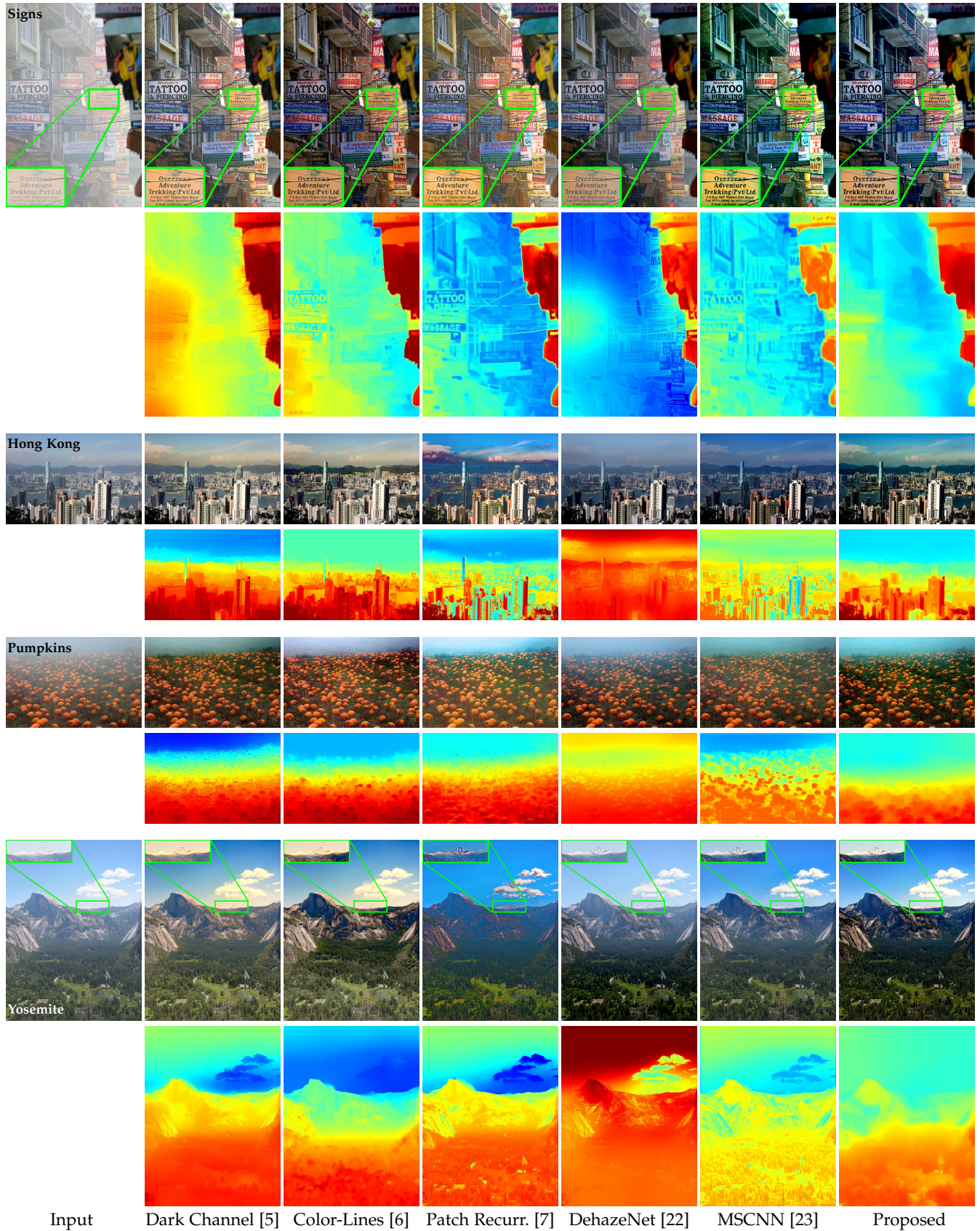


Fig. 10. **Natural images:** A comparison of several end-to-end single image dehazing methods. The leftmost column is the input image and the rest of the columns show the output of different methods. The transmission maps are shown below each output image.





Fig. 11. **Natural images:** A comparison of several end-to-end single image dehazing methods. The leftmost column is the input image and the rest of the columns show the output of different methods.

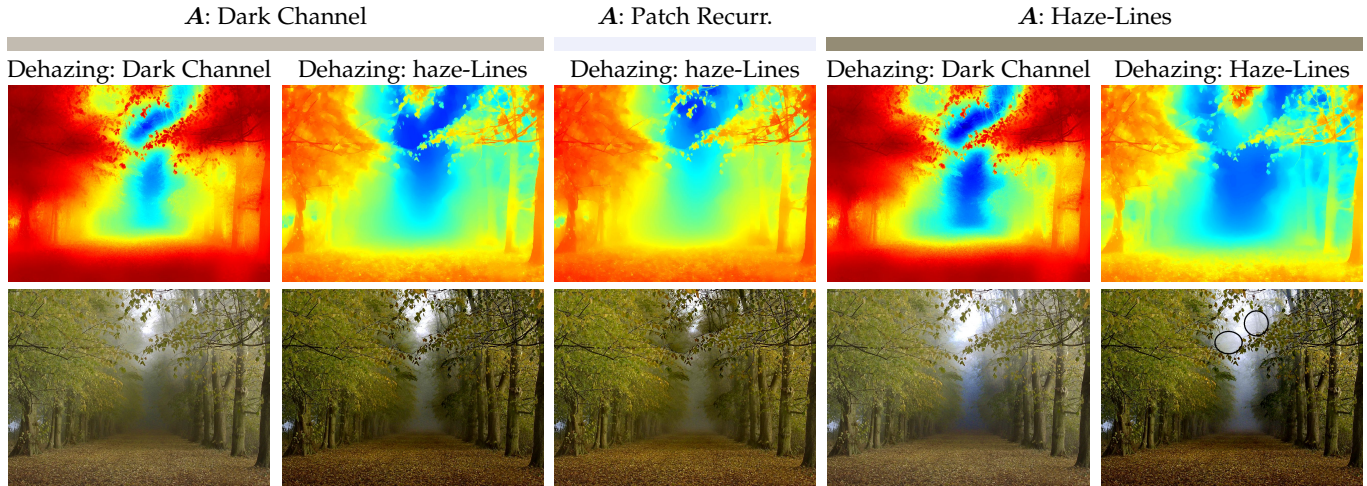


Fig. 12. **Intermixed atmospheric light and transmission estimations.** The input hazy image is shown in Fig. 9. The two left columns show dehazing results using the atmospheric light value estimated by the Dark Channel [5], which is more accurate for this image than the haze-lines [11] estimation. The center column shows the dehazing result using the atmospheric light estimate by Patch Recurrence [7], which is even closer to the ground truth than the Dark Channel. As a result, the haze removal process performs better on the farthest regions of the scene in the second and third columns from the left. This is evident in the area circled in black in the rightmost image. When comparing the second and third columns from the left, the second has dark shadows due to over-correction, whereas the center column maintains a better balance between dehazing and dynamic range.

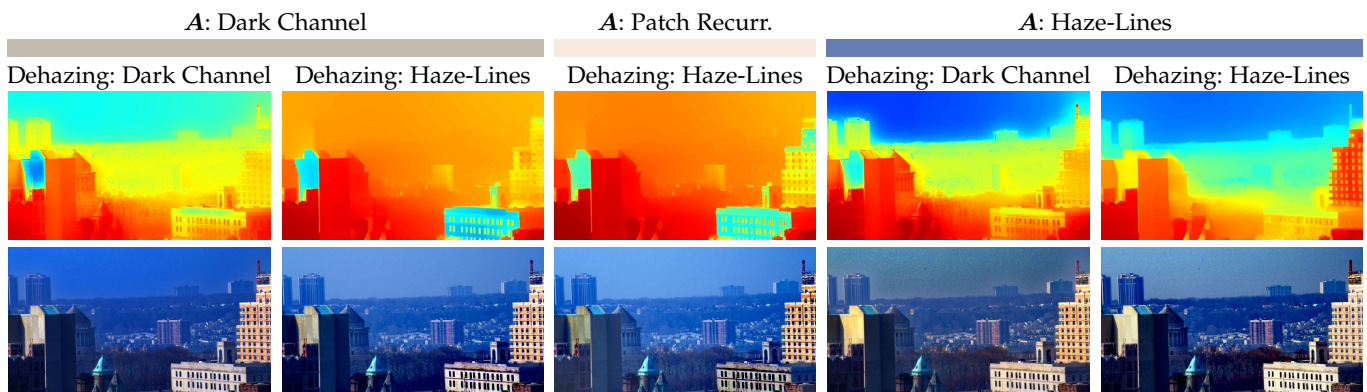


Fig. 13. **Intermixed atmospheric light and transmission estimations.** The input hazy image is shown in Fig. 8. The two left columns show dehazing results using the atmospheric light value estimated by the Dark Channel [5], while the center column shows the dehazing result using the atmospheric light value estimated by the Path Recurrence [7] (which is superior to [5] as shown in Section 5.2). Both of these estimations are not accurate for this image. This affects the next step in the pipeline, and as a result the respective transmission maps do not capture the 3D structure of the scene. The two rightmost column show dehazing results with the atmospheric light value estimated using Haze-Lines [11], which is more accurate for this image. Indeed, the corresponding transmission maps seem to better capture the scene distances.

The qualitative comparison of the dehazed images is somewhat subjective, but some insights can be gained from comparing the estimated transmission maps. For example, the top image in Fig. 10 shows the dehazing results along with the estimated transmission maps of a street scene. This is a challenging photo since many of the image gradients do not correspond to depth discontinuities, i.e., the street signs. The transmission is a function of the distance of the objects in the world from the camera, and therefore the signs should appear as planar surfaces rather than showing their text in the transmission. While none of the methods fully achieves this goal, the proposed method has less visible text in the transmission map.

The atmospheric light and the transmission estimation are two consecutive steps in the haze removal process, and different methods can be used for each step. Figs. 12 and 13 show dehazing results of intermixed techniques for atmospheric light and dehazing: Dark Channel [5] and Haze-Lines [10], [11]. Using different atmospheric light values shows the effect of the atmospheric light estimation on the output dehazed image. Fig. 12 shows the *Forest* image, for which the estimated atmospheric light using haze-lines [11] is darker than the Dark Channel estimation. This estimation error leads to an error in the transmission map, and some distant regions look faded in the output, as seen in the area circled in black. Fig. 13 depicts an example where the atmospheric light was wrongly estimated by the Dark Channel for the *Schechner* image, yet more accurately estimated by haze-lines. The wrong value estimated by [5] leads to a completely wrong transmission map in the two left columns, while the transmission in the two right columns approximately describe the scene structure.

## 5.4 Limitations

Unfortunately, when solving Eq. (20) with a closed form weighted-least squares (WLS) solution, the solution does not always conform to the lower bound in Eq. (18). This effect is demonstrated in Fig. 14, where the input image shown on the top left and the WLS solution is shown in the middle column. The bottom left is a binary mask showing where the estimated transmission is smaller than the lower bound derived in Eq. (18). This condition is violated in many image regions, resulting in a dark output and loss of details, as seen in the areas circled in red. Solving Eq. (20) under the constraint  $\hat{t} \geq t_{LB}$ , where  $t_{LB}$  is defined in Eq. (18) is possible using an interior point method [32], which prolongs the computation time, but also improves the accuracy of the solution. The dehazed image and the transmission map using the more rigorous optimization technique are shown on the rightmost column of Fig. 14. Indeed, the transmission is larger in the areas that are circled in red in the top middle image, and they have much more details in the top left solution. This is a trade-off between accuracy and run-time: using closed-form WLS takes on average 3.1sec/MPixel, while using the constrained WLS takes on average 121.9sec/MPixel (both implementations are in MATLAB, un-optimized, and run on a single core of desktop computer with 4th generation Intel Core i7 processor @3.4GHz and 32GB RAM).

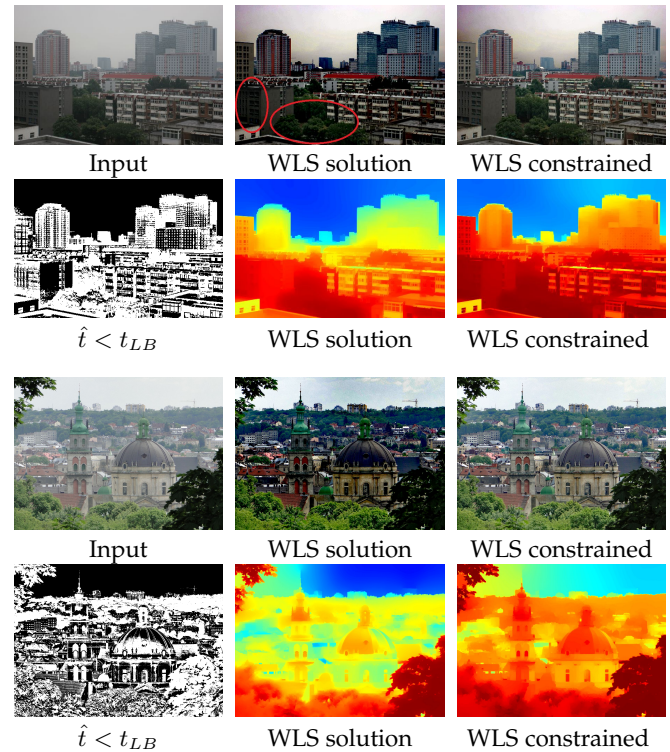


Fig. 14. **Limitations:** The input image is shown on the top left. The estimated transmission map obtained by a closed form weighted-least squares (bottom center) often violates the lower bound derived in Eq. (18). On the bottom left, the binary mask depicts the areas in which the transmission is smaller than the lower bound. As a result, the dehazed image (top center) is too dark and some details are missing, e.g. in the foliage. When solving a constrained optimization, the results on the right column show an image with much more details.

TABLE 1  
Comparison of SSIM and CIEDE2000 errors over O-HAZE [33], a dehazing benchmark with real hazy and haze-free outdoor images. HL\* on the bottom row is the constrained WLS solution, which is proposed in Section 5.4. For SSIM, higher is better, while for CIEDE2000 lower is better.

User / Method	SSIM	CIEDE2000
Dark Channel [5]	0.735	20.745
Meng et al. [17]	0.753	16.968
Color-Lines [6]	0.707	19.854
DehazeNet [22]	0.666	17.348
Night Fusion [27]	0.747	16.431
MSCNN [23]	0.765	<b>14.670</b>
HL [10], [11]	0.750	17.088
HL*	<b>0.781</b>	16.579

## 5.5 Quantitative Evaluation

It is quite difficult to record ground-truth clear scenes for the hazy images, due to varying illumination conditions. Nonetheless, such an effort was recently made using a smoke machine [33], where 45 scenes were captured in both clear conditions and in fog. Table 1 summarizes the average performance of various dehazing methods in terms of SSIM [34] and CIEDE2000 [35], which evaluate both structure and color fidelity, respectively. The last two rows show the baseline dehazing method proposed in [10], [11] and the improvement we proposed in Sec. 5.4. The latter is optimal in terms of SSIM, even compared to a learning-based method which used additional training data.



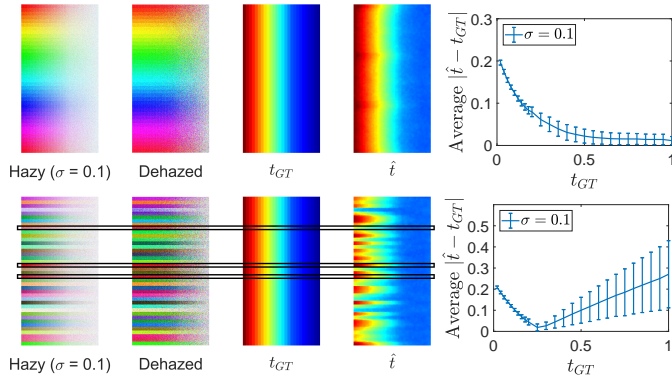


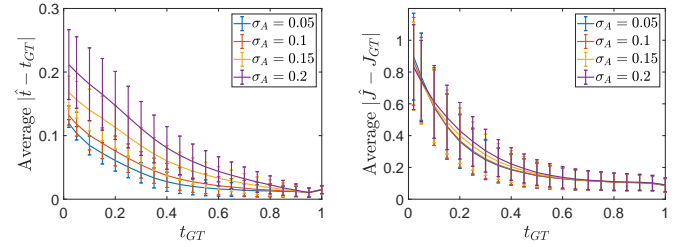
Fig. 15. **Robustness analysis:** Synthetic haze is added to images with colored stripes (leftmost column) in order to analyze the model's behavior when the transmission is small. The true transmission  $t_{GT}$  is monotonically decreasing from the left side of the image to the right. The dehazed result and estimated transmission  $\hat{t}$  are shown on the second and fourth columns, respectively. The graph on the right analyzes the transmission estimation error as a function of the magnitude of the transmission. The experiment is repeated for two different color distributions (on different rows), and demonstrates that the estimation accuracy depends on the color distribution in the image. However,  $\hat{t}$  is generally reliable for medium transmission values.

## 5.6 Robustness Analysis

When haze is considerably thick, the projection onto the RGB space will concentrate nearby the atmospheric light point. Due to the presence of noise in real images, the points near the atmospheric light will shift from their true location and the estimated transmission might not be accurate.

In order to evaluate the model's robustness, we generated a synthetic image and a synthetic transmission map as demonstrated in Fig. 15. The image consists of horizontal color stripes, where each row in Fig. 15 demonstrates a different color distribution. The transmission map is piecewise constant, decreasing from 1 on the left to 0.02 on the right in steps of 0.05 if  $t_{GT} > 0.2$  and 0.02 otherwise. The atmospheric light is randomly sampled with a normal distribution  $\sim \mathcal{N}(0.9, 0.01)$  (each color channel separately), where the pixel values are scaled to the range  $[0, 1]$ . The hazy image is synthesized according to Eq. (1) and random Gaussian noise is added to each element with  $\mu = 0, \sigma = 0.1$ , as well as Poisson noise, which was shown to decrease SNR in hazy scenes [36]. The estimated transmission  $\hat{t}$  is shown to the right of the true transmission. As evident from the dehazed image, when the transmission value is extremely small (on the right image border), the pixel values cannot be accurately restored. The graph on the right shows the average error in the transmission estimation, as well as the standard deviation. For transmission values of 0.15, the mean estimation error is about 50% of the transmission value, and the error increases as the transmission decreases.

The top row of Fig. 15 consists of colors with different hues. In practice, natural images may contain colors with similar hues and different saturation values, and some of these different colors may be co-linear with the atmospheric light. Such a case is demonstrated in the bottom row of Fig. 15, where the colors are uniformly and randomly samples from the range  $[0, 1]$ . The estimated transmission of each color depends on the other colors present in the image, as evident by the co-linear colors marked by black rectangles. The transmission value is recovered properly only for



a) Transmission estimation error b) Radiance estimation error

Fig. 16. **Dehazing sensitivity to atmospheric light:** a) The average transmission estimation error is shown as a function of the transmission magnitude, for various standard deviation values of the atmospheric light ( $\sigma_A$ ). As  $\sigma_A$  increases, so does the estimation error of the transmission. b) A similar error analysis is shown for the dehazed image. Unlike the transmission, the errors of the restored pixels values are not sensitive to the magnitude of  $\sigma_A$ .

one of the colors marked by the rectangles, whereas  $\hat{t}$  of the other colors is under-estimated. The average transmission estimation error when  $t_{GT} = 1$  is higher than areas in which  $t_{GT} < 0.6$ . However, the standard deviation when  $t_{GT} = 1$  is quite high, which means that  $\hat{t}$  is sometimes estimated correctly, but sometimes isn't, depending on the color content of the image. This graph explains why this method was found to best perform in medium haze [12].

## 5.7 Dehazing Sensitivity to Atmospheric Light

The proposed haze removal process consists of two successive steps: estimation of the atmospheric light (Sec. 4.1) and then of the transmission (Sec. 4.2), which are combined in Sec. 4.3 to generate the dehazed output  $\hat{J}$  (for simplicity we omit the spatial index  $x$ ). Eq. 21 describes the dependency of  $\hat{J}$  on the atmospheric light  $A$  and the transmission  $\hat{t}$ , which itself depends on  $A$  via the clustering. In order to estimate the effect of errors in the atmospheric light estimation, we simulated a hazy image similarly to Fig. 15(top), and dehazed it using a noisy atmospheric light, with increasing standard deviation. Fig. 16 summarizes the average results over several iterations. While the transmission errors (Fig. 16 left) increase significantly when the noise levels of  $A$  increase, the dehazed image (Fig. 16 right) is less affected, probably due to the complicated dependency on Eq. 21. Note that a standard deviation of 0.2 is quite high considering the empirical results in Fig. 8.

## 5.8 Failure Cases

Fig. 17 shows two failure cases, where the lighting in non-uniform and the sky show un-natural colors in both output images. On the bottom row, the transmission of the lake is severely under-estimated, leading to over-correction in contrast to the boats. Moreover, the transmission of the sky is estimated to be quite different on both sides of the tower, and as a result the sky on the left contains haze residues in the output image, while the sky on the right is over-corrected.

## 6 CONCLUSION

We introduced a novel non-local method for single image dehazing. The method is based on the assumption that an image can be faithfully represented with just a few hundreds of distinct colors. In RGB space, this corresponds to a few

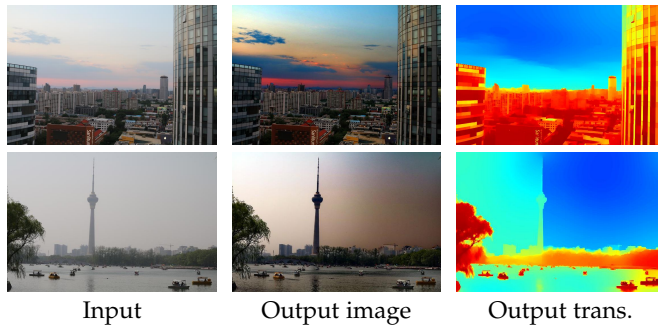


Fig. 17. **Failure cases:** Non-uniform lighting leads to artifacts in the restored images.

hundreds tight color clusters. We showed that in a hazy image, these tight color clusters change because of haze and form lines in RGB space that pass through the atmospheric light coordinate.

We proposed an efficient algorithm to predict the atmospheric light and then identify these haze-lines and estimate a per-pixel transmission based on them. We take into consideration the variance of our estimation in the regularization process, so only pixels that comply with the model assumptions contribute to the result. We show that a naïve regularization may fail, and propose a constrained weight-least-squares solution to overcome this limitation.

In contrast to previous methods our algorithm is pixel-based and not patch-based, and does not require any training. This makes our algorithm deterministic, fast, more robust and less prone to issues such as the choice of patch size, patch tiling, and patches with non-uniform content. Our method was tested and found to work well on many real-world images.

## ACKNOWLEDGMENTS

TT was supported by the The Leona M. and Harry B. Helm-sley Charitable Trust and The Maurice Hatter Foundation. Part of this research was supported by ISF grant 1917/2015. DB was supported by Apple Graduate Fellowship.

## REFERENCES

- [1] S. G. Narasimhan and S. K. Nayar, "Vision and the atmosphere," *International Journal of Computer Vision*, vol. 48, no. 3, pp. 233–254, Jul 2002.
- [2] W. E. K. Middleton, *Vision through the atmosphere*. Toronto: University of Toronto Press, 1952.
- [3] Y. Y. Schechner, S. G. Narasimhan, and S. K. Nayar, "Instant dehazing of images using polarization," in *Proc. IEEE CVPR*, 2001.
- [4] S. G. Narasimhan and S. K. Nayar, "Chromatic framework for vision in bad weather," in *Proc. IEEE CVPR*, 2000.
- [5] K. He, J. Sun, and X. Tang, "Single image haze removal using dark channel prior," in *Proc. IEEE CVPR*, 2009.
- [6] R. Fattal, "Dehazing using color-lines," *ACM Trans. Graph.*, vol. 34, no. 1, p. 13, 2014.
- [7] Y. Bahat and M. Irani, "Blind dehazing using internal patch recurrence," in *Proc. IEEE ICCP*, 2016.
- [8] K. Tang, J. Yang, and J. Wang, "Investigating haze-relevant features in a learning framework for image dehazing," in *Proc. IEEE CVPR*, 2014.
- [9] M. T. Orchard and C. A. Bouman, "Color quantization of images," *Signal Processing, IEEE Transactions on*, vol. 39, no. 12, pp. 2677–2690, 1991.
- [10] D. Berman, T. Treibitz, and S. Avidan, "Non-local image dehazing," in *Proc. IEEE CVPR*, 2016.
- [11] —, "Air-light estimation using haze-lines," in *Proc. IEEE ICCP*, 2017.
- [12] Y. Li, Y. Shaodi, M. S. Brown, and R. T. Tan, "Haze visibility enhancement: A survey and quantitative benchmarking," *Computer Vision and Image Understanding*, 2017.
- [13] R. T. Tan, "Visibility in bad weather from a single image," in *Proc. IEEE CVPR*, 2008.
- [14] J.-P. Tarel and N. Hautiere, "Fast visibility restoration from a single color or gray level image," in *IEEE Proc. ICCV*, 2009.
- [15] K. Nishino, L. Kratz, and S. Lombardi, "Bayesian defogging," *Int. Journal of Computer Vision (IJCV)*, vol. 98, no. 3, pp. 263–278, 2012.
- [16] K. B. Gibson and T. Q. Nguyen, "An analysis of single image defogging methods using a color ellipsoid framework," *EURASIP Journal on Image and Video Processing*, vol. 2013, no. 1, 2013.
- [17] G. Meng, Y. Wang, J. Duan, S. Xiang, and C. Pan, "Efficient image dehazing with boundary constraint and contextual regularization," in *Proc. IEEE ICCV*, 2013.
- [18] M. Sulami, I. Geltzer, R. Fattal, and M. Werman, "Automatic recovery of the atmospheric light in hazy images," in *Proc. IEEE ICCP*, 2014.
- [19] I. Omer and M. Werman, "Color lines: Image specific color representation," in *Proc. IEEE CVPR*, 2004.
- [20] C. O. Ancuti and C. Ancuti, "Single image dehazing by multi-scale fusion," *IEEE Trans. on Image Processing*, vol. 22, no. 8, pp. 3271–3282, 2013.
- [21] L. K. Choi, J. You, and A. C. Bovik, "Referenceless prediction of perceptual fog density and perceptual image defogging," *IEEE Transactions on Image Processing*, vol. 24, no. 11, pp. 3888–3901, 2015.
- [22] B. Cai, X. Xu, K. Jia, C. Qing, and D. Tao, "Dehazenet: An end-to-end system for single image haze removal," *IEEE Trans. on Image Processing*, vol. 25, no. 11, pp. 5187–5198, 2016.
- [23] W. Ren, S. Liu, H. Zhang, J. Pan, X. Cao, and M.-H. Yang, "Single image dehazing via multi-scale convolutional neural networks," in *European Conference on Computer Vision*, 2016.
- [24] N. Silberman, D. Hoiem, P. Kohli, and R. Fergus, "Indoor segmentation and support inference from rgbd images," *Computer Vision—ECCV 2012*, pp. 746–760, 2012.
- [25] B. Li, X. Peng, Z. Wang, J. Xu, and D. Feng, "Aod-net: All-in-one dehazing network," in *Proc. IEEE ICCV*, 2017.
- [26] Y. Li, R. T. Tan, and M. S. Brown, "Nighttime haze removal with glow and multiple light colors," in *Proc. IEEE ICCV*, 2015.
- [27] C. Ancuti, C. O. Ancuti, C. De Vleeschouwer, and A. C. Bovik, "Night-time dehazing by fusion."
- [28] A. Chakrabarti, D. Scharstein, and T. Zickler, "An empirical camera model for internet color vision," in *BMVC*, 2009.
- [29] H. Lin, S. J. Kim, S. Ssstrunk, and M. S. Brown, "Revisiting radiometric calibration for color computer vision," in *Proc. IEEE ICCV*, 2011.
- [30] G. Marsaglia, "Choosing a point from the surface of a sphere," *Ann. Math. Statist.*, vol. 43, no. 2, pp. 645–646, 04 1972.
- [31] R. L. Lagendijk, J. Biemond, and D. E. Boeke, "Regularized iterative image restoration with ringing reduction," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 36, no. 12, pp. 1874–1888, 1988.
- [32] A. Altman and J. Gondzio, "Regularized symmetric indefinite systems in interior point methods for linear and quadratic optimization," *Optimization Methods and Software*, vol. 11, no. 1-4, pp. 275–302, 1999.
- [33] C. O. Ancuti, C. Ancuti, R. Timofte, and C. D. Vleeschouwer, "O-HAZE: a dehazing benchmark with real hazy and haze-free outdoor images," in *Proc. IEEE CVPR, NTIRE Workshop*, 2018.
- [34] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE trans. on image processing*, vol. 13, no. 4, pp. 600–612, 2004.
- [35] G. Sharma, W. Wu, and E. N. Dalal, "The ciede2000 color-difference formula: Implementation notes, supplementary test data, and mathematical observations," *Color Research and Application*, vol. 30, no. 1, pp. 21–30, 2005.
- [36] T. Treibitz and Y. Y. Schechner, "Resolution loss without imaging blur," *J. Opt. Soc. Am. A*, vol. 29, no. 8, pp. 1516–1528, 2012.



**Dana Berman** Received her BSc degree in Electrical Engineering and her BA in Physics Technion-Israel Institute of Technology in 2007, and her MSc in Electrical Engineering from Tel Aviv University in 2013. She is a PhD Candidate at Tel Aviv University.



**Tali Treibitz** received her BA degree in Computer Science and her PhD degree in Electrical Engineering from the Technion-Israel Institute of Technology in 2001 and 2010, respectively. Between 2010-2013 she has been a post-doctoral researcher in the department of computer science and engineering, in the University of California, San Diego and in the Marine Physical Lab in Scripps Institution of Oceanography. Since 2014 she is heading the marine imaging lab in the school of Marine Sciences in the University

of Haifa. She was the recipient of the Google Anita Borg Scholarship in 2009 and the Weizmann Institute of Science National Postdoctoral Award for Advancing Women in Science in 2010.



**Shai Avidan** received the PhD degree from the School of Computer Science, Hebrew University, Jerusalem, Israel, in 1999. He is currently a professor on the Faculty of Engineering, Tel-Aviv University, Tel-Aviv, Israel. In between, he worked for Adobe, Mitsubishi Electric Research Labs, and Microsoft Research. He has published extensively in the fields of object tracking in video and 3D object modeling from images. He is also interested in Internet vision applications

such as privacy-preserving image analysis, distributed algorithms for image analysis, and image retargeting.