

Dehazing using Color-Lines

RAANAN FATTAL

The Hebrew University of Jerusalem

Photographs of hazy scenes typically suffer from having low-contrast and offer a limited visibility of the scene. This paper describes a new method for single-image dehazing that relies on a generic regularity in natural images where pixels of small image patches typically exhibit a one-dimensional distribution in RGB color space, known as color-lines. We derive a local formation model that explains the color-lines in the context of hazy scenes and use it for recovering the scene transmission based on the lines' offset from the origin. The lack of a dominant color-line inside a patch or its lack of consistency with the formation model allows us to identify and avoid false predictions. Thus, unlike existing approaches that follow their assumptions across the entire image, our algorithm validates its hypotheses and obtains more reliable estimates where possible.

In addition, we describe a Markov random field model which is dedicated for producing complete and regularized transmission maps given noisy and scattered estimates. Unlike traditional field models that consist of local coupling, the new model is augmented with long-range connections between pixels of similar attributes. These connections allow our algorithm to properly resolve the transmission in isolated regions where nearby pixels do not offer relevant information.

An extensive evaluation of our method over different types of images and its comparison to state-of-the-art methods over established benchmark images shows a consistent improvement in the accuracy of the estimated scene transmission and recovered haze-free radiances.

Categories and Subject Descriptors: I.3.3 [Computer Graphics]: Picture/Image Generation—*Display algorithms*; I.4.1 [Computer Graphics]: Digitization and Image Capture—*Radiometry*

Additional Key Words and Phrases: image dehazing, contrast enhancement, transmission estimation

ACM Reference Format:

Raanan Fattal 20013. Dehazing using Color Lines. ACM Trans. Graph. 28, 4, Article 106 (August 2009), 11 pages.

DOI = 10.1145/1559755.1559763

<http://doi.acm.org/10.1145/1559755.1559763>

Raanan Fattal acknowledges the Israeli Science Foundation (ISF) for supporting this research.

Author's addresses: raananf@cs.huji.ac.il.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies show this notice on the first page or initial screen of a display along with the full citation. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, to redistribute to lists, or to use any component of this work in other works requires prior specific permission and/or a fee. Permissions may be requested from Publications Dept., ACM, Inc., 2 Penn Plaza, Suite 701, New York, NY 10121-0701 USA, fax +1 (212) 869-0481, or permissions@acm.org.

© 2009 ACM 0730-0301/09/14-ART106 \$10.00

DOI 10.1145/1559755.1559763

<http://doi.acm.org/10.1145/1559755.1559763>

1. INTRODUCTION

Small dust particles or liquid droplets in the air, collectively known as aerosols, scatter the light in the atmosphere. This light deflection reduces the direct scene transmission and replaces it with a layer of previously-scattered ambient light known as airlight or veiling light. Consequently, photographs taken in hazy or dusty weather conditions, and even ones taken in relatively clear days but capture long distances, are often of low-contrast and offer a limited visibility of the scene. A similar difficulty is encountered in underwater photography.

Most image dehazing methods remove the layer of haze by recovering the direct scene radiance. These methods rely on a physical image formation model that describes the hazy image as a convex combination between the scene radiance and the atmospheric light. As we detail in Section 3, the coefficients of this linear combination correspond to the scene transmission (visibility) at each image pixel. In case of RGB images, this model consists of four unknowns per pixel, the scene radiance at each color channel and the transmission value, whereas the input image supplies only three constraints, the intensity of each channel.

In order to resolve this indeterminacy many methods require additional information about the scene, such as multiple images taken at different weather conditions [Narasimhan and Nayar 2000] or polarization angles [Schechner et al. 2001] and knowing the scene geometry [Kopf et al. 2008]. More recently, methods that alleviate these input requirements were developed. This is achieved either by relaxing the physical model, for example by seeking for an image of maximal contrast [Tan 2008], or by introducing additional assumptions over hazy scenes. For example, Fattal [2008] resolves the indeterminacy by assuming a local lack of correlation between the transmission and surface shading functions. While this approach is capable of providing physically-consistent estimates, it cannot be applied at regions where the two functions do not vary sufficiently. He et al. [2009] robustly estimate the transmission from pixels with a dark (low-intensity) color channel. This approach requires that such pixels are found across the entire image. Large regions of bright surfaces in the image bias towards under-estimated transmission.

In this paper we propose a new method for single-image dehazing that takes advantage of a generic regularity in natural images in which pixels of small image patches typically exhibit one-dimensional distributions in RGB color space, known as color-lines [Omer and Werman 2004]. We use this observation to define a local image formation model that reasons the color-lines in the context of hazy images and allows recovering the scene transmission based on the lines' offset from the origin. Moreover, the unique pixel distribution predicted by the formation model allows us to identify patches that do not exhibit proper color-lines and discard them. In contrast to existing approaches that follow their assumptions across the entire image, our algorithm validates its hypotheses and hence obtains more reliable transmission estimates where possible. In this paper we focus on estimating the transmission accurately and assume the atmospheric light vector is given.

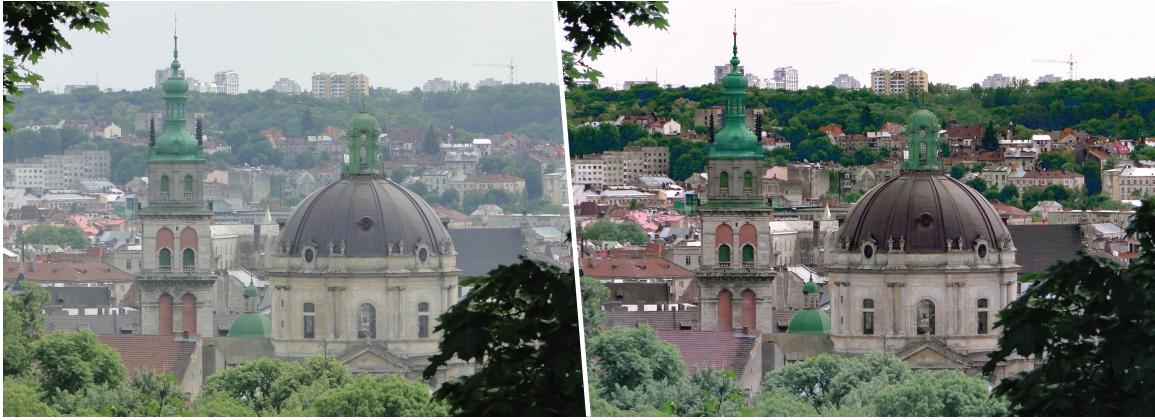


Fig. 1. Old Town of Lviv. Input image on the left, our result on the right.

In the last step of the algorithm, we interpolate and regularize these partial estimates into a complete transmission map using a dedicated Markov random field model. Unlike traditional field models which consist of regular coupling between nearby pixels, we augment the field model with long-range couplings. As we demonstrate, this new model better resolves the transmission in isolated regions where nearby pixels do not offer relevant information.

Finally, we report the results of an extensive evaluation of our method and its comparison to state-of-the-art techniques. This evaluation consists of a large number of benchmark images of different quality and resolution. We also use various types of synthetic images with known ground-truth in order to analyze the method's performance at different levels of noise and haze thickness. Our method shows a consistent improvement in the accuracy at which both the scene transmission and radiance are estimated.

2. PREVIOUS WORK

Image dehazing is a challenging problem which draws considerable attention from several research communities. Due to its ambiguous nature, many of the methods developed require additional data on top of the hazy image. Tan and Oakley [2000] assume the terrain geometry is known and estimate the pose of forward-looking airborne camera in order to obtain the transmission in the scene. A user-assisted registration process, between the image and a known scene geometry, is described by Kopf et al. [2008]. Schechner et al. [2001] remove haze effects given two or more photographs taken at different polarization angles. The polarization angle affects the magnitude of the polarized airlight and given a parameter, relating these changes to optical thickness, the polarized airlight is removed. Schwartz et al. [2006] estimate this parameter automatically by assuming that higher spatial-bands of the scene radiance are uncorrelated with the polarized haze. The success of the polarization-based approach depends on the extent at which the airlight is polarized in the scene. Nayar and Narasimhan [1999] estimate the scene structure from multiple images with and without haze, assuming the surface radiance remains unchanged. In a later work Narasimhan and Nayar [2003] describe a user interactive tool for removing weather effects.

A different line of work alleviates the input requirements by following various assumptions over hazy scenes. Oakley and Bu [2007] assume a constant layer of airlight and estimate its thickness, from a single image, based on an expected proportionality

between the local sample mean and the standard deviation of pixel intensities which is typically encountered in natural images. In this work we derive a localized model predicting this behavior and use it for recovering spatially-varying airlight layer. The dark-object subtraction method of Chavez [1988] also removes a uniform layer of haze by subtracting the color of the darkest object. This color is used as an approximation for the airlight present in the scene and it is found manually by inspecting offsets in the image histograms. Zhang et al. [2002] automate and extend this process for multi-spectral images acquired by satellite sensors. Du et al. [2002] assume the haze contribution resides in the lower part of the image spectrum and eliminate it based on a reference haze-free image.

More recent methods extract a spatially-varying layer of haze from a single image by following more refined assumptions over the scene. Tan [2008] extracts the haze by maximizing the resulting image contrast as well as transmission smoothness. This method generates compelling images with enhanced contrast, however it may also result in a physically-invalid excessive haze removal. Tarel and Hautière [2009] also promote high image contrast yet circumvent the time-consuming optimization by computing the transmission explicitly, based on an envelope function that ensures positive output pixels.

Fattal [2008] estimates the transmission based on an lack-of-correlation assumption between the transmission and shading functions. As we mentioned earlier, this approach requires a sufficient variation in these functions in order to obtain a reliable transmission estimate. Kratz and Nishino [2009; 2012] model the gradient distribution of the scene depth and radiance functions using heavy-tail distributions and recover these functions by further assuming statistical independence between the two.

He et al. [2009] generalize the dark-object subtraction method by inferring the transmission, locally, from dark-channel pixels found within a small neighborhood. While the prior that pixels with at least one dark channel can be found nearby holds in many regions of the image, often there are large regions where only bright pixels are available. Gibson and Nguyen [2011] explain the effectiveness of this approach using principal component analysis and minimum volume ellipsoid approximation. Carr and Hartley [2009] combine the dark-channel prior with a piecewise planar prior over the scene geometry using the alpha-expansion energy minimization framework. Matlin et al. [2012] combine the dark-channel approach with non-parametric denoising. More recently Gibson and Nguyen [Gibson and Nguyen 2013a] suggested a new dark prior for image de-

hazing within this framework. Unlike the dark-channel prior that assumes zero minimal value, the new prior seeks for the darkest pixel average inside each ellipsoid. This assumption may also be inaccurate over pixels that correspond to bright objects.

3. NEW APPROACH

3.1 Background

Aerosols present in the atmosphere deflect the light from its linear propagation to other directions in a process known as light scattering. Repeated scattering events across the medium reduce the visibility by creating a semi-transparent layer of ambient light, known as airlight [Koschmieder 1924]. This physical scenario is expressed by the following image formation model

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

where $I(\mathbf{x})$ is the input image, $J(\mathbf{x})$ is the scene radiance, i.e., the light reflected from its surfaces, and $\mathbf{x} = (x, y)$ denotes the pixel coordinates. The direct transmission of the scene radiance, $t(\mathbf{x})J(\mathbf{x})$, corresponds to the light reflected by the surfaces in the scene and reaching the camera directly, without being scattered. The airlight, $(1 - t(\mathbf{x}))A$, corresponds to the ambient light that replaces the direct scene radiance. The *atmospheric light* vector A describes the intensity of the ambient light. The use of a constant atmospheric light is a valid approximation when the aerosol reflectance properties as well as the dominant scene illumination are approximately uniform across the scene. We consider RGB images and hence Eq. (1) is a three-dimensional vector equation, where each coordinate corresponds to a different color channel. The scalar $0 \leq t(\mathbf{x}) \leq 1$ denotes the transmission along the camera ray at each pixel. These values correspond to the fraction of light crossing the medium, along camera rays, without being scattered. Unlike the atmospheric light A , the transmission is allowed to vary across the image and hence Eq. (1) applies to scenes of arbitrary optical depth and scattering coefficient (e.g., due to changes in aerosol density).

Many image dehazing algorithms use the image formation model in Eq. (1) to dehaze images by recovering J . This includes the recent single-image methods that perform this operation solely based on I ; either by first estimating the transmission [Fattal 2008; He et al. 2009; Tarel and Hautière 2009], or together with J in a joint optimization [Tan 2008; Carr and Hartley 2009; Kratz and Nishino 2009]. In this paper we follow the former group of methods and estimate the transmission first. Both these strategies require knowing the global atmospheric light vector A which can be estimated by various procedures, namely [Tan 2008; Fattal 2008; He et al. 2009; Tarel and Hautière 2009]. In this work we focus on estimating the transmission accurately and assume A is known. Finally, note that Eq. (1) assumes the input pixels values $I(\mathbf{x})$ are radiometrically-linear. Thus, similarly to other methods that rely on this formation model, our method requires the reversal of the acquisition nonlinearities.

3.2 Local Color-Line Model

Natural environments are typically composed of distinct objects, each with its own surface reflectance properties. Modeling natural images as a collage of projected surfaces showed success in matching various empirical statistics, see the work of Huang and Mumford [1999] and the dead-leaves model of Lee et al. [2001]. Motivated by these findings we assume that many small image patches correspond to mono-chromatic surfaces and admit the following

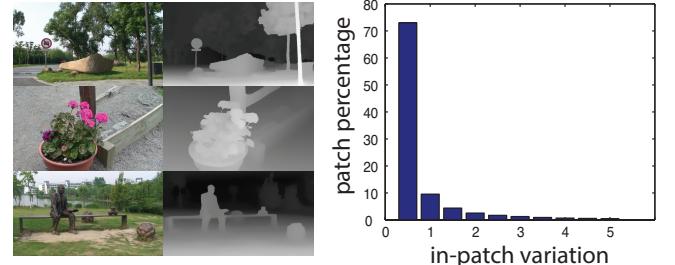


Fig. 2. Piecewise smoothness of the transmission. Bars show the percentage of image patches (vertical axis) in which the transmission values deviate from their average by less than a certain percent (horizontal axis).

factorization of the scene radiance

$$J(\mathbf{x}) = l(\mathbf{x})\bar{R}, \mathbf{x} \in \Omega \quad (2)$$

where \bar{R} is an RGB vector describing the relative intensity of each color channel of the reflected light, i.e., $\|\bar{R}\| = 1$. The scalar $l(\mathbf{x})$ describes the magnitude of the radiance at each pixel \mathbf{x} in the patch Ω . This assumption is successfully used in various dehazing methods [Schechner et al. 2001; Narasimhan and Nayar 2003; Fattal 2008]. While this model applies to more general surfaces, in case of purely diffuse surfaces \bar{R} corresponds to the surface reflectance coefficients and l to the incident light projected onto the surface. For simplicity we refer to \bar{R} as the surface reflectance or albedo and to l as the shading.

Natural environments are further characterized by being composed of nearly-planar object surfaces. This analogous collage description is also supported by studies of range images [Huang et al. 2000] and optical flow fields [Roth and Black 2007]. In addition, the density of dust, water droplets and other aerosols varies smoothly in space due to diffusion processes that govern these particles [Hinds 1999]. The combined effect of these regularities is inherited by the scene transmission due to the following relation

$$t(\mathbf{x}) = \exp\left(-\int_0^{d(\mathbf{x})} \beta(\mathbf{r}_x(s))ds\right), \quad (3)$$

where $d(\mathbf{x})$ is the depth and $\mathbf{r}_x(s)$ parametrizes the ray at pixel \mathbf{x} . The function $\beta(\cdot)$ denotes the scattering coefficient (in three-dimensional space).

Thus, since we expect piecewise smooth scene depths d and a smooth aerosol density, which in turn leads to a smooth scattering coefficient β , the rule of function composition implies that the resulting transmission $t(\mathbf{x})$ is also piecewise smooth function which is smooth at pixels that correspond to the same object. The assumption of piecewise smooth geometry is also used by Carr and Hartley [2009] and Kratz and Nishino [2009].

Figure 2 shows that in 72% of the patches of images of outdoor scenes, the transmission does not vary from its average by more than 0.5%, i.e., $|t - \bar{t}|/t < 0.05$ where \bar{t} is the average transmission in the patch, and that in 82.5% of the patches the variation is below 1%. This statistic was gathered from the outdoor depth maps acquired by Zhang et al. [2009] under the assumption of a constant scattering coefficient (more detail about this dataset is given at Section 4).

By taking into account both the transmission smoothness with the surface albedo constancy we use the following models to describe small image patches

$$I(\mathbf{x}) = tl(\mathbf{x})\bar{R} + (1 - t)A = l(\mathbf{x})R + (1 - t)A, \mathbf{x} \in \Omega \quad (4)$$

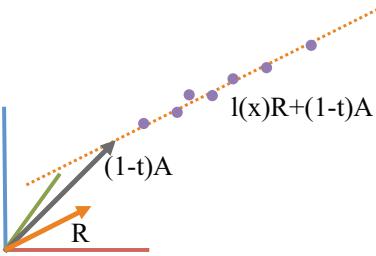


Fig. 3. Intersection between the atmospheric light (dark gray arrow) and a color-line (dashed-line) in RGB space. The point of intersection is of distance $(1 - t)$ along A .

where t is a fixed transmission value in the patch Ω and $R = t\bar{R}$.

Pixels of a patch Ω obeying this model differ only by the surface shading $l(\mathbf{x})$. Thus, their values $\{I(\mathbf{x}) : \mathbf{x} \in \Omega\}$ are distributed along a one-dimensional line in RGB space. This *patch color-line* is parameterized by the pixel shading l , its orientation coincides with the patch albedo R , and it is shifted from the origin by the airlight contribution, $(1 - t)A$. This configuration is illustrated in Figure 3. Studies of haze-free natural images report the existence of color-lines in RGB space, however, unlike our scenario these lines pass through the origin, see [Omer and Werman 2004].

Model Validation. The formation model in Eq. (4) does not apply for every image patch. For example, it is highly unlikely that both the albedo and depth (and hence the transmission) will be smooth in patches containing a boundary between different objects. Thus, the unique linear pixel distribution in RGB space predicted by our model makes it possible to identify and discard patches that do not obey it. In Section 3.4 we describe various criteria, derived from Eq. (4), that we use to prune patches.

This is in contrast to existing approaches, such as [Fattal 2008] and [He et al. 2009], where no verification of the model validity is made (besides a rather conservative elimination of pixels based on signal-to-noise estimates in case of [Fattal 2008]). More specifically, it is always possible to find an airlight-albedo separation that results in zero-correlation in Fattal's method and, similarly, every non-negative value is a valid dark-channel value, whether it is produced solely by the airlight or not. In sections 3.4 and 4, we show how the ability to verify the assumptions made over the image plays a central role in the overall robustness and accuracy of the method. We proceed by explaining how the transmission is estimated from the color-line model in patches where valid lines are found.

3.3 Transmission Estimation

In the next section we describe the way we recover color-lines inside small image patches and assume here that the line found is given by $lD + V$, where $D, V \in \mathbb{R}^3$, and $l \in \mathbb{R}$ is now considered as the free line parameter. Thus, given the color-line, we recover the transmission by finding its offset from the origin, which according to Eq. (4), is of length $1 - t$ along A (see Figure 3). More specifically, we search for the offset $s \in \mathbb{R}$ along A that shifts the line such that it passes through the origin, i.e., there exists $l \in \mathbb{R}$ such that $lD + V - sA = 0$. This is geometrically equivalent to intersecting the color-line $lD + V$ with the line passing through the origin in the orientation of the atmospheric light vector, sA . In practice we compute this intersection by solving

$$\min_{l,s} \|lD + V - sA\|^2, \quad (5)$$

where we relax the exact geometric operation by a minimization problem that copes with inaccuracies in the estimated D and V (and perhaps A). This quadratic objective is minimized by solving a 2-by-2 linear system (Eq. (10) at the Appendix) which gives s (and l). According to Eq. (4), the patch transmission is given by $t = 1 - s$. This value is expected to be a physically-consistent estimate in patches with approximately constant surface albedo and transmission.

Relation to Existing Methods. As we mentioned earlier, Oakley and Bu [2007] estimate a constant layer of haze based on an expected proportionality between the local sample mean and the standard deviation of the pixel intensities. This proportionality is also predicted by our color-line model, and its bias can be estimated by the procedure we describe here. However, unlike the Oakley and Bu's model, our localized patch-based model allows us to estimate a spatially-varying scene transmission.

Gibson and Nguyen [2013a] model the pixel histogram using ellipsoids, computed using principal-component analysis. The scene transmission is estimated as the one that minimizes the centroid of the dehazed color ellipsoid, i.e., by searching for the darkest image on average. Unlike our local color-line model the ellipsoid axes do not directly participate in this process and the transmission is not recovered from their offset from the origin. Thus, the two methods follow different assumptions and consist of different transmission estimation procedures.

3.4 Dehazing Algorithm

In this section we explain the steps that we carry out in order to dehaze an image using the local patch model in Eq. (4) and its associated transmission estimation procedure in Eq. (5). We begin with a brief overview of the algorithm. An outer loop of the algorithm scans the input image and considers small windows of pixels as candidate patches that obey Eq. (4). As discussed in the previous section, pixels that correspond to a nearly-planar mono-chromatic surface lie on a color-line in RGB space described by Eq. (4). Therefore, in each patch we run a RANSAC procedure that searches for a line supported by a significant number of pixels. We then check whether the line found is consistent with our formation model by testing it against a list of conditions posed by the model. A line that passes all these tests successfully is then used for estimating the transmission according to Eq. (5). The resulting value then is assigned to all the pixels that support the color-line found. We do not estimate the transmission in patches where we fail to find a line that meets all the conditions. Thus, it is likely that not all the image pixels receive a transmission estimate.

At the last step of the algorithm, we interpolate and regularize the transmission over the entire image using a dedicated Gauss-Markov random field model. Given the complete transmission map, we recover the output image J from I according to Eq. (1). We proceed by describing each of these steps and provide the details of our implementation. The parameter values quoted here apply for images with pixels values between zero and one.

Image Scan. Estimating the transmission at every possible image window is costly and redundant due to their overlap. We use a procedure that limits the number of overlapping transmission estimations while attempting to achieve a uniform coverage of the image. The idea is to scan a non-overlapping grid of square patches that cover the entire image and, since some patches are likely to be discarded, this process is repeated at different grid offsets. In this process we keep track of the number of transmission estimates obtained at each pixel and skip patches in which the center pixel received enough estimates (three or more in our implementation).

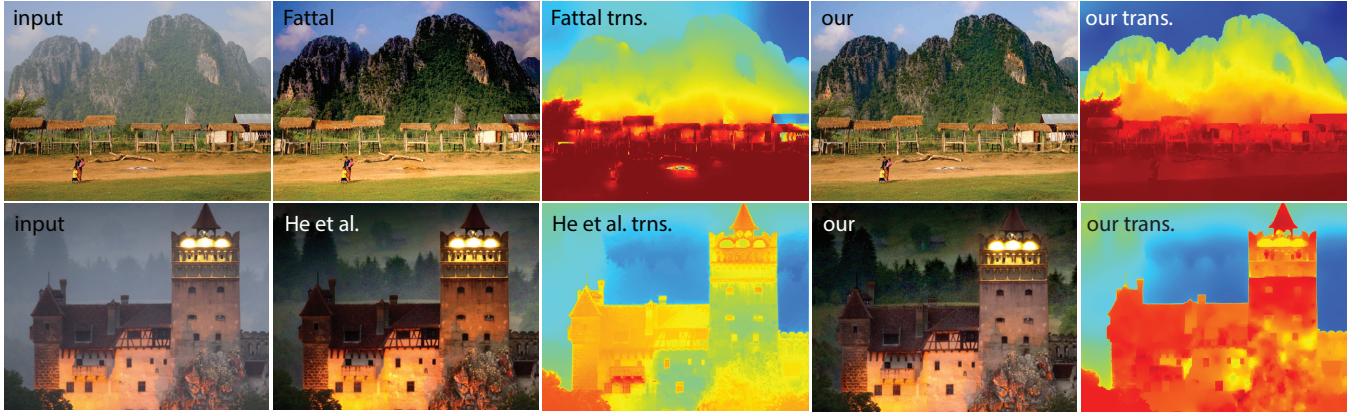


Fig. 4. Evaluation against the methods of Fattal [2008] and He et al. [2009]. Left-to-right are: the input images, dehazed images and transmission produced by Fattal (top) and He et al. (bottom) and our output radiance and transmission.

This way, the more challenging regions of the image are scanned multiple times and less work is performed in other regions. In our implementation we use patches of 7-by-7 pixels and scan the image four times by offsetting the grids by half the patch size, 3 pixels at each axis.

Color-Line Recovery. We estimate the color-line robustly using a RANSAC procedure. This process consists of picking random pairs of pixels in a patch (30 in our implementation), counting the number of patch pixels that lie close to the color-line defined by each pair, and picking the line that receives the largest number of supporting pixels. Then, we check whether the color-line found is consistent with our formation model by running it through a list of accept/reject tests. In case the line passes all the tests, it is used for estimating the transmission over the supporting pixels in the patch.

More formally, given two pixels, $\mathbf{x}_1, \mathbf{x}_2 \in \Omega$, randomly selected from a patch Ω , we consider the candidate line $lD + V$ defined by

$$D = I(\mathbf{x}_2) - I(\mathbf{x}_1), \quad \text{and} \quad V = I(\mathbf{x}_1). \quad (6)$$

Each line is associated with pixels $\mathbf{x} \in \Omega$ that support it, i.e., pixels in which $I(\mathbf{x})$ is sufficiently close to the line. This is measured by projecting $I(\mathbf{x}) - V$ onto the plane perpendicular to D and computing the norm of the projected vector. In our implementation we associate a pixel with the line if the norm falls below 2×10^{-2} . In order for this line to be considered as the patch's color-line we require it to meet each of the following conditions.

Significant line support. A small number of supporting pixels implies that either the line fails to represent the patch pixels or that most of its pixels do not obey Eq. (4) as its underlying assumptions do not hold. Therefore, we discard lines with less than 40% pixel support in the patch. If the line passes this test, we redefine the set of patch pixels Ω to be the subset of pixels that support it and do not consider the rest of the pixels in the following tests.

The local formation model derived in Section 3.2 predicts a unique behavior over the patch pixels and the line on which they lie. Not every line found is consistent with this model and hence we apply the following tests to identify and reject lines that cannot be reasoned by our model.

Positive reflectance. The color-line orientation D , as discussed in Section 3.3, corresponds to the surface reflectance vector R in Eq. (4). Therefore, we discard lines in which negative values are found in its orientation vector D . More precisely, since we ob-

tain D up to an arbitrary factor, we identify this inconsistency when D 's show mixed signs.

Large intersection angle. The operation of computing the intersection of two lines, as we do in Eq. (5), becomes more sensitive to noise as their orientation gets closer. At the Appendix we show that the error in the estimated transmission grows like $O(\theta^{-1})$, where θ is the angle between the line orientation D and atmospheric light vector A . Thus, we discard lines with $\theta < 15^\circ$ and weigh the confidence of the estimated transmission accordingly when interpolating these values to a complete transmission map (explained below). Figure 5 shows an example of patches with small and large intersection angles.

Unimodality. According to the collage model, discussed in Section 3.2, the image is expected to be made of pixels that correspond to piecewise nearly-planar mono-chromatic surfaces. The window patches we are examining may contain interfaces between two or more surfaces (edges in the image). It may be the case that in such patches a line connecting the two clusters of pixels will be proposed, however these pixels cannot be reasoned by Eq. (4) and the line must be rejected.

We identify these cases by examining the modality of the pixels' distribution along the line found by computing

$$\frac{1}{|\Omega|} \sum_{\mathbf{x} \in \Omega} \cos(a \langle I(\mathbf{x}) - V, D \rangle + b), \quad (7)$$

where the scalars a and b are set to shift and stretch the line parameters $\langle I(\mathbf{x}) - V, D \rangle$ of the patch pixels such that their extents coincide with the interval $[0, \pi]$. The $\langle \cdot, \cdot \rangle$ denotes the dot-product in RGB space. This measure consists of projecting the line parameters onto a function which is positive at the two ends, 0 and π , and negative in the middle (third Fourier mode). Therefore, Eq. (7) vanishes over uniformly distributed pixels and becomes positive when the pixels are concentrated near the endpoints. In our implementation we discard lines in which this value is above 7×10^{-2} .

Close intersection. Eq. (5) searches for a point on the airlight line and a point on the color-line which are closest to one another. While two arbitrary lines in three-dimensional space do not necessarily intersect, the lines predicted by our model are expected to do so. This requirement introduces another line admissibility test; we discard lines that produce intersection error, i.e., a minimal value in Eq. (5), which is above 5×10^{-2} .

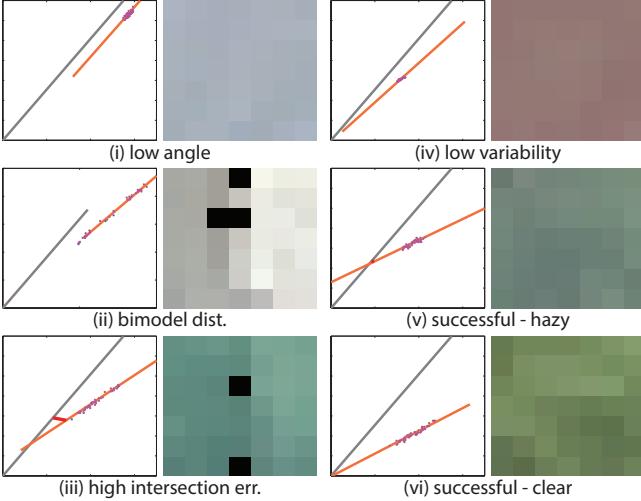


Fig. 5. Example color-lines. Left-column, top-to-bottom are: (i) a discarded patch in which the angle between color-line found and A is below 1° (resulting with $t = -1/2$), (ii) a patch containing an edge and hence discarded for its bimodal pixels distribution and (iii) a patch containing glossy object, the green dome in Figure 1, and discarded for its high intersection error in Eq. (5) (shown by the red line). Right-column, top-to-bottom are: (iv) a patch rejected for its low variability in shading, and (v) a successful estimation of the transmission in a hazy patch (positive transmission) and (vi) a clear patch (intersection at origin). The black pixels and blue dots correspond to pixels that do not support the color-line found. The plots show the pixels and lines projected to the red-green plane.

Valid transmission. Similarly, the intersection computed by solving Eq. (5) may not result in a valid transmission value, $0 \leq t \leq 1$. Thus, we discard patches in which the intersection results in values outside this admissible range.

Sufficient shading variability. As we noted in Section 3.3, the color-line is parameterized by the shading of each pixel, $l(\mathbf{x})$. Thus, the variability in the shading within the patch determines the length of the segment occupied by its pixels along the color-line. In presence of noise, the shorter this segment is, the less reliable the estimated line orientation D becomes. Thus, in principle it is preferable to discard patches whose pixels occupy very short segments. Note however that the segment length also depends intrinsically on the transmission in the patch since the latter multiplies the shading in Eq. (4). This means that the lower the transmission is, the shorter this segment becomes. Thus, in our decision of whether to use or discard a patch, we measure the segment length with respect to the transmission estimated from it. We ensure this self-consistency by computing the standard deviation of the line parameters normalized by the estimated patch transmission value,

$$\sqrt{\text{Var}_\Omega[\langle I(\mathbf{x}) - V, D \rangle]} / t, \quad (8)$$

where Var_Ω denotes the empirical variance, computed from the patch pixels. In our implementation we discard the patch if this value fall below 2×10^{-2} .

Figure 5 shows example color-lines that fail some of these tests as well as ones that succeed in estimating t . As discussed in Section 3.2, existing methods do not verify their assumptions and may therefore obtain wrong estimations. Figure 4 demonstrates this in relation to the methods of Fattal [2008] and He et al. [2009]. The former underestimates the transmission both at the mountains and

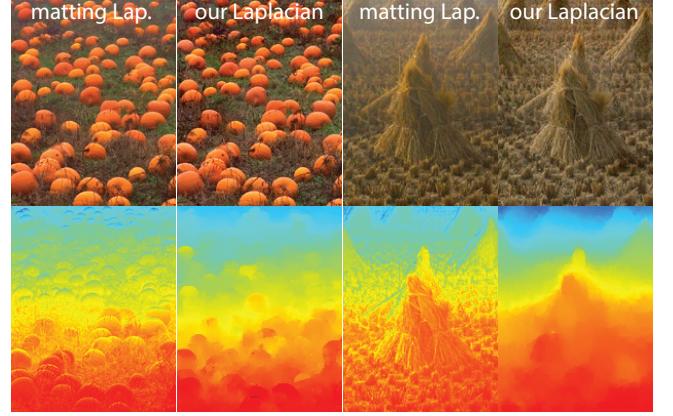


Fig. 6. Comparison between the matting Laplacian and our MRF. Assuming local linear relation between $I(\mathbf{x})$ and $t(\mathbf{x})$ transfers small-scale image content into $t(\mathbf{x})$ and weakens them in $I(\mathbf{x})$. In contrast, our transmission maps are piecewise smooth and hence the fine image details are preserved. The average transmission levels estimated by both methods are approximately equal in these examples and hence differences in clarity should be attributed to the detail transfer effect.

the hut's roof (these regions obtain transmission values which are as low as the sky). Our method rejects the roof's patches due to the small-angle condition and achieves more accurate results. The dark-channel assumption is not met across the bright castle pixels and hence the method of He et al. underestimates the transmission and over-correct the result. Once again, these biases are confirmed by inspecting the transmission maps where the method of He et al. produces highly-varying estimates across the castle's pixels which share roughly the same distance from the camera. The over-corrected pixels correspond to the lower transmission values estimated (color-coded in green).

3.5 Transmission Interpolation and Regularization

While the procedure described above typically manages to resolve the transmission over a fairly large portion of the image pixels, there still remains a significant number of pixels where it fails to provide an estimate. Moreover, the list of conditions we use to prune patches are necessary, but not enough to guarantee that the line found obeys our model. Therefore, similarly to Fattal [2008] and He et al. [2009], we obtain a complete transmission map and cope with errors due to noise and modeling inaccuracies by applying a Laplacian-based interpolation and regularization step to which we feed the partially estimated transmission values $\hat{t}(\mathbf{x})$ obtained at the previous step.

This regularization is based on imposing the smoothness of the input image $I(\mathbf{x})$ over the output transmission map $t(\mathbf{x})$ by maximizing the following Gauss-Markov random field (GMRF) model

$$P(t) \propto \exp \left(-\sum_{\Omega} \sum_{\mathbf{x} \in \Omega} \frac{(t(\mathbf{x}) - \hat{t}(\Omega))^2}{(\sigma_t(\Omega))^2} - \sum_{\mathbf{x}} \sum_{\mathbf{y} \in N_{\mathbf{x}}} \frac{(t(\mathbf{x}) - t(\mathbf{y}))^2}{\|I(\mathbf{x}) - I(\mathbf{y})\|^2} \right), \quad (9)$$

where Ω runs over all the patches in which a transmission estimate $\hat{t}(\Omega)$ is available, and $N_{\mathbf{x}}$ denotes the set of four-nearest neighbors of each pixel \mathbf{x} in the image.

The data term, left sum in Eq. (9), results from modeling the error in the estimated transmission as Gaussian noise with variance, $\sigma_t(\Omega)$, which expresses the amount of uncertainty in the estimated

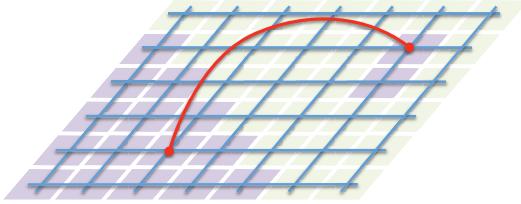


Fig. 7. Augmented Markov random field model. Blue lines illustrate the regular four-point coupling between nearby pixels in the model. The red line indicates a connection between distant pixels of similar color. In case the two-pixel island lacks its own transmission estimate, the red link can resolve these values.

values. In the Appendix we derive this model by assuming that the error in the estimated color-line (due to noise in the input pixels) is a zero-mean Gaussian variable with variance σ^2 , and obtain that $\sigma_t(\Omega) = \sigma \|A - D\langle D, A \rangle\| (1 - \langle D, A \rangle^2)^{-1}$. The pixel noise level σ can be tuned in case of known acquisition conditions such as ISO setting, aperture size and exposure time.

The regularization term, right sum in Eq. (9), penalizes for variation in $t(\mathbf{x})$ according to the smoothness modulus of $I(\mathbf{x})$, i.e., the lower $\|I(\mathbf{x}) - I(\mathbf{y})\|^2$ is, the stronger the requirement for low $(t(\mathbf{x}) - t(\mathbf{y}))^2$ becomes. This requirement follows from the fact that according to the haze formation model in Eq. (1), spatial variations in both $t(\mathbf{x})$ and $J(\mathbf{x})$ produce variations in $I(\mathbf{x})$. Hence, the smoothness of $I(\mathbf{x})$ can be used as an upper-bound for that of $t(\mathbf{x})$. In summary, this regularization term allows the transmission map to exhibit sharp profiles along edges in the input image and requires it to be smooth where the input is smooth.

Note that the competition between the smoothness and data terms is strong only at pixels where a reliable transmission estimate is available (small σ_t). This competition gets weaker where the estimates are less reliable and it vanishes where no estimates are available, in which case the MRF acts as a pure interpolation mechanism.

Maximizing P is done by minimizing the quadratic form $-\log P$ which boils down to solving a linear system consisting of a sparse Laplacian matrix with strictly negative off-diagonal elements (known as M-matrix). In contrast the matting Laplacian, used by He et al. [2009] for regularizing the transmission, is known to achieve less smoothing; see the discussion in [Levin et al. 2008]. In principle, this behavior follows from the fact that the matting Laplacian is derived under the assumption of linear relation between the transmission (alpha-channel in the original context) and the input pixels $I(\mathbf{x})$, meaning that small variations in the latter will induce variations in the former. Images are intrinsically more content-rich compared to transmission maps, mainly due to changes in the surface shading and albedo. Attributing these variations to the transmission leads to their unwanted reduction in the dehazed image J . Figure 6 shows the contrast reduction created by using the matting Laplacian for regularization.

The regularization term in Eq. (9) couples nearby pixels and is responsible for the interpolation of the transmission to pixels \mathbf{x} lacking their own estimate, $\hat{t}(\mathbf{x})$. However, occasionally there are islands of strongly-connected pixels which are *weakly* connected to their surrounding pixels due to color mismatch, i.e., large $\|I(\mathbf{x}) - I(\mathbf{y})\|^2$ in the denominator of the regularization term in Eq. (9). This scenario takes place between pixels of distinct objects. In case no transmission estimate exists inside the island, its pixels may receive irrelevant values from their surrounding pixels which correspond to a different object in the scene. We avoid these

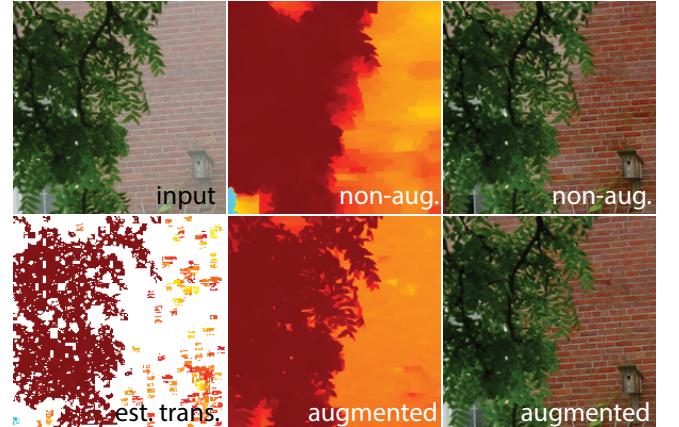


Fig. 8. Interpolation and regularization using standard and augmented Markov random fields. Top-left image is the input and below it is the estimated transmission. The second column shows the resulting transmission maps with (bottom) and without (top) augmenting the model with long-range connections. Right column shows the resulting dehazed images.

wrong assignments by searching for similar pixels within a wider perimeter and augmenting $N_{\mathbf{x}}$ with these additional coordinates. This augmented GMRF is illustrated in Figure 7.

In our implementation, we find these connections by randomly sampling pixels \mathbf{y} inside windows whose size is 15% of the image size and once we find a pixel \mathbf{y} such that $\|I(\mathbf{x}) - I(\mathbf{y})\| < 0.1$ we stop the search and add it to $N_{\mathbf{x}}$. For efficiency reasons we stop the search after five unsuccessful attempts and limit this augmentation to a subsampled grid of every fourth pixel in each image axis. Hence, this process increases the number of connections by a small factor of 1/64 and increases the GMRF construction and solve time by less than 25%. Note that since we do not perform a complete search within these windows but use few random samples, this procedure does not undermine the overall linear running time of our algorithm. We note that the use of long-range connections was explored in the context of image denoising for capturing high-order relations efficiently in [Li and Huttenlocher 2008].

Finding a small number of long-range connections is enough to resolve all the island's pixels due to their strong inner connectivity and weak dependency on the surrounding. Figure 8 shows how the transmission in regions surrounded by tree leafs is resolved better by the augmented GMRF.

4. RESULTS

We report here the evaluation of our method over a large dataset of over 40 images that includes the benchmark images used by Tan [2008], Fattal [2008], Kopf et al. [2008], He et al. [2009], Tarel and Hautière [2009], Kratz and Nishino [2009], Nishino et al. [2012] and Gibson and Nguyen [2013a; 2013b] to evaluate their methods. All the tests shown in the paper as well as many other can be found in the supplemental material¹. We strongly encourage the reader to explore this in-depth comparison.

The images generated by our method were produced by the same set of parameters quoted in the previous sections. The thresholds were determined by a learning procedure in which we searched

¹The supplemental material containing all the tests and comparisons made can be found at: http://www.cs.huji.ac.il/~raananf/projects/dehaze_cl

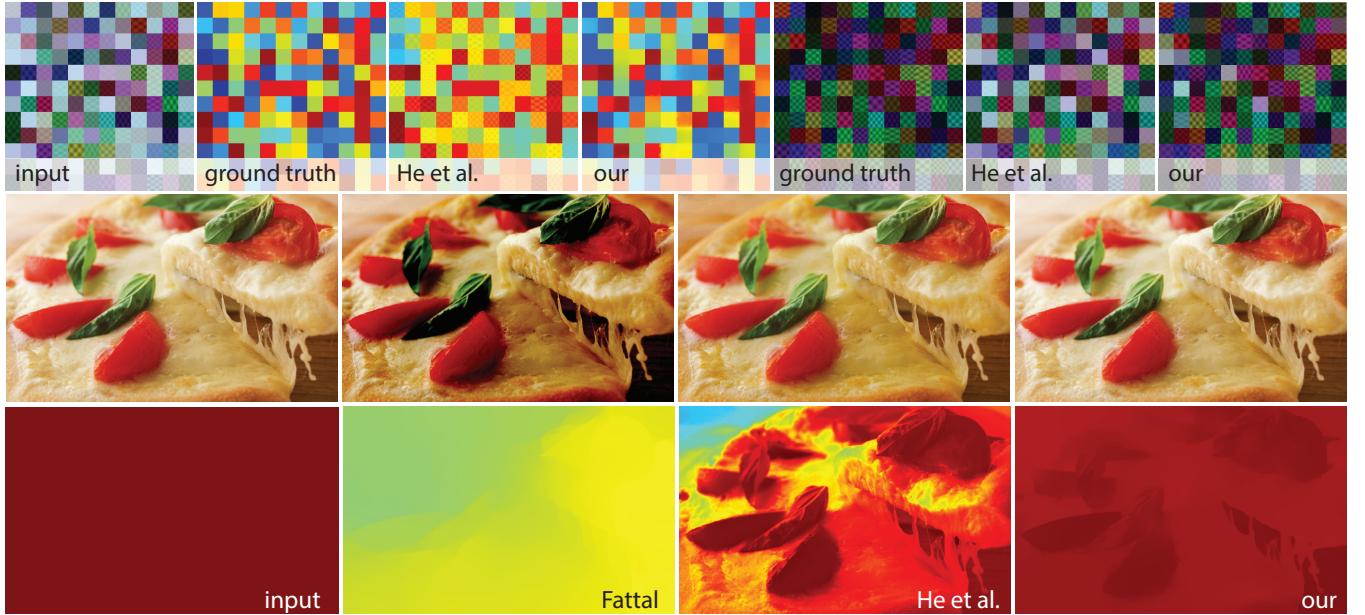


Fig. 9. Images with known ground-truth. Top-row shows a comparison between He et al. [2009] and our method over an artificial dead-leaves-like image obeying the dark-channel property. Middle-row shows the results by Fattal [2008], He et al. [2009] and our methods over a haze-free image. The corresponding transmission maps are shown at the bottom row (red indicates $t = 1$). In this test the output radiance and transmission should be identical to the input.

for the optimal values that achieve the highest accuracy over a set of three images with known ground-truth transmission (Road1, Flower1, and Lawn1). We used the fixed value of $\sigma = 1/30$ to produce all our dehazed images even though they arrived from multiple sources with unknown noise level. Finally, we applied our method with the atmospheric light vectors A used by others (depending on the source of the image), and when unavailable we recovered this value by manually selecting the haziest pixel in the image. The values of the atmospheric light vectors A that we used are specified in the supplemental material.

Qualitative comparison. Figure 14 (last page) shows a number of the comparisons we made against state-of-the-art methods where several trends can be pointed out. The method of Fattal [2008] produces results of variable quality, suffering from occasional severe over- and under-estimations in the transmission. This can be attributed to its inability to validate its assumptions and its limited operation across the image due to a conservative signal-to-noise criterion. These failures are seen in the Red Bricks House image where it over corrects the red bricks and under corrects the grass as well as in the false variations it produces in the Stadium image (see supp. mat.). Moreover, this approach shows a limited ability to dehaze distant regions in the Wheat Field, Arial, New York and Manhattan images (see the supplemental material). A severe over-correction is seen in the Mountain image shown in Figure 4.

As pointed out by He et al. [2009], while the method of Tan [2008] removes haze robustly, it also tends to underestimate the transmission and produce over-saturated results, see for example the Manhattan and Red Bricks House images. A somewhat similar behavior is seen in the Red Bricks House and Swan images dehazed by the method of Nishino et al. [2012], which over corrects the bricks and swans.

The method of He et al. [2009] is known for its robustness. However, in regions where no color channel vanishes it underestimates the transmission and also produces over-corrected results, as seen

Table I. Accuracy comparison with known ground-truth.

	Fattal [2008]	He et al. [2009]	our
Squares	0.083 / 0.097	0.11 / 0.15	0.03 / 0.06
DC Sqrs.	0.056 / 0.061	0.115 / 0.17	0.025 / 0.05
Pizza	0.42 / 0.21	0.164 / 0.073	0.0255 / 0.012
Fruit	0.171 / 0.064	0.011 / 0.016	0.0025 / 0.003

Table reports the mean L_1 error of the estimated transmission t (left value) and output image J (right value).

in Figure 4. As discussed in Section 3.5, the matting Laplacian regularization that He et al. [2009] use transfers some of the fine image detail into the transmission. This leads to an overall reduction of contrast in $J(x)$ which can be observed at the distant regions of the Cityscape, Hong Kong, Manhattan, Snow Mountain and Wheat Field images as well as in the Logos and Red Bricks House. In the supplemental material we compare between the transmission maps generated by the different methods.

Finally, the method of Kopf et al. [2008] produces well-balanced results with some under performance at heavily hazed regions. We should note however that unlike the rest of the methods mentioned here, this method requires a user-aligned scene geometry.

The results produced by our method are not free from errors. However, they do not appear as severe as some of the errors produced by Fattal [2008] and, in general, show higher contrast and clarity compared to He et al. [2009]. Similarly to the rest of the methods, our method has a limited effectiveness at regions of very low visibility such as in the case of Staten Island seen in the Manhattan image. The amplification of noise at these regions is another noticeable drawback. However, in most cases it compares favorably to the alternatives in this respect.

Quantitative comparison. In order to quantitatively evaluate the performance of our method we tested it over different types of images in which the transmission is known. In the first test we synthesized artificial scenes composed of distinct squares where we



Fig. 10. Real-world images with known transmission. The input image (not shown here) was synthesized using the bottom-right image and the transmission map above it. The results produced by Fattal [2008], He et al. [2009] and our methods are shown to the right.

randomly sampled the reflectance coefficients, illumination function and a constant transmission value and plugged these values in Eq. (4) to simulate haze. The resulting images are meant to mimic the way natural images images are modeled by Lee et al. [2001]. We used this procedure twice and in the second image we generated (the DC Squares) we made sure that, when sampling the reflectance values, at least one channel is set to zero in order to meet the dark-channel prior as well. The images produced in this test are shown in Figure 9 as well as the results obtained by the method of He et al. [2009] and our method. The L_1 errors produced on both images (with and without the dark-channel constraint) are reported in Table I.

In another test we applied different dehazing methods over lucid, haze-free, images in which case we expect $t(\mathbf{x}) = 1$ to be the solution. Figure 9 shows one of these images and Table I provides the errors produced by the methods of He et al. [2009], Fattal [2008] and ours. In both tests our method outperforms the competing techniques. All the images participating in the tests described in this section can be found at the supplemental material.

In order to obtain a more realistic evaluation we synthesized hazy images of natural scenes using pairs of real-world photographs and their corresponding depth maps. By assuming the media scattering coefficient β is constant in space, we obtain the transmission from Eq. (3) by $t(\mathbf{x}) = e^{-\beta d(\mathbf{x})}$, where $d(\mathbf{x})$ is the optical depth at each pixel \mathbf{x} . Note that the resulting transmission maps are not constant in image space and exhibit non-trivial variations along depth discontinuities. We produced 12 such test images using the depth maps found in [Scharstein and Szeliski 2002; Zhang et al. 2009; Kim et al. 2013] and used them to compare our method with the methods of He et al. [2009], Fattal [2008]. Figure 10 shows the results obtained over one of these test images. Table II summarizes the L_1 errors in the estimated transmission and dehazed image $J(\mathbf{x})$ produced by the different methods. In this test our method achieves the highest accuracy and Fattal [2008] is the least accurate method on many images.

We further used these images for gathering the statistics reported in Figure 2, as well as for studying the sensitivity of the three meth-

ods to the level of noise and the thickness of the haze present in the image.

Table III reports the errors obtained over sequences of images produced with an increasing level of scattering coefficient (three levels of β differing by a factor of 3). As β increases and the haze becomes thicker, the method of Fattal [2008] loses accuracy both in its transmission estimate and dehazed output $J(\mathbf{x})$. In contrast, the method of He et al. [2009] and our method estimate the transmission more accurately at higher β values, yet similarly to Fattal [2008]'s method both methods recover $J(\mathbf{x})$ less accurately. The increase in the transmission accuracy can be explained by the reduction in the contribution of the direct transmission, $t(\mathbf{x})J(\mathbf{x})$ in Eq. (1). The latter is the (sole) component in which inaccuracies in the dark-channel assumption can appear. In case of our method, pixels of heavily-hazed patches cluster closer to the atmospheric light line, sA , and hence the intersection point between this line and the patch color-line is less sensitive to errors in the recovered color-line orientation vector D . Nevertheless, in both cases the increased accuracy of $t(\mathbf{x})$ does not lead to higher accuracy in the dehazed image $J(\mathbf{x})$. This follows from the more extreme correction involved in removing thick layers of haze when extracting $J(\mathbf{x})$ from Eq. (1).

In order to assess the influence of noise, we added an identically-distributed zero-mean Gaussian noise to each color channel of each image pixel independently. This test was conducted with three different noise levels, $\sigma = 0.01, 0.025$ and 0.05 . Table IV reports the errors obtained by the three methods as well as by using the true transmission to dehaze the noisy input image. While our method obtains more accurate transmission estimates in case of low σ , the method of Fattal [2008] achieves lower errors at high noise levels. Figure 11 shows one of the images used in this test with $\sigma = 0.05$ where our method managed to achieve stronger dehazing in the farther regions of the scene. However, there are regions in this image where our method under-estimated the transmission and, by subtracting the blueish haze, resulted in unnatural yellowish output (such as in the case of the distance trees).

Table II. Accuracy comparison over real-world images with known transmission.

	Fattal [2008]	He et al. [2009]	our
Road1	0.319 / 0.078	0.097 / 0.032	0.069 / 0.020
Road2	0.347 / 0.096	0.086 / 0.026	0.061 / 0.019
Flower1	0.089 / 0.017	0.190 / 0.065	0.047 / 0.012
Flower2	0.074 / 0.013	0.203 / 0.058	0.042 / 0.009
Lawn1	0.317 / 0.053	0.118 / 0.030	0.078 / 0.015
Lawn2	0.323 / 0.061	0.115 / 0.034	0.064 / 0.015
Mansion	0.147 / 0.044	0.074 / 0.030	0.042 / 0.015
Church	0.377 / 0.105	0.070 / 0.033	0.038 / 0.018
Couch	0.089 / 0.020	0.069 / 0.019	0.089 / 0.019
Dolls	0.043 / 0.068	0.036 / 0.055	0.031 / 0.046
Moebius	0.111 / 0.027	0.235 / 0.091	0.145 / 0.047
Reindeer	0.070 / 0.018	0.126 / 0.043	0.066 / 0.015

Table reports the mean L_1 error of the estimated transmission t (left value) and output image J (right value).

Table III. Sensitivity to scattering level over real-world images with known transmission.

scattering	Fattal [2008]	He [2009]	our
Road1			
low	0.083 / 0.019	0.122 / 0.029	0.075 / 0.017
medium	0.319 / 0.078	0.097 / 0.032	0.070 / 0.020
high	0.604 / 0.150	0.055 / 0.039	0.043 / 0.024
Lawn1			
low	0.104 / 0.019	0.158 / 0.030	0.040 / 0.009
medium	0.317 / 0.053	0.118 / 0.030	0.076 / 0.015
high	0.442 / 0.090	0.064 / 0.034	0.050 / 0.017
Mansion			
low	0.033 / 0.009	0.108 / 0.031	0.040 / 0.010
medium	0.147 / 0.044	0.074 / 0.030	0.042 / 0.015
high	0.533 / 0.153	0.039 / 0.031	0.029 / 0.022
Church			
low	0.079 / 0.022	0.148 / 0.039	0.045 / 0.013
medium	0.377 / 0.105	0.070 / 0.033	0.036 / 0.017
high	0.771 / 0.193	0.027 / 0.032	0.023 / 0.027
Reindeer			
low	0.018 / 0.004	0.150 / 0.042	0.057 / 0.013
medium	0.070 / 0.018	0.126 / 0.043	0.067 / 0.016
high	0.303 / 0.082	0.072 / 0.044	0.053 / 0.023

Table reports the mean L_1 error of the estimated transmission t (left value) and output image J (right value).

Two color channel images. Finally, while we derived our method for three color-channel images, most of the derivation holds for two-channel images including the line intersection formula in Eq. (5). The lack-of-intersection criterion, however, trivializes as every two non-parallel lines intersect in two-dimensional space. Figure 12 shows the result obtained when we evaluate the transmission based on two channels by dropping the red channel of the Hong Kong image. While there is some over-estimation in the recovered transmission, the method remains effective for two color channel images.

Running times. We implemented our method in C and run it on a 2.6 GHz computer (running on a single core). Estimating the transmission in a one mega-pixel image takes us 0.4 seconds and constructing and solving the GMRF takes another 5 seconds. He et al. [2009]’s method requires 10 to 20 seconds to process a 600×400 pixel image on a 3.0 GHz machine. We attribute these longer running times to the construction and solution of the matting Laplacian, which unlike our Laplacian, its entries are computed based on



Fig. 11. Dehazing a noisy image. Image produced with i.i.d. zero-mean Gaussian noise with $\sigma = 0.05$. At the left column (top to bottom) are the input image, the haze-free image used to produce this test, and the result obtained by dehazing the noisy image using the true transmission. At the right column we see the results of Fattal [2008], He et al [2009] and ours.

patches rather than individual pixels. Moreover, this matrix is not an M-matrix which makes it harder to solve.

He et al. [2009] as well as Gibson and Nguyen [2013b] reduce their running time to 0.3 seconds per one mega-pixel image using the guided and bilateral image filtering. The edge-avoiding wavelets [Fattal 2009] was shown to accelerate edge-aware interpolation problems with scattered data such as our partial transmission maps. We used this method successfully to compute the transmission and reach an overall running time of 0.55 seconds per one mega-pixel image (0.15 seconds for the interpolation). At the supplemental material we provide several comparisons between the different smoothing methods. While solving the Laplacian system achieves a greater accuracy (mostly on low-resolution images), the tests show that in many cases negligible visual differences are observed. All the time quotes mentioned here grow linearly with the image dimension.

5. CONCLUSIONS

We presented a new single-image dehazing method based on the color-lines pixel regularity in natural images. We derived a local formation model that reasons this regularity in hazy scenes and described how it is used for estimating the scene transmission. Unlike existing dehazing methods that follow their assumptions across the entire image, the new formation model allows us to dismiss parts of the image that violate the underlining assumptions and achieve higher overall accuracy. We also proposed an augmented GMRF model with long-range coupling in order to better resolve the transmission in isolated pixels that lack their own estimate. Finally, we reported the results of an extensive evaluation of the algorithms on different types of problems that demonstrate its high accuracy. Besides practical contributions, at the theoretical level of image un-

Table IV. Sensitivity to noise over real-world images with known transmission.

σ	True t	Fattal [2008]	He [2009]	our
Road1				
0.01	0.015	0.198 / 0.062	0.100 / 0.039	0.068 / 0.027
0.025	0.038	0.127 / 0.055	0.106 / 0.056	0.083 / 0.053
0.05	0.074	0.114 / 0.074	0.136 / 0.084	0.120 / 0.095
Lawn1				
0.01	0.017	0.234 / 0.047	0.115 / 0.037	0.056 / 0.023
0.025	0.043	0.136 / 0.046	0.108 / 0.054	0.056 / 0.054
0.05	0.084	0.085 / 0.072	0.115 / 0.080	0.114 / 0.100
Mansion				
0.01	0.012	0.101 / 0.036	0.067 / 0.030	0.048 / 0.022
0.025	0.030	0.077 / 0.039	0.057 / 0.040	0.062 / 0.041
0.05	0.059	0.077 / 0.062	0.083 / 0.063	0.081 / 0.070
Church				
0.01	0.017	0.223 / 0.079	0.067 / 0.038	0.052 / 0.033
0.025	0.042	0.155 / 0.071	0.058 / 0.052	0.085 / 0.062
0.05	0.081	0.153 / 0.092	0.088 / 0.082	0.121 / 0.104
Reindeer				
0.01	0.011	0.059 / 0.020	0.119 / 0.044	0.077 / 0.026
0.025	0.026	0.053 / 0.032	0.109 / 0.049	0.083 / 0.041
0.05	0.053	0.058 / 0.057	0.117 / 0.063	0.106 / 0.069

Table reports the mean L_1 error of the estimated transmission t (left value) and output image J (right value). Leftmost column provides the accuracy obtained when dehazing the image using the known transmission.

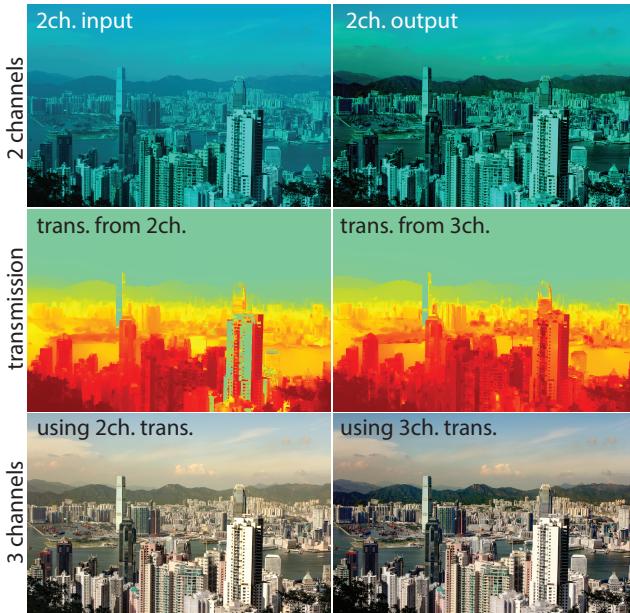


Fig. 12. Operating on two-color channel images. Top-left is the input with no red channel. To its right is the recovered two-channel image. Middle row compares the transmission estimated from the two-color images and three-color (original Hong Kong image from Figure 14) images. Below are the dehazed three-channel Hong Kong image using these maps.

derstanding, this work supports the relevance of dead-leaves type of models to hazy natural scenes.

Limitations. Our method relies on specific assumptions based on which we derive Eq. (4). While we propose a list of conditions for identifying patches that do not obey Eq. (4), this list is not sufficient to guarantee a correct classification. As an example,



Fig. 13. Highly specular and artificially-lit scene. Underestimated transmission is seen along the shore of lit buildings.

Figure 13 shows a night scene with many artificial colored lights and specular highlights. The transmission estimated in this scene is severely underestimated across the shore of lit buildings which is over-corrected by our method. Furthermore, even when classifying patches correctly we may still obtain too few estimates across the image. We should note however that our reported evaluation demonstrates that the color-line assumption is, in general, a reliable and competitive prior for hazy scenes.

While our method achieves higher accuracy in low noise levels ($\sigma < 0.01$), Table IV shows that at high noise levels $\sigma \geq 0.05$, our method becomes less accurate than competing approaches.

Figure 13 shows another difficult problem, shared by other dehazing techniques, which is the treatment the sky receive. In many cases the atmospheric light is very close to the sky color and hence the latter is wrongly treated as a thick layer of haze. As future work we intend to address this problem both by detecting the sky as well as devising a proper treatment for it. Another limiting factor that we intend to address is the inability to cope with spatially-varying atmospheric light vector, i.e., $A = A(\mathbf{x})$.

Finally, unlike the methods of He et al. [2009] and Tarel and Hautière [2009] our method cannot operate on mono-chromatic images where the notion of color-lines trivializes.

Appendix

Here we analyze the dependency of the error in the estimated transmission on the angle between the patch-line orientation D and the atmospheric light vector A . The transmission is recovered by minimizing Eq. (5), which boils down to solving the following system

$$\begin{bmatrix} \|D\|^2 & -\langle A, D \rangle \\ -\langle A, D \rangle & \|A\|^2 \end{bmatrix} \begin{bmatrix} l \\ s \end{bmatrix} = \begin{bmatrix} -\langle D, V \rangle \\ \langle A, V \rangle \end{bmatrix}. \quad (10)$$

Since $\|D\|$ is chosen arbitrarily let us assume that $\|D\| = \|A\|$ and hence, with no loss of generality, let us further assume the two are $\|D\| = \|A\| = 1$. In this case, the solution for Eq. (10) is given by

$$\begin{bmatrix} l \\ s \end{bmatrix} = \frac{1}{1 - \langle D, A \rangle^2} \begin{bmatrix} 1 & \langle A, D \rangle \\ \langle A, D \rangle & 1 \end{bmatrix} \begin{bmatrix} -\langle D, V \rangle \\ \langle A, V \rangle \end{bmatrix} \quad (11)$$

Now let us denote the error in the estimated line offset vector by E , i.e., $V = (1-t)A + E$. In which case the estimated transmission, $\hat{t} = 1 - s$, is given by

$$\begin{aligned} 1 - \frac{\langle A, (1-t)A + E \rangle - \langle D, (1-t)A + E \rangle \langle D, A \rangle}{1 - \langle D, A \rangle^2} \\ = 1 - (1-t) \frac{1 - \langle D, A \rangle^2}{1 - \langle D, A \rangle^2} + \frac{\langle A, E \rangle - \langle D, E \rangle \langle D, A \rangle}{1 - \langle D, A \rangle^2}, \end{aligned} \quad (12)$$

where the terms besides the last reduce to the true transmission t and the last term corresponds to the estimation error. Note that if $E = \mathbf{0}$ then this error vanishes, meaning that the line may have an arbitrary orientation D and yet the exact transmission t will be

recovered. This follows from the fact that we recover the transmission based on the patch-line's offset from the origin.

Having assumed that $\|A\| = \|D\| = 1$ the similarity between the orientation of the two can be measured by the length of $\Delta = A - D$. Thus, the error term in Eq. (12) becomes

$$\begin{aligned} & \frac{\langle D, E \rangle + \langle \Delta, E \rangle - \langle D, E \rangle - \langle D, E \rangle \langle D, \Delta \rangle}{1 - (1 + \langle D, \Delta \rangle)^2} \\ &= \frac{\langle \Delta, E \rangle - \langle D, E \rangle \langle D, \Delta \rangle}{-2\langle D, \Delta \rangle - \langle D, \Delta \rangle^2}. \end{aligned} \quad (13)$$

Now since

$$1 = \|A\|^2 = \|D + \Delta\|^2 = \|D\|^2 + 2\langle D, \Delta \rangle + \|\Delta\|^2 = 1 + 2\langle D, \Delta \rangle + \|\Delta\|^2, \quad (14)$$

we get $\langle D, \Delta \rangle = O(\|\Delta\|^2)$, and therefore the transmission error in Eq. (13) is approximately

$$\frac{O(\|\Delta\|) - \langle D, E \rangle O(\|\Delta\|^2)}{O(\|\Delta\|^2) - O(\|\Delta\|^4)} = O(\|\Delta\|^{-1}). \quad (15)$$

Finally, since

$$\|\Delta\|^2 = \|A - D\|^2 = \|A\|^2 - 2\langle D, A \rangle + \|D\|^2 = 2 + 2\cos(\theta) \approx \theta^2, \quad (16)$$

for small angle θ between the D and A , we conclude that the error in the transmission grows like $O(\theta^{-1})$. Figure 15 shows a numerical simulation, where we synthesized patches with color-lines that form different angles with A and added Gaussian noise with $\sigma = 0.01$. The graphs confirm the prediction of our analysis, namely, that the transmission \hat{t} estimated from Eq. (12) obeys $\text{Var}[\hat{t}]^{-1} = \theta^2$.

In practice, we use this transmission estimate to define the Gaussian Markov random field model in Eq. (9) from which we obtain a complete regularized transmission map. In this model we specify the confidence in the estimated values based on the relation between A and D in the corresponding patch. This score is derived by modeling the patch-line error E as a zero-mean Gaussian variable and, since it appears in linear form in the transmission error term (last term in Eq. (12)), we get a zero-mean Gaussian noise in the estimated transmission. More specifically, by rewriting its numerator as $\langle A - D\langle D, A \rangle, E \rangle$ we obtain the following standard deviation in the estimated transmission

$$\sigma \frac{\|A - D\langle D, A \rangle\|}{1 - \langle D, A \rangle^2} \quad (17)$$

which we plug in Eq. (9), where σ is the standard-deviation of E .

REFERENCES

- CARR, P. AND HARTLEY, R. 2009. Improved single image dehazing using geometry. In *Digital Image Computing: Techniques and Applications, 2009. DICTA '09*. 103–110.
- CHAVEZ, P. S. 1988. An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data. *Remote Sensing of Environment* 24, 450–479.
- DU, Y., GUINDON, B., AND CIHLAR, J. 2002. Haze detection and removal in high resolution satellite image with wavelet analysis. *IEEE Transactions on Geoscience and Remote Sensing* 40, 1, 210–217.
- FATTAL, R. 2008. Single image dehazing. In *ACM SIGGRAPH papers. SIGGRAPH '08*. ACM, New York, NY, USA, 72:1–72:9.
- FATTAL, R. 2009. Edge-avoiding wavelets and their applications. In *ACM SIGGRAPH 2009 papers*. ACM, New York, NY, USA, 22:1–22:10.
- GIBSON, K. AND NGUYEN, T. 2011. On the effectiveness of the dark channel prior for single image dehazing by approximating with minimum volume ellipsoids. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 1253–1256.
- GIBSON, K. AND NGUYEN, T. 2013a. An analysis of single image defogging methods using a color ellipsoid framework. *EURASIP Journal on Image and Video Processing* 2013, 1, 37.
- GIBSON, K. AND NGUYEN, T. 2013b. Fast single image fog removal using the adaptive wiener filter. *International Conference on Image Processing*.
- HE, K., SUN, J., AND TANG, X. 2009. single image haze removal using dark channel prior. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 1956–1963.
- HINDS, W. 1999. *Aerosol technology: properties, behavior, and measurement of airborne particles*. Wiley-Interscience. Wiley.
- HUANG, J., LEE, A., AND MUMFORD, D. 2000. Statistics of range images. *Proceedings. IEEE Conference on Computer Vision and Pattern Recognition*, 324–331.
- HUANG, J. AND MUMFORD, D. 1999. Statistics of natural images and models. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 1*, 637–663.
- KIM, C., ZIMMER, H., PRITCH, Y., SORKINE-HORNUNG, A., AND GROSS, M. 2013. Scene reconstruction from high spatio-angular resolution light fields. *ACM SIGGRAPH 2009 papers* 32, 4, 73:1–73:12.
- KOPF, J., NEUBERT, B., CHEN, B., COHEN, M., COHEN-OR, D., DEUSSEN, O., UYTTENDAELE, M., AND LISCHINSKI, D. 2008. Deep photo: model-based photograph enhancement and viewing. In *ACM SIGGRAPH Asia papers*. ACM, New York, NY, USA, 116:1–116:10.
- KOSCHMIEDER, H. 1924. Theorie der horizontalen sichtweite. In *Beitr. zur Phys. d. freien Atm.* 171–181.
- KRATZ, L. AND NISHINO, K. 2009. Factorizing scene albedo and depth from a single foggy image. In *IEEE International Conference on Computer Vision (ICCV)*. 1701–1708.
- LEE, A. B., MUMFORD, D., AND HUANG, J. 2001. Occlusion models for natural images: A statistical study of a scale-invariant dead leaves model. *International Journal of Computer Vision* 41, 35–59.

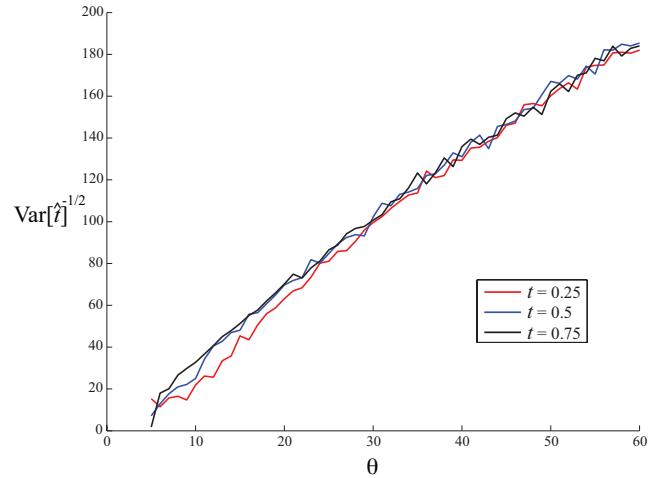


Fig. 15. Graph showing the variance of the transmission estimated from 1000 noisy patches with $\sigma = 0.01$. The variance is shown as function of the angle θ between the patch color-line and A . This test is repeated from three different transmission level and in all cases a linear dependence is observed between θ^{-1} and the error's standard-deviation.

- LEVIN, A., LISCHINSKI, D., AND WEISS, Y. 2008. A closed-form solution to natural image matting. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 30, 2, 228–242.
- LI, Y. AND HUTTENLOCHER, D. P. 2008. Sparse long-range random field and its application to image denoising. In *European Conference on Computer Vision*. Springer-Verlag, Berlin, Heidelberg, 344–357.
- MATLIN, E. AND MILANFAR, P. 2012. Removal of haze and noise from a single image. *Proc. SPIE* 8296, 82960T–82960T–12.
- NARASIMHAN, S. AND NAYAR, S. 2000. Chromatic framework for vision in bad weather. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* 1, 598–605.
- NARASIMHAN, S. G. AND NAYAR, S. K. 2003. Interactive (De)weathering of an Image using Physical Models. In *ICCV Workshop on Color and Photometric Methods in Computer Vision*.
- NAYAR, S. K. AND NARASIMHAN, S. G. 1999. Vision in bad weather. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 820.
- NISHINO, K., KRATZ, L., AND LOMBARDI, S. 2012. Bayesian defogging. *International Journal of Computer Vision* 98, 3, 263–278.
- OAKLEY, J. P. AND BU, H. 2007. Correction of simple contrast loss in color images. *IEEE Transactions on Image Processing* 16, 2, 511–522.
- OMER, I. AND WERMAN, M. 2004. Color lines: image specific color representation. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 946–953.
- ROTH, S. AND BLACK, M. J. 2007. On the spatial statistics of optical flow. *International Journal of Computer Vision* 74, 1, 33–50.
- SCHARSTEIN, D. AND SZELISKI, R. 2002. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *International Journal of Computer Vision* 47, 1-3 (Apr.), 7–42.
- SCHECHNER, Y., NARASIMHAN, S. G., AND NAYAR, S. K. 2001. Instant dehazing of images using polarization. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 325–332.
- SHWARTZ, S., NAMER, E., Y., Y., AND SCHECHNER. 2006. Blind haze separation. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 1984–1991.
- TAN, K. AND OAKLEY, J. P. 2000. Enhancement of color images in poor visibility conditions. *Proceedings of International Conference on Image Processing* 2, 788–791.
- TAN, R. T. 2008. Visibility in bad weather from a single image. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.
- TAREL, J.-P. AND HAUTIÈRE, N. 2009. Fast visibility restoration from a single color or gray level image. In *IEEE International Conference on Computer Vision (ICCV'09)*. Kyoto, Japan, 2201–2208.
- ZHANG, G., JIA, J., WONG, T.-T., AND BAO, H. 2009. Consistent depth maps recovery from a video sequence. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 31, 6, 974–988.
- ZHANG, Y., GUINDON, B., AND CIHLAR, J. 2002. An image transform to characterize and compensate for spatial variations in thin cloud contamination of landsat images. *Remote Sensing of Environment* 82, 173–187.

Received September 2008; accepted March 2009

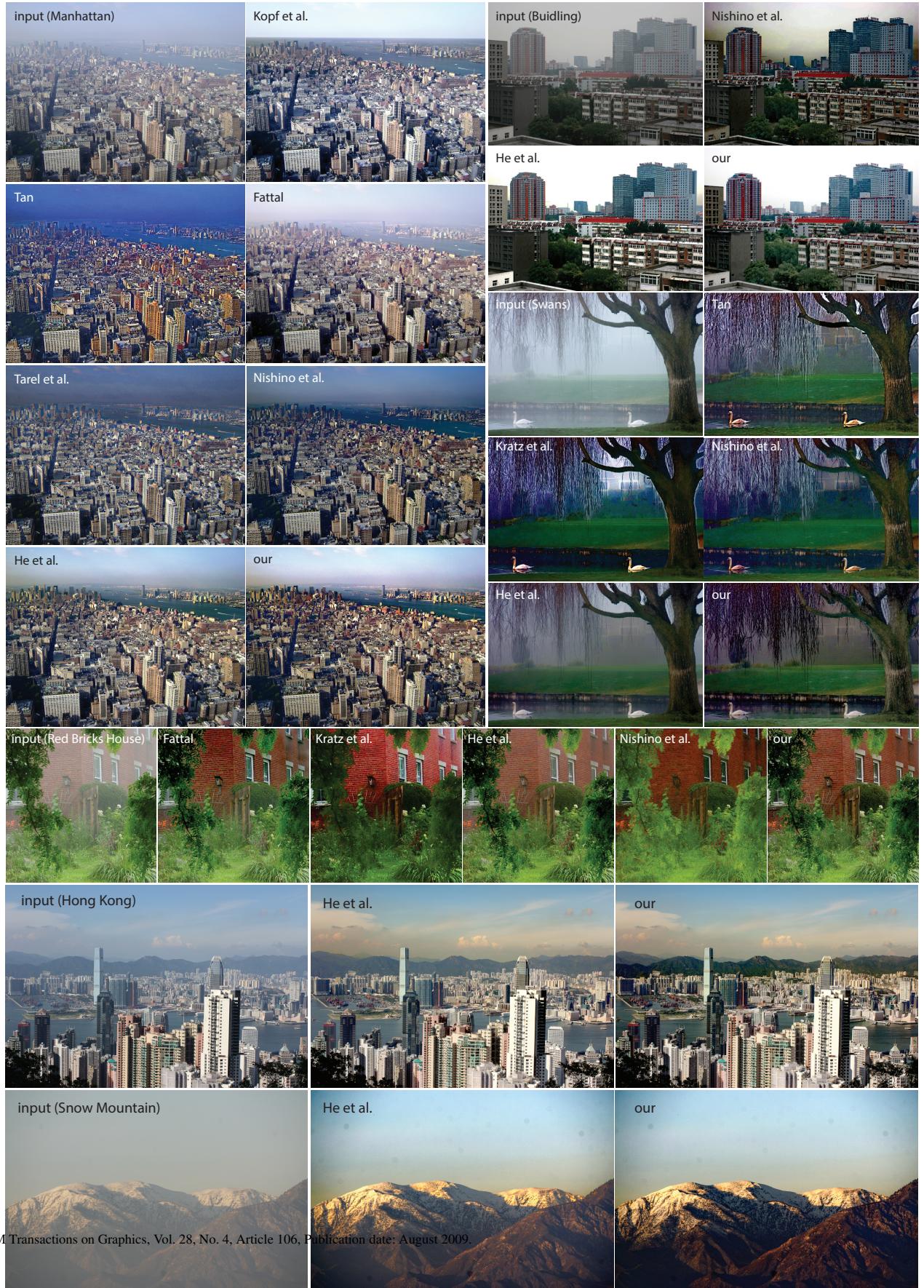


Fig. 14. Comparison with state-of-the-art. Top-left is the Manhattan image, top-right the Buildings image, below it the Swans. Third- row from the bottom is the Red Bricks House, Hong Kong below it and the Snow Mountain image is shown at the bottom.