Non-local Haze Propagation with an Iso-Depth Prior

Incheol Kim and Min H. Kim(✉)

KAIST, 291 Daehak-ro, Yuseong-gu, Daejeon 34141, Korea
minhkim@vclab.kaist.ac.kr

Abstract. The primary challenge for removing haze from a single image is lack of decomposition cues between the original light transport and airlight scattering in a scene. Many dehazing algorithms start from an assumption on natural image statistics to estimate airlight from sparse cues. The sparsely estimated airlight cues need to be propagated according to the local density of airlight in the form of a transmission map, which allows us to obtain a haze-free image by subtracting airlight from the hazy input. Traditional airlight-propagation methods rely on ordinary regularization on a grid random field, which often results in isolated haze artifacts when they fail in estimating local density of airlight properly. In this work, we propose a non-local regularization method for dehazing by combining Markov random fields (MRFs) with nearest-neighbor fields (NNFs) extracted from the hazy input using the PatchMatch algorithm. Our method starts from the insightful observation that the extracted NNFs can associate pixels at the similar depth. Since regional haze in the atmosphere is correlated with its depth, we can allow propagation across the iso-depth pixels with the MRF-based regularization problem with the NNFs. Our results validate how our method can restore a wide range of hazy images of natural landscape clearly without suffering from haze isolation artifacts. Also, our regularization method is directly applicable to various dehazing methods.

Keywords: Dehazing · Non-local regularization · Image restoration

1 Introduction

The atmosphere in a landscape includes several types of aerosols such as haze, dust, or fog. When we capture a landscape photograph of a scene, often thick

The initial version of this manuscript was published in conference proceedings of International Conference on Computer Vision, Theory and Applications (VISAPP 2017) [1]. This paper was invited to be published in the Communications in Computer and Information Science series by Springer afterward. It is a revised and extended version of the conference paper to additionally introduce more scientific analysis and evaluation on the proposed method.

Electronic supplementary material The online version of this chapter (https://doi.org/10.1007/978-3-030-12209-6_11) contains supplementary material, which is available to authorized users.

A. P. Cláudio et al. (Eds.): VISIGRAPP 2017, CCIS 983, pp. 213–238, 2019.
https://doi.org/10.1007/978-3-030-12209-6_11

aerosols scatter light transport from the scene to the camera, resulting in a hazy photograph. A haze-free image could be restored if we could estimate and compensate the amount of scattered energy properly. However, estimating haze from a single photograph is a severely ill-posed problem due to the lack of the scene information such as depth.

An image processing technique that removes a layer of haze and compensates the attenuated energy is known as *dehazing*. It can be applied to many outdoor imaging applications such as self-driving vehicles, surveillance, and satellite imaging. The general dehazing algorithm consists of two main processes. We first need to approximate haze initially by utilizing available haze clues based on a certain assumption on natural image statistics, such as a dark channel prior [2]. In this stage, most of dehazing algorithms tend to produce an incomplete transmission map from the hazy image. Once we obtain rough approximation of haze, we need to propagate the sparse information to the entire scene to reconstruct a dense transmittance map used for recovering a haze-free image.

Difficulty of dehazing arises from the existence of ambiguity due to the lack of the scene information. First, the initial assumption on image statistics on natural colors in particular is insufficient to cover the wide diversity of natural scenes in the real world, resulting in *incomplete* haze estimation. No universal image statistics on natural colors can handle the dehazing problem. Moreover, as shown in Fig. 1, state-of-the-art propagation algorithms with a common grid random field often suffer from *haze-isolation artifacts* [3–5]. Meanwhile, the amount of haze in the atmosphere at each pixel is determined by its depth. In order to handle abrupt changes of haze density, we need the scene depth information, even though it is unavailable in single-image dehazing.

In this paper, we propose a non-local regularization method for dehazing that can propagate sparse estimation of airlight to yield a dense transmission map without suffering from the typical isolation problem (Fig. 1). Our regularization approach is developed by combining Markov random fields (MRFs) with nearest-neighbor fields (NNFs) using the PatchMatch algorithm [6]. Our main insight is that the NNFs searched in a hazy image associate pixels at the similar depth. Since no depth information is available in single-image dehazing, we utilize the NNF information to infer depth cues by allowing non-local propagation of latent scattered light, which is exponentially proportional to depth [7]. To the best of our knowledge, this approach is the first work that combines MRF regularization with NNFs for dehazing. This proposed regularization method can be used with any other dehazing algorithms to enhance haze regularization.

2 Related Work

Previous works on dehazing can be grouped into three categories: multiple image-based, learning-based and single image-based approaches.

Multiple Image-Based Dehazing. Since removing haze in the atmosphere is an ill-posed problem, several works have attempted to solve the problem using

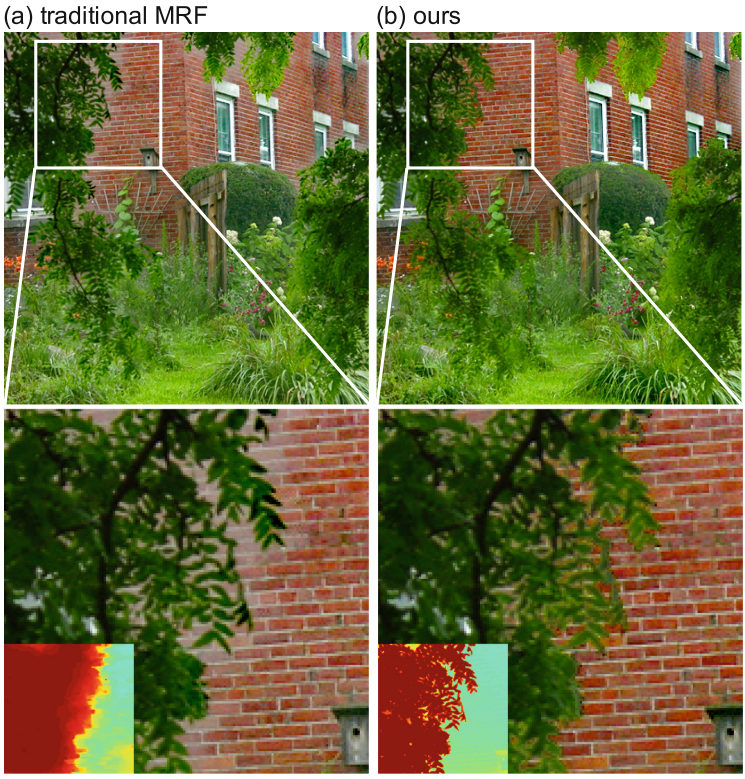

Fig. 1. Comparison of dehazing results using (a) regularization of haze using traditional MRFs commonly used in state-of-the-art dehazing algorithms [3–5] and (b) our regularization using *MRFs with iso-depth NNFs* (Insets: corresponding transmission maps). The proposed method for single-image dehazing can propagate haze more effectively than traditional regularization methods by inferring depth from NNFs in a hazy image. Images courtesy of Kim and Kim [1]. (Color figure online)

multiple input images, often requiring additional hardware. Schechner et al. capture a set of linearly polarized images. They utilize the intensity changes of the polarized lights to infer the airlight layer [8]. Narasimhan et al. [7,9] employ multiple images with different weather conditions to restore the degraded image using an irradiance model. Kopf et al. [10] remove haze from an image with additionally known scene geometry, instead of capturing multiple images. These haze formation models stand on the physics of light transport to provide sound accuracy. However, these applications could be limited at the cost of acquiring multiple input images.

Learning-Based Dehazing. Learning-based methods have been proposed to mitigate the ill-posed dehazing problem using a trained prior knowledge. From training datasets, they attempt to earn a prior on natural image statistics to factorize the haze layer and the scene radiance from the hazy image.

Tang et al. [11] define haze-relevant features that are related to the properties of hazy images, and train them using the random forest regression. Zhu et al. [12] obtain the color attenuation prior using supervised learning. They found that the concentration of haze is positively correlated with the difference between brightness and saturation, and they train a linear model via linear regression. However, no general statistical model can predict the diverse distributions of natural light environments; hence, they often fail to restore hazy-free images that are not similar to the trained dataset.

Single Image-Based Dehazing. Owing to the ill-posedness of the dehazing problem, single image-based methods commonly rely on a certain assumption on statistics of natural images. Most prior works have made an assumption on the statistics of natural *scene radiance* [2,4,13–16]. Tan [13] and Tarel and Hautiere [14] restore visibility by maximizing local contrast, assuming that clean color images have a high contrast, but this causes overly saturated results. He et al. [2] exploit image statistics where a natural image in the sRGB color space should include a very low intensity within a local region. However, it often overestimates the amount of haze if there is a large area having bright pixels. Nishino et al. [15] employ scene-specific priors, a heavy-tailed distribution on chromaticity gradients of colors of natural scenes, to infer the surface albedo, but they also often produce over-saturated results.

Developing the natural image prior further, Fattal [4] assumes that in the sRGB space, the color-line of a local patch within a clear image should pass through the origin of the color space. This can yield a clear and naturally-looking result, but it requires per-image tweaking parameters such as the gamma value and the manual estimation of the atmospheric light vector. Li et al. [17] suggest a nighttime dehazing method that removes a glow layer made by the combination of participating media and light source such as lamps. Recently, a non-local transmission estimation method was proposed by Berman et al. [5], which is based on the assumption that colors of a haze-free image can be approximated by a few hundred distinct colors forming tight clusters in the RGB space.

In addition, an assumption on *light transport* in natural scenes is also used. Fattal assumes that shading and transmission are statistically independent [3], and Meng et al. [18] impose boundary conditions on light transmission. In particular, our airlight estimation follows the traditional approach based on dimension-minimization approach [3], which allows for robust performance in estimating airlight.

Haze Regularization. Most single-image dehazing methods estimate per-pixel haze using a patch-wise operator. Since the operator often fails in a large portion of patches in practice, regularizing sparse haze estimates is crucial to obtain a dense transmission map for restoring a haze-free image. Grid Markov random fields are most commonly used in many dehazing algorithms [3,5,13,15,19], and filtering methods are also used, for instance, matting Laplacian [2], guided filtering [20], and a total variation-based approach [14,18]. These regularization methods only account for local information, they often fail to obtain sharp depth-discontinuity along edges if there is an abrupt change in scene depth.

Recently, Fattal [4] attempts to mitigate this isolation problem by utilizing augmented Markov random fields, which extend connection boundaries of MRFs. However, this method does not search neighbors in every region in an image since only pixels within a local window are augmented. For this reason, the augmented MRFs cannot reflect all non-local information in the image, and in some cases, isolation artifacts still remain. Berman et al. [5] non-locally extend the boundary in estimating haze; however, they still regularize an initial transmission map by using Gaussian MRFs (GMRFs) with only local neighbors. As a result, severe isolation problems occur in a region where there is an abrupt change of scene depth. In our regularization method, we extend neighbors in MRFs with NNFs to allow non-local propagation across iso-depth pixels to obtain sharp edge-discontinuity when inferring latent transmission values.

3 Initial Estimation of Haze

We first estimate the initial density of haze following a traditional dimension-reduction approach using linear subspaces [3,7]. To help readers understand the formulation of the dehazing problem, we briefly provide the foundations of the traditional haze formation model.

Haze Formation Model. Haze is an aerosol that consists of ashes, dust, and smoke. Haze tends to present a gray or bluish hue [7], which leads to decrease of contrast and color fidelity of the original scene radiance. As the amount of scattering increases, the amount of degradation also increases in light transport. This phenomenon is defined as a *transmission* function that represents the portion of light from the scene radiance that is not affected by scattering in participating media.

The relationship between the scattered light and the attenuated scene radiance has been expressed as a linear interpolation via a transmission term commonly used in many dehazing algorithms [3,4,7,9]:

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

where $I(x)$ is the linear signal intensity at pixel x, $J(x)$ is unknown scene radiance, $t(x)$ is the transmission ratio, describing the portion of remaining light when the reflected light from a scene surface reaches to the observer through the medium, and A is a global atmospheric vector, which is unknown. The atmospheric vector A represents the color vector orientation and intensity of atmospheric light in the linear sRGB color space, and along with the interpolation term $(1 - t(x))$, the right additive term in Eq. (1) defines the intensity of airlight at an arbitrary pixel x. Additionally, the atmospheric vector is independent of scene locations, i.e., the atmospheric light is globally constant.

The amount of scattering is closely related to the distance that light travels, i.e., the longer light travels, the more scattering occurs. Therefore, the transmission decays as light travels. Suppose that haze is homogeneous; this phenomenon then can be written as follows: $t(x) = e^{-\beta d(x)}$, where β is a scattering coefficient

of the atmosphere [9] that controls the amount of scattering, and $d(x)$ is the scene depth at the pixel x.

The goal of haze removal is to estimate the transmission ratio t and the atmospheric vector A so that scene radiance J can be recovered from the transmission t and the atmospheric vector A as follows:

$$J(x) = \frac{I(x) - (1 - t(x))A}{\max(\varepsilon, t(x))}, \tag{2}$$

where ε is a small value to prevent division by zero.

Linear Color Space. As shown in Eq. (1), the hazy image formation is the linear combination of a scene radiance and haze factor. Dehazing is a process to perform subtraction from the input intensity by the amount of haze. The required condition for the process is that the pixel intensity must be linearly proportional to the incident radiance based on physics, which is only valid in the *linear* sRGB color space. However, *gamma correction* is already baked in the pixel intensities in general color images. If one subtracts the haze factor from the original intensity in the nonlinear sRGB space, dehazing results appear inconsistent with different levels of pixel intensities. Consequently, manual tweaking parameters are often required, as in Fattal [4]. Different from existing dehazing methods [2–4,17], we first perform inverse gamma correction to linearize the pixel values before recovering the scene radiance; i.e., we use a linearized image I_L by applying a power function with an exponent of the standard display gamma to an sRGB value: $I(x) = \{I'(x)\}^{\gamma}$, where $I'(x)$ is a non-linear RGB value, and γ is a display gamma (e.g., $\gamma = 2.2$ for the standard sRGB display), instead of I during the transmission estimation and regularization processes, and we then perform gamma correction for display.

Haze Estimation. Since airlight is energy scattered in air, airlight tends to be locally smooth in a scene, i.e., local airlight remains constant in a similar depth. In contrast, the original radiance in a scene tend to vary significantly, naturally showing a variety of colors. When we isolate the scene radiance into a small patch in an image, the variation of scene radiances within a patch tends to decrease significantly to form a cluster with a similar color vector, assuming that the real world scene is a set of small planar surfaces of different colors. Then, one can estimate a transmission value with certain natural image statistics within a patch based on the local smoothness assumption on scene depths.

Following this perspective of the traditional approach [3], we also define a *linear subspace* that presents local color pixels in the color space. A linear subspace in each patch comprises two bases: a scene radiance vector $J(x)$ at the center pixel x and a global atmospheric vector A. In this space, a scene depth is piecewise smooth, and the local pixels share the same atmospheric vector. Now we can formulate dehazing as finding these two unknown basis vectors, approximating the transmission value $t(x)$ that is piecewise smooth due to the local smoothness of a scene depth. Figure 2 depicts the estimation process for an overview.

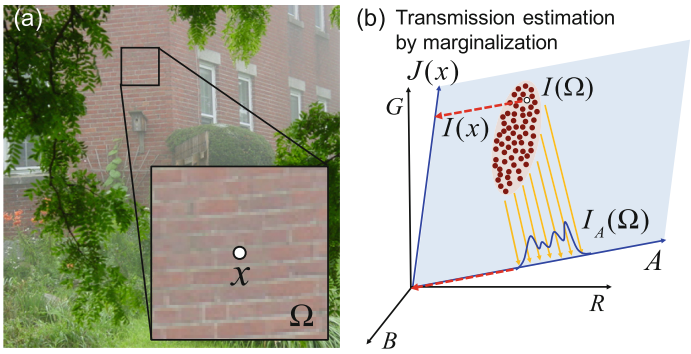

Fig. 2. (a) Extracting a patch from a hazy image. $I(\Omega)$ is a set of linearized color pixels in patch Ω that has a center pixel of x. The white dot indicates the center pixel x. (b) We initially estimate the amount of haze using linear subspaces [3,7]. A is an atmospheric vector of the image (a), $I(x)$ is the linearized center pixel x depicted as the white dot, and $J(x)$ is the scene radiance vector of the pixel $I(x)$. Pixel intensity $I(x)$ is a linear interpolation of the vector A and $J(x)$, and hence lies on the linear subspace [the blue plane in (b)] spanned by those two vectors. The red dots describe pixels extracted from $I(\Omega)$. These pixels are projected onto vector A to obtain a marginal distribution with respect to A. The red arrow from the cluster denotes the amount of airlight that is determined from the minimum value of the marginal distribution. Images courtesy of Kim and Kim [1]. (Color figure online)

Atmospheric Vector Estimation. Airlight is a phenomenon that acts like a light source, which is caused by scattering of participating media in the atmosphere [7]. The atmospheric vector represents the airlight radiance at the infinite distance in a scene, i.e., the color information of airlight itself. Therefore, the atmospheric vector does not include any scene radiance information, and it only contains the airlight component. The region full of airlight is the most opaque area in a hazy image. We follow a seminal method of airlight estimation by He et al. [2]. The atmospheric vector A is estimated by picking up the pixels that have the top 0.1% brightest dark channel pixels and then choosing the pixel among them that has the highest intensity in the input image. However, if there are saturated regions such as sunlight or headlights, maximum filtering of the dark channel could be incorrect since those regions might have the highest (saturated) dark channel. Also, we assume that the most opaque region is the most brightest within an image, and we therefore discard the pixels that are within aforementioned saturated regions. We then select the 0.1% pixels among the rest as He et al. [2]'s method does, so that we can estimate the atmospheric vector consistently. We subsequently average the chosen pixels to reject noise.

Transmission Estimation. We first assume that transmission is piecewise smooth. In Eq. (1), the portion of haze at a pixel x is determined by the term $(1 - t(x))$ that indicates the amount of haze to be removed. We determine the amount of haze from given color signals within a patch. Suppose the

given color signals in each patch are linear combinations of two unknown bases, J and A, that form a linear subspace. If we project the given pixels onto the atmospheric vector A, we can estimate the contribution of the haze signal mixed into the input signals in the patch.

Supposing $I_A(\Omega)$ is a set of *scalar projections* of color vectors $I(\Omega)$ onto an atmospheric vector A in patch Ω (Fig. 2), where the pixel x is located at the center, then it can be written as following Fattal's method [3]:

$$I_A(\Omega) = I(\Omega) \cdot \frac{A}{\|A\|}, \quad I_A(\Omega) \in \mathbb{R}^{1\times|\Omega|}. \tag{3}$$

We assume the airlight within a patch to be constant while the scene radiance might vary. To focus only on the airlight component, it is necessary to obtain a *marginal distribution* of the surrounding pixels with respect to the basis vector A, as shown in Fig. 2(b).

The marginal distribution $I_A(\Omega)$ describes the histogram of airlight components within a patch. This distribution would have had a very low minimum value if it had not been influenced by piecewise constant airlight. However, if we take the minimum projected value, there could be a large chance to take an outlying value as the minimum. We use the i-th percentile value from the projected pixel distribution to reject outliers effectively to achieve robust performance:

$$I_A^{\min}(\Omega) = \underset{k\in\Omega}{P_i}\left(I_A(k)\right), \quad I_A^{\min}(\Omega) \in \mathbb{R}^1, \tag{4}$$

where P_i represents an i-th percentile value ($i = 2$).

The minimum percentile scalar projection onto an atmospheric vector corresponds to the amount of haze of a pixel from its patch, and thus the minimum projection corresponds to the haze component part in Eq. (1), which is $(1 - t(x)) \leftarrow I_A^{\min}(\Omega)$.

Additionally, projection onto the atmospheric vector requires two bases (a pixel and an atmospheric vectors) to be orthogonal. However, pixels within a patch are not necessarily orthogonal to the atmospheric vector, so projection needs to be compensated for non-orthogonality. If a color vector has a small angle with its atmospheric vector, then its projection will have a larger value due to the correlation between the two vectors. We attenuate $I_A^{\min}$ by a function with respect to the angle between a pixel vector and an atmospheric vector that is given by

$$t(x) = 1 - f\left(\bar{\theta}\right) \cdot I_A^{\min}(\Omega), \tag{5}$$

where $\bar{\theta}$ is a normalized angle between a pixel vector and an atmospheric vector within $[0, 1]$. The attenuation function $f()$ is given by

$$f\left(\bar{\theta}\right) = \frac{e^{-k\bar{\theta}} - e^{-k}}{1 - e^{-k}}, \tag{6}$$

where the function has a value of $[0, 1]$ in the range of $\bar{\theta}$. In this work, we set $k = 1.5$ for all cases. This function compensates transmission values by

attenuating the value $I_A^{\min}$ since the function has a value close to 1 if $\bar{\theta}$ has a small value. See Fig. 3(c). Figure 4 shows the impact of our attenuation function. Our attenuation prevents over-estimation of transmission where orthogonality between the atmospheric vector and a color vector does not hold: Thus, we can avoid over-saturated dehazed results.

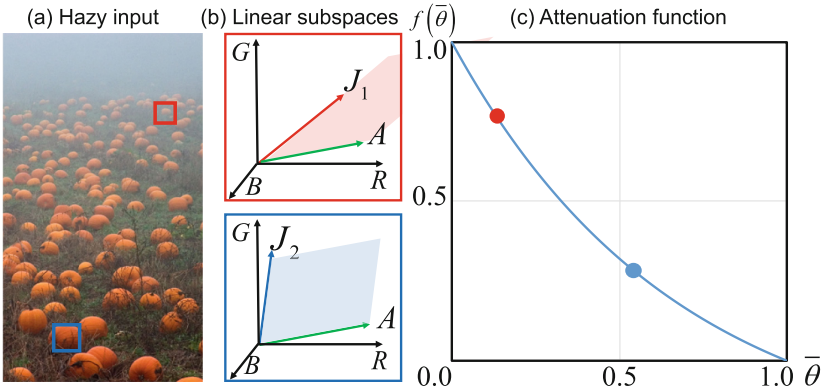

Fig. 3. (a) A hazy input image. (b) Each single pixel from the red and blue boxes is plotted in the RGB space along with the atmospheric vector A. J_1 and J_2 in each plot correspond to the two pixels extracted. (c) The attenuation function defined as Eq. (6) is plotted as above. The red and blue dots indicate the amount of attenuations of the red and blue patches. This plot shows that the amount of attenuation increases as an angle between a color vector and an atmospheric vector decreases. Images courtesy of Kim and Kim [1]. (Color figure online)

The size of a patch is crucial in our method. If the size is too small, then the marginal distribution does not contain rich data from the patch, resulting in unreliable estimation such as clamping. On the contrary, an excessively large patch might include pixels in different scene depth and our estimation stage takes the minimum value in the marginal distribution, and hence the transmission estimate will be overestimated. In our implementation, we use patches of 15-by-15 pixels and it showed consistent results regardless of the size of an image.

Removing Outliers. While our transmission estimation yields reliable transmission estimates in most cases, however, there are a small number of cases that does not obey our assumption. We take them as outliers and mark them as invalid transmission values, and then interpolate them in the regularization stage (see Sect. 4).

Distant regions in an image such as sky, and objects whose color is grayish have a similar color of haze. In the RGB color space, the angle between an atmospheric vector and the color vector of a pixel in those regions is very narrow and the image pixel's luminance is quite high. In this case, unreliable estimation is inevitable since there is a large ambiguity between the color of haze and scene

radiance. As a result, unless we do not reject those regions, the transmission estimate will be so small that those regions will become very dim in turn. For this reason, we discard the transmission estimates, where the angle between an image pixel and an atmospheric vector is less than 0.2 rad, the pixel's luminance (L^*) is larger than 60 in the CIELAB space, and the estimated transmission value is lower than a certain threshold: 0.4 for scenes having a large portion of distant regions and 0.1 for others.

When estimating an atmospheric light, we assumed that the most opaque region in an image is the brightest area of the whole scene. However, pixels brighter than the atmospheric light can exist due to very bright objects such as direct sunlight, white objects, and lamps in a scene. Those pixels do not obey our assumption above, and hence this leads to wrong transmission estimation. Therefore, we discard pixels whose luminance is larger than the luminance of the atmospheric light.

4 Non-local Regularization Using Iso-Depth Neighbor Fields

Once we calculate the initial estimates of transmission for every pixel, we filter out invalid transmission values obtained from extreme conditions. The transmission estimation and outlier detection stages might often yield incomplete results with blocky artifacts. We, therefore, need to propagate valid transmission values in the image.

MRF Model. As we mentioned above, the transmission is locally smooth. Therefore, in order to obtain a complete transmission map having sharp-edge discontinuities, we need to propagate the incompletely estimated transmission map using Markov random fields. The probability distribution of one node in an MRF is given by

$$p\left(t\left(x\right)\left|\hat{t}\left(x\right)\right.\right) = \phi\left(t\left(x\right), \hat{t}\left(x\right)\right) \psi\left(t\left(x\right), t\left(y\right)\right), \tag{7}$$

where $t\left(x\right)$ is a latent transmission variable at pixel x, $\hat{t}\left(x\right)$ is an initially estimated transmission value (see Sect. 3), $\phi()$ is the data term of the likelihood between $t(x)$ and $\hat{t}(x)$, and $\psi()$ is a smoothness prior to relate latent transmission $t(x)$ with neighboring transmission $t(y)$, where y is a neighboring pixel of x. While the data term $\phi()$ describes the fidelity of observations by imposing a penalty function between the latent variable and the observed value, the regularization term $\psi()$ enforces smoothness by penalizing the errors between one latent variable and its neighboring variables.

The data term $\phi()$ is given by

$$\phi\left(t\left(x\right), \hat{t}\left(x\right)\right) = \exp\left(-\frac{\left(t\left(x\right) - \hat{t}\left(x\right)\right)^2}{\sigma_{\hat{t}}(\Omega)^2}\right), \tag{8}$$

where $\sigma_{\hat{t}}(\Omega)$ is the variance of observation values $\hat{t}$ within patch Ω that has the center at pixel x. See Fig. 5. The data term models error between a variable

and observation with in-patch observation variance noise via a Gaussian distribution. The in-patch variance of observation values implies that the greater the variance of in-patch observation is, the more uncertain the observation values are, resulting in giving less influence from the data term on the distribution.

The regularization term $\psi()$ is written as

$$\psi\left(t\left(x\right),t\left(y\right)\right)=\prod_{y\in N_x}\exp\left(-\frac{\left(t\left(x\right)-t\left(y\right)\right)^2}{\left\|I\left(x\right)-I\left(y\right)\right\|^2}\right), \tag{9}$$

where $I()$ is a linearized pixel intensity of an image, and pixel y is in a set of neighbors N_x of pixel x. The regularization term encourages smoothness among one variable and its neighboring variables by penalizing pairwise distances between them, where the distribution of the distances follows a Gaussian distribution. If $\left(t\left(x\right)-t\left(y\right)\right)^2$ is large, then it indicates that the distance between $t\left(x\right)$ and its neighbor $t\left(y\right)$ is large, and hence the cost from the regularization term will also become large, which enforces strong smoothness between them. $\left\|I\left(x\right)-I\left(y\right)\right\|^2$ in the denominator of the prior term controls the amount of smoothness by exploiting information from an input image. This property implies that if two image pixels are similar, then their transmission values are likely to be similar as well. On the contrary, it gives sharp-edge discontinuity in transmission values along edges since the value of the denominator becomes large when the difference between two pixels is large.

In fact, the probability distribution of an MRF over the latent variable t is modeled via the Gaussian distribution. In this case, the MRF is formalized by using a Gauss-Markov random field (GMRF), which can be solved by not only using computationally costly solvers, but also by a fast linear system solver [3,21].

Finally, we formulate a cost function by taking the negative log of the posterior distribution [Eq. (7)] following Fattal's method [3,4], which is written by

$$E\left(t\right)=\sum_x\left\{\frac{\left(t\left(x\right)-\hat{t}\left(x\right)\right)^2}{\sigma_{\hat{t}}\left(\Omega\right)^2}+\sum_{y\in N_x}\frac{\left(t\left(x\right)-t\left(y\right)\right)^2}{\left\|I\left(x\right)-I\left(y\right)\right\|^2}\right\}. \tag{10}$$

The regularization process is done by minimizing the cost function, which is solved by differentiating the function with respect to t and setting it to be zero.

Iso-Depth Neighbor Fields. In conventional grid MRFs, a prior term [Eq. (9)] associates adjacent four pixels as neighbors. However, pixels in a patch lying on an edge may be isolated when the scene surface has a complicated shape. In Fig. 5(a), the leaves in the left side of the image have a complicated pattern of edges, and the bricks lie behind the leaves. If we model a grid MRF on the image, then pixels on the tip of the leaves will be isolated by the surrounding brick pixels. In this case, smoothness of the leaf pixels will be imposed mostly by the brick pixels, where there is a large depth discontinuity between them. In other words, a large scene depth discrepancy exists in the patch, and thus if some pixels lying on the edge are only connected to their adjacent neighbors, the

Fig. 4. Impact of our attenuation term in transmission estimation. As the red arrows indicate, our attenuation term prevents over-estimation of transmission which results in over-saturation. (Color figure online)

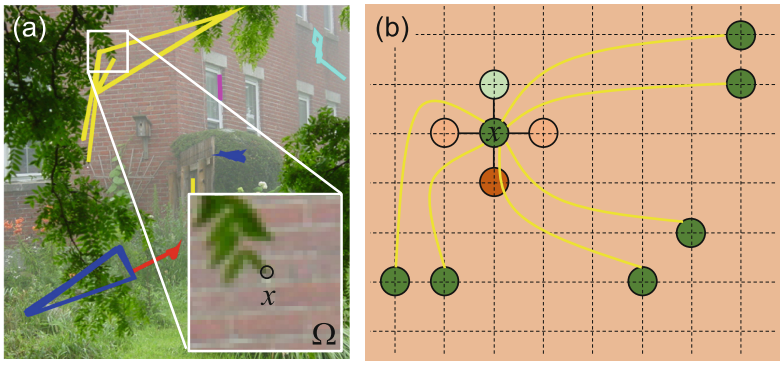

Fig. 5. (a) The picture shows some sampled NNFs that associate pixels having similar scene depths. The line with the same color denotes association of pixels in the same NNF. (b) An MRF model of the node x from the patch in (a) associated with adjacent four neighbors and distant neighbors in the NNF. Since the node x is located in the end point of the leaf, its adjacent pixels have very different transmission values due to a large depth discontinuity. As (a) shows, the neighbors connected with the same NNF have very similar scene depths, and hence they give a more accurate regularization cue than the adjacent neighbors do. Images courtesy of Kim and Kim [1]. (Color figure online)

prior term will enforce wrong smoothness due to the large depth discrepancy. As a result, those regions will be overly smoothed out due to the wrong connection of neighbors.

Algorithm 1. Dehazing via Non-Local Regularization.

Require: an image I
Ensure: a result image J and a transmission map t
$\hat{A} \leftarrow$ atmosphericVectorEstimate(I)
$\{I_L, A\} \leftarrow$ inverseGammaCorrect($\{I, \hat{A}\}$)
for pixels $x = 1$ to n **do**
  $I_A(\Omega) \leftarrow I_L(\Omega) \cdot \frac{A}{\|A\|}$
  $I_A^{\min}(\Omega) \leftarrow \min_{k \in \Omega} (I_A(k))$
  $t'(x) \leftarrow 1 - f(\bar{\theta}) \cdot I_A^{\min}(\Omega)$
  $\hat{t}(x) \leftarrow$ outlierReject($t'(x), A, I_L$ (x))
end for
$NNF \leftarrow$ PatchMatch(I)
$t \leftarrow$ regularize($NNF, \hat{t}, I$)
$J_L \leftarrow (I - (1-t)A)/t$
$J \leftarrow$ gammaCorrect(J_L)

We investigate neighbors extracted from a nearest-neighbor field (NNF) using the PatchMatch algorithm and found that the NNF associates pixels at similar scene depths. This insightful information gives a more reliable regularization penalty since the neighboring nodes in the NNF are likely to have similar transmission estimates. We validate our method through evaluation using synthetically generated hazy images along with their ground truth depth maps. Figure 6 shows the synthetic hazy scenes and their corresponding depths. We compute the absolute depth difference between a ground truth depth pixel and its iso-depth pixels associated by NNFs. The histograms in Fig. 6 show that the NNFs link one pixel to others having similar depth values. Thus, we add more neighbors belonging to the same NNF to the smoothness term and perform statistical inference on the MRF along with them. We note that these long-range connections in regularization are desirable in many image processing applications, addressed by other works [4,22]. After regularization, we use the weighted median filter [23] to refine the transmission map. Algorithm 1 summarizes our dehazing algorithm as an overview.

5 Results

We implemented our algorithm in a non-optimized MATLAB environment except the external PatchMatch algorithm [6], and processed it on a desktop computer with Intel 4.0 GHz i7-4790K CPU and 32 GB memory. For the case of the house image of resolution 450×440 shown in Fig. 1(b), our algorithm took

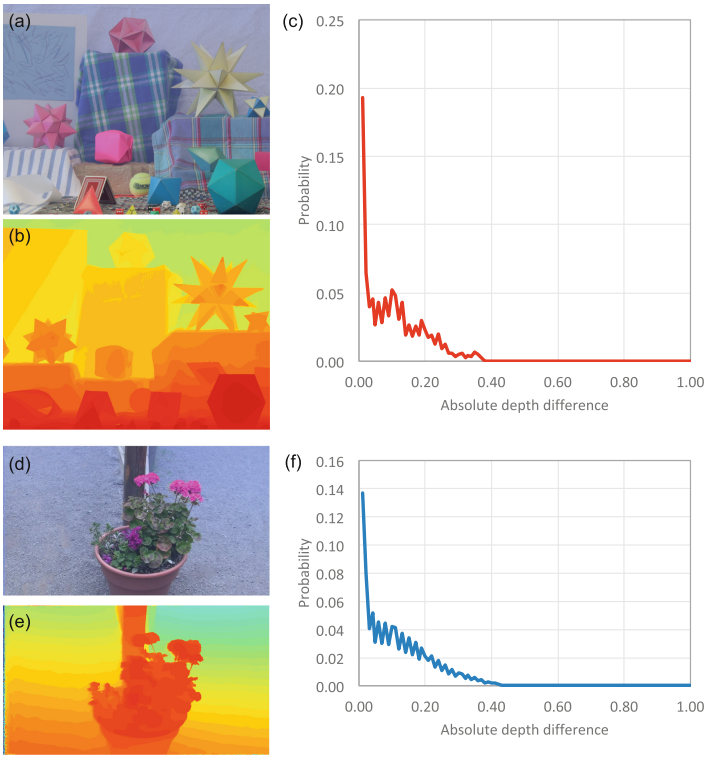

Fig. 6. Images (a) and (d) are synthetically generated hazy scenes, and images (b) and (e) are their ground truth depth maps. We computed the absolute transmission difference between a transmission value of a pixel and its iso-depth pixel's transmission values associated by PatchMatch [6]. Plots (c) and (f) are the distributions of the differences. Plot (c) shows that the portion of the absolute difference below 0.2 occupies 86% of entire NNFs, while Plot (f) shows the case of 81%.

6.44 s for running the PatchMatch algorithm to seek 17 neighbors, 8.32 s to estimate an atmospheric vector, transmission values and rejecting outliers, 43.43 s for our regularization stage, and 0.65 s for running the weighted median filter and recovering the scene radiance, taking approximately 58.84 s in total. We evaluated our algorithm with a large number of outdoor hazy images obtained from Fattal's method [4] to prove robustness, and we also present comparisons with state-of-the-art dehazing methods. Refer to the supplemental materials for more results.

Regularization. We compare results of our method with those of state-of-the-art methods in terms of regularization. Berman's method [5] regularizes initial transmission estimates with a grid GMRF as shown in the second columns in Fig. 7. Due to the lack of non-local information in regularization, certain regions suffer from the haze isolation problem as mentioned above. Other than using a

grid MRF, Fattal's method [4] takes an augmented GMRF model for regularization, which extends neighbor fields within a local window. However, it does not connect more neighbors for all pixels due to time complexity. As a result, certain regions are not fully recovered from the haze isolation problem. Figure 7 validates that our method successfully removes haze even from a scene having abrupt depth changes with complicated patterns.

Figure 8 shows the intermediate stages in our regularization process of transmission (d)–(g), along with our result of the house scene (c). We start our regularization from Fig. 8(d) that has outliers [represented as black pixels in Fig. 8(d)]. In particular, Fig. 8(e) and (f) compare the impact of NNFs in the MRF regularization. When we regularize the initial estimate with only GMRFs, certain regions with complex scene structures are over-smoothed due to the wrong smoothness penalty as Fig. 8(e) shows. We account for additional neighbors from NNFs to obtain a clearer transmission map shown in Fig. 8(f). Figure 8(g) shows the final transmission map that we refine with a weighted median filter [23].

We also compare our regularization method with representative matting methods: the matting Laplacian method [24] and the guided filter method [20] in Fig. 9. While we use the guide image as just a guide to smooth and enforce sharp gradient along edges on transmission estimates, both methods are based on the assumption that an output and an input guidance form a linear relationship. As described in Sect. 3, scene radiance varies largely while transmission does the opposite. Consequently, the two methods follow the behavior of the scene radiance, which results in distorting the given estimates. As a result, our regularization method yields an accurate transmission map with clear-edge discontinuities while the others overestimate the transmission estimates in turn.

Qualitative Comparison. Figure 10 qualitatively validates the robust performance in dehazing the common reference dataset of hazy scenes [4]. We compare the performance of our dehazing algorithm with three state-of-the-art methods [2,4,5]. We were motivated to achieve consistent performance of dehazing with less parameter controls like other image processing algorithms [25,26]. Figure 10 shows results using the single set of parameters as described in Sect. 3. Our method shows competitive results to other method [4] that requires manual tweaking parameters per scene to achieve plausible results. For close-up images of the results, refer to the supplemental material.

Time Performance. Table 1 compare the computational performance of our method with traditional grid GMRFs and our iso-depth GMRFs using images shown in Fig. 10. We also shows computational costs of obtaining only NNFs with 17 neighbors using PatchMatch [6] in the third row. Dehazing with iso-depth NNF-GMRFs takes 10.58 times more time; however, iso-depth NNFs give richer information in regularization, resulting in more exact scene radiance recovery.

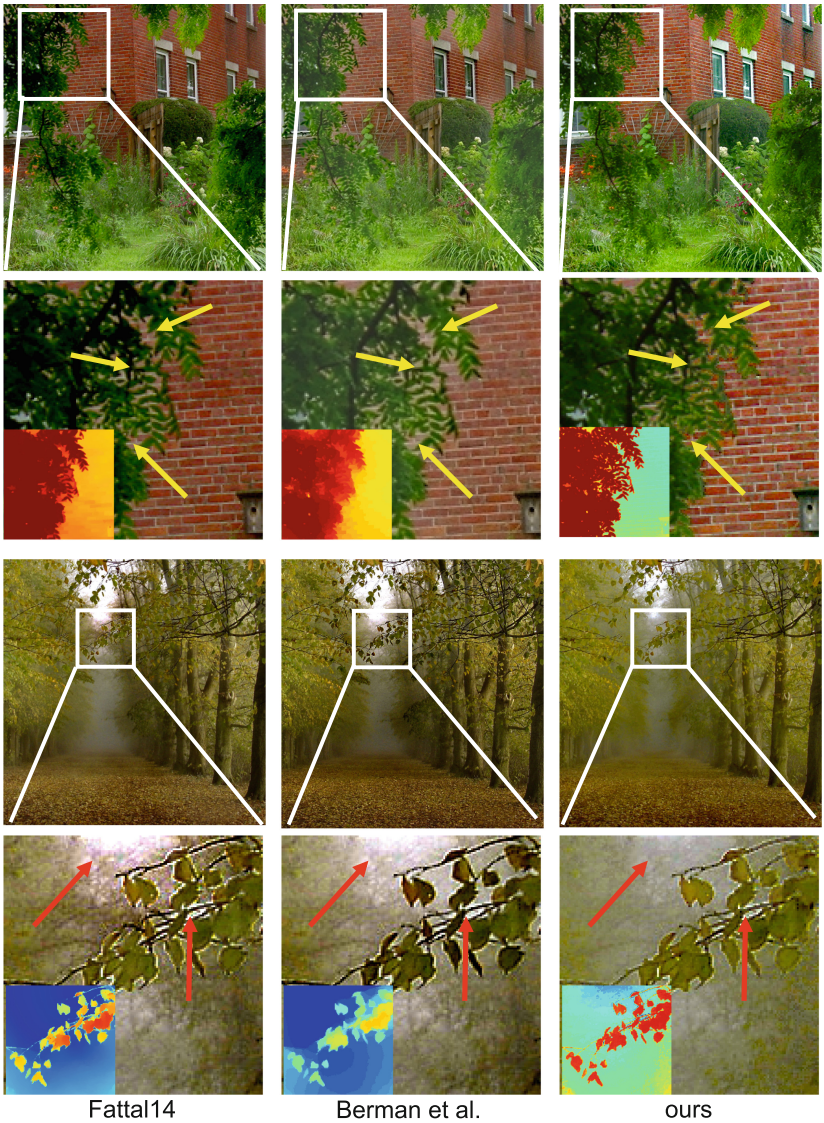

Fig. 7. Comparisons of dehazing in terms of regularization. The two columns from left are results from other two methods: Fattal's method [4] using augmented GMRFs and Berman's method [5] using traditional GMRFs, and the third column is our results (Insets: corresponding transmission maps). While other methods often fail to obtain sharp edge-discontinuities in the images, our method yields clear recovered scene radiance maps as shown above. Notable regions are pointed with arrows. Images courtesy of Kim and Kim [1].

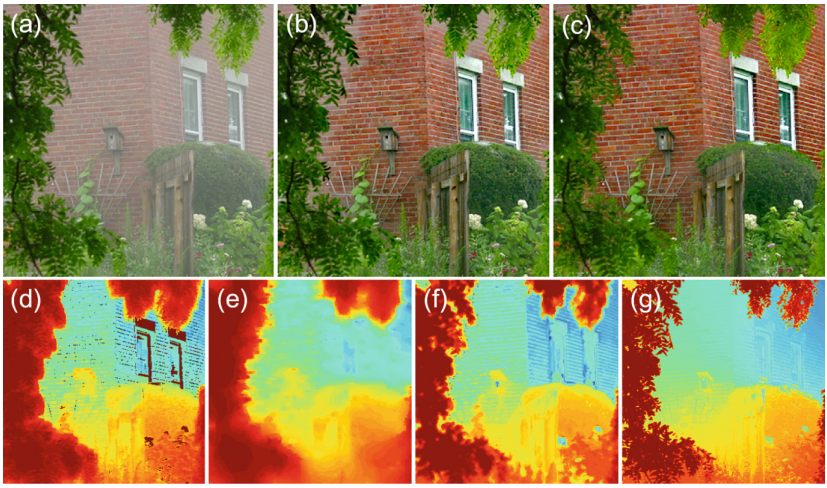

Fig. 8. We present an example before and after applying our dehazing and regularization method. (a) The hazy input image. (b) The recovered scene radiance map with the transmission map regularized by grid MRFs (e). (c) The recovered scene radiance map with the final transmission map (g). Images (d)–(g) compare transmission maps to show the influence of using iso-depth NNFs. All regularizations are done using GMRFs. (d) The initial transmission estimates including discarded pixels (the black pixels). (e) The regularized transmission map without NNFs. (f) The regularized transmission map with NNFs. (g) The final refined map of (f) using the weighted median filter. Images courtesy of Kim and Kim [1].

Quantitative Comparison. We compare our method with the entire synthetic hazy image dataset provided by Fattal [4]. The synthetic hazy images were generated by datasets that contain clear indoor and outdoor scenes, and their corresponding depth maps. Table 2 reports the quantitative comparison of our method with other methods: He et al. [2], Fattal [4], and Berman et al. [5]. Additionally, we present the statistics in Table 2 and Fig. 11. It says that our method shows the best performance in dehazed images, and is strongly competitive to state-of-the-arts in transmission maps. We also show the dehazed images used for the quantitative comparison in Fig. 12. Our method shows competitive and consistent results particularly in dehazed images.

Table 1. Comparison of time performance of dehazing with the traditional grid GMRFs and our GMRFs with iso-depth NNFs (unit: second). Refer to Fig. 10 for processed images. The third row shows computational costs of only seeking NNFs with 17 neighbors using PatchMatch [6] in our method. The table courtesy of Kim and Kim [1].

Dehazing	House	Forest	ny17	Train	Snow	Castle	Cones	Average
With grid GMRFs	6.43	26.55	27.51	7.74	18.88	12.84	6.41	15.19
With NNF-GMRFs	58.84	305.48	305.06	73.06	191.76	129.18	61.12	160.64
(Computing NNFs only)	(6.44)	(31.82)	(28.48)	(7.15)	(18.54)	(11.01)	(7.31)	(15.82)

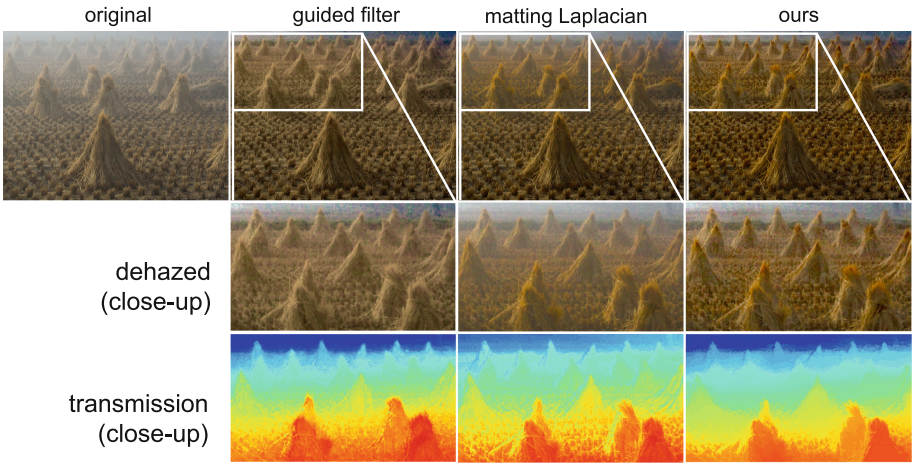

Fig. 9. We compare our regularization with other methods. The leftmost one is the original image of cones. The first row shows dehazed results with our transmission estimation step and each regularization method written at the lower right. We cropped the dehazed images in the first row to highlight the influence of regularization methods in the second row. The third row presents a sequence of cropped transmission maps in the same manner as the second row. Images courtesy of Kim and Kim [1].

Table 2. Quantitative comparisons of our method with other methods [2,4,5]. The error values are computed from the entire synthetic hazy image dataset provided by Fattal [4]. All figures represent mean L_1 error of the estimated transmission t (the left value) and output image J (the right value). Red figures indicate the best results, and blue for the second best. For a fair comparison, parameters for each method, such as display gamma for sRGB linearization and the airlight vector, were optimized for the highest performance. The table courtesy of Kim and Kim [1].

	He [2]	Fattal [4]	Berman [5]	Ours
church	0.0711/0.1765	0.1144/0.1726	0.1152/0.2100	0.1901/0.1854
couch	0.0631/0.1146	0.0895/0.1596	0.0512/0.1249	0.0942/0.1463
flower1	0.1639/0.2334	0.0472/0.0562	0.0607/0.1309	0.0626/0.0967
flower2	0.1808/0.2387	0.0418/0.0452	0.1154/0.1413	0.0570/0.0839
lawn1	0.1003/0.1636	0.0803/0.1189	0.0340/0.1289	0.0604/0.1052
lawn2	0.1111/0.1715	0.0851/0.1168	0.0431/0.1378	0.0618/0.1054
mansion	0.0616/0.1005	0.0457/0.0719	0.0825/0.1234	0.0614/0.0693
moebius	0.2079/0.3636	0.1460/0.2270	0.1525/0.2005	0.0823/0.1138
reindeer	0.1152/0.1821	0.0662/0.1005	0.0887/0.2549	0.1038/0.1459
road1	0.1127/0.1422	0.1028/0.0980	0.0582/0.1107	0.0676/0.0945
road2	0.1110/0.1615	0.1034/0.1317	0.0602/0.1602	0.0781/0.1206
average	0.1181/0.1862	0.0839/0.1180	0.0783/0.1567	0.0836/0.1152

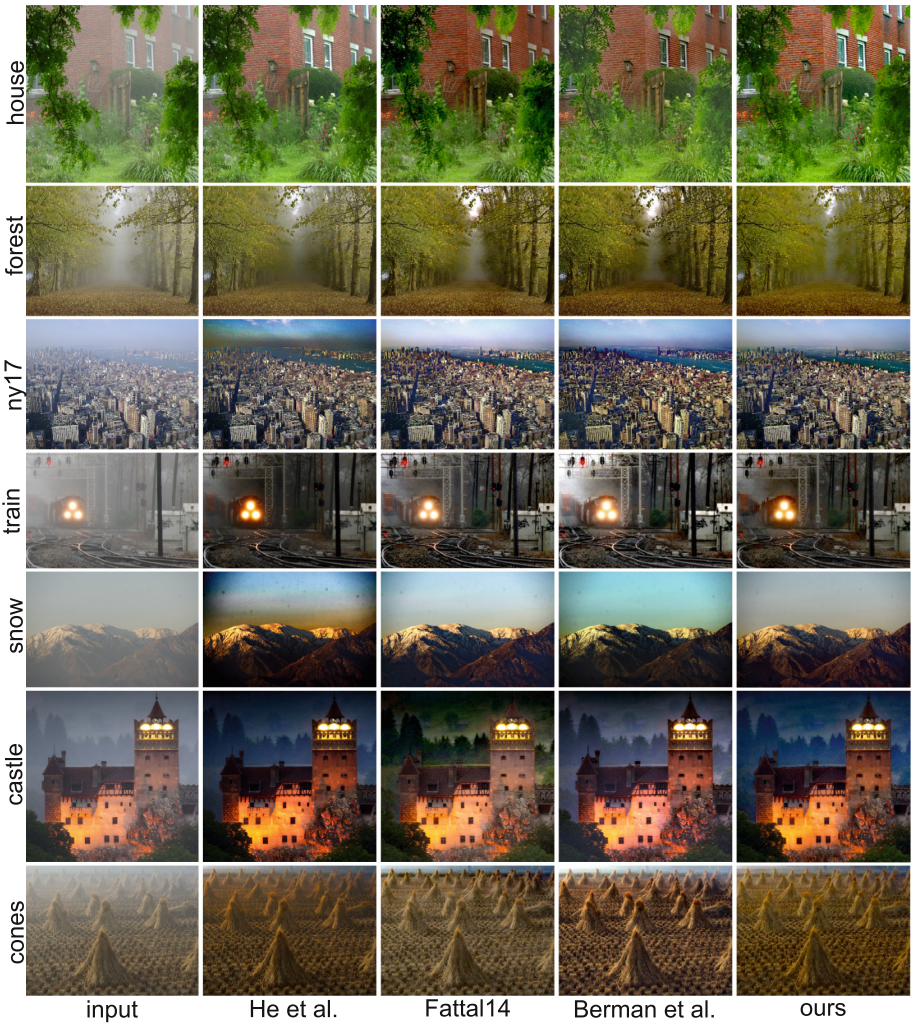

Fig. 10. Validation of consistency of dehazing. The first column shows input images. The second, third, and fourth columns are results from He et al. [2], Fattal [4], Berman et al. [5], respectively. The fifth column presents our method's results. We use the set of parameters as described in Sect. 3. For the images in the third and fifth rows, we only set the threshold of lower bound transmission to 0.4 and the others to 0.1 for removing narrow angle outliers. Our method is competitive to other method [4] that requires with manual tweaking parameters to achieve plausible results. Refer to the supplemental material for more results. Images courtesy of Kim and Kim [1].

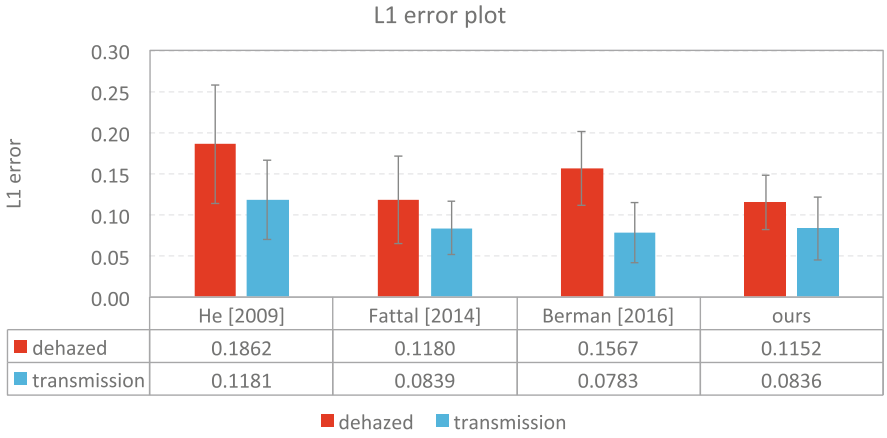

Fig. 11. Mean L1 error plots of 11 pairs of transmission maps and dehazed (Table 2) results, respectively. Our method shows the best performance in dehazed images, and is strongly competitive to state-of-the-arts in transmission maps.

Fig. 12. Dehazed results for the quantitative comparison shown in Table 2. The first column shows synthetic hazy images generated from the ground truth dataset [4] in the second column with their corresponding depth maps. The remaining columns are recovered scene radiance maps by each method. Our method yields consistent results compared with other methods. Parameters for each method were optimized for the highest performance for a fair comparison. Images courtesy of Kim and Kim [1].

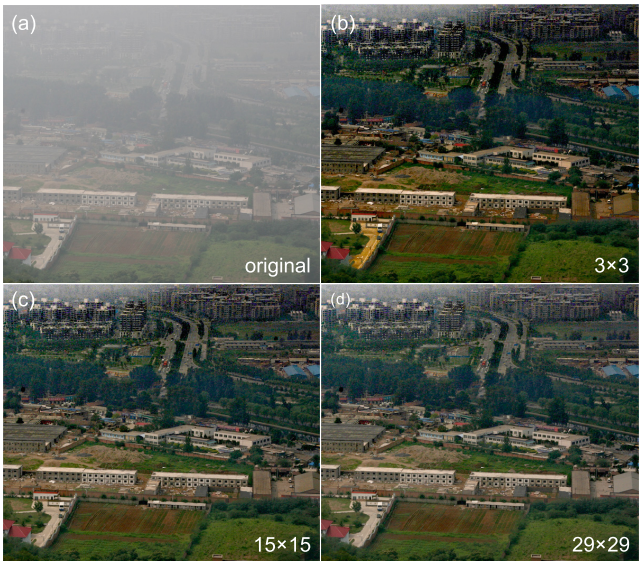

Fig. 13. Comparisons to show the influence of a patch size in estimating transmission. (a) The original canon image. (b) The dehazed image with a patch size of 3×3 where severe color clamping happens. (c) The dehazed image with a patch size of 15×15, which is our choice for all results. (d) The dehazed image with a patch size of 29×29 in which the airlight in distant regions is underestimated. Images courtesy of Kim and Kim [1].

Fig. 14. Validation of our narrow angle outlier rejection method described in Sect. 3. In the second column, the distant region represented as sky has an infinite depth, and hence our transmission estimation stage estimates its transmission as being close to zero, which yields overly saturated results. We obtained consistent results by our outlier rejection stage, as shown in the third column. Images courtesy of Kim and Kim [1].

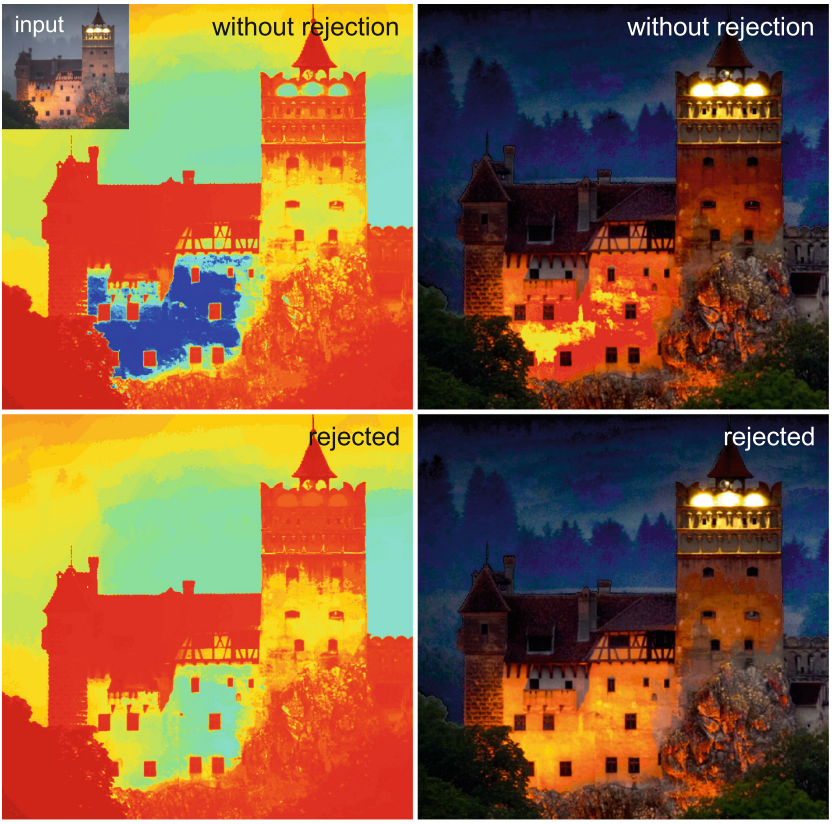

Fig. 15. Validation of our saturated outlier rejection method described in Sect. 3. The inset is an input hazy image. The first column shows estimated transmission maps without our rejection (upper left) and with our rejection (bottom left). The corresponding dehazed scenes are shown in the second column. The bright regions (the direct light at the upper right and the wall of the castle in the middle) are overly saturated. Our outlier rejection succeeds to produce a consistent result by discarding those regions.

Impact of Patch Size. Figure 13 shows the results of dehazing under varying patch sizes. Image (a) is an input image of canon, the size of which is 600×524. Image (b) is severely over-saturated since the size of patches is so small that each patch cannot contain rich information of scene structures, i.e., the patch failed to reject the influence of highly-varying nature of scene radiance. On the other hand, as shown in image (d), its airlight is underestimated since patches are too large to hold the assumption that transmission is piecewise constant.

This underestimation result is exacerbated in distant regions where their scene depth changes rapidly. In our experiment, we found that the patch size of 15×15 works properly for most scenes, and therefore we take the same patch size for all results in this paper.

Outlier Removal. We validate our outlier-rejection process. Figure 14 shows the regions in infinite scene depths occupy a large portion of the image that is full of airlight in the two input images. In these regions, there is a large ambiguity between airlight and scene radiance, and hence our method fails to produce a naturally looking result as the second column shows. After we discard outliers having a narrow angle between the atmospheric vector and the input color pixel, we could obtain high-quality scene radiance maps in the third column. We also show the influence of saturated intensity outliers as mentioned in Sect. 3. We estimated an atmospheric vector under the assumption that the atmospheric light is the brightest all over a scene. As Fig. 15 presents, without rejecting saturated intensity outliers, transmission of those pixels will be severely overestimated due to their high luminance. We can also reject those regions by increasing a patch size; however, this will cause underestimation of airlight and cannot handle a large area as well.

6 Limitations

While our method produces consistent results for most cases; however, there are a small number of cases where our atmospheric vector estimation stage fails. Figure 16 shows an example of our algorithm's failure in finding the correct atmospheric light. There are clouds in the image that occupy relatively large regions but are not saturated, and therefore in the atmospheric vector estimation stage, our method selects pixels in cloud regions as candidates of the atmospheric light, which is not correct. For this reason, our transmission estimation stage severely overestimates the amount of airlight, particularly in distant regions in the scene as shown in Fig. 16(b). We validate the limitation by picking up the atmospheric vector of the image manually, and our algorithm yields a naturally-looking result, as the Fig. 16(c) presents. In addition, if there is a large region that is grayish and thereby has a narrow angle between an atmospheric vector and the region color, our algorithm fails to find correct transmission estimates since there are too many outliers according to our outlier rejection stage, which leads to unreliable regularization. We leave these problems as future work.

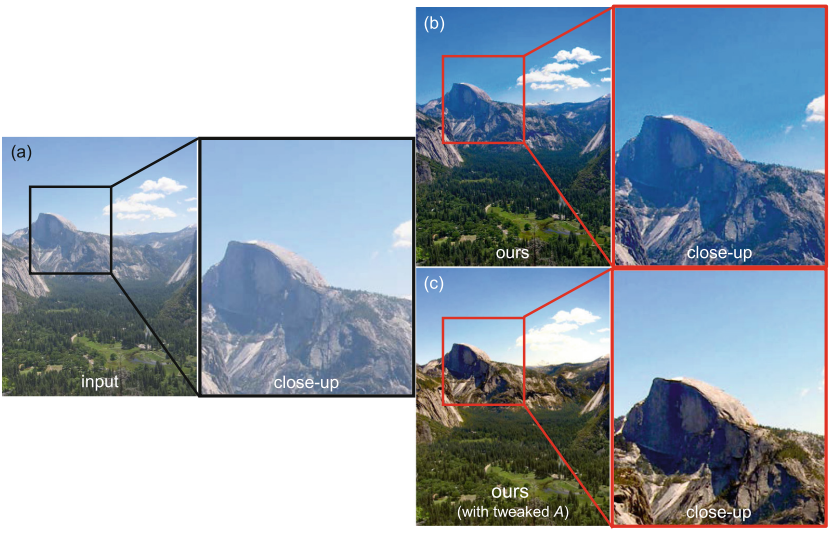

Fig. 16. Our failure case with a landscape image. Image (a) shows the input image, and Image (b) presents our result with the same set of parameters described. Image (c) is our result produced with the manually-tweaked atmospheric vector. (Color figure online)

7 Conclusion

We have presented a single-image dehazing method with our novel non-local regularization using iso-depth neighbor fields. While traditional dehazing methods often suffer from haze isolation artifacts due to improper propagation of the haze cues in the transmission map, our dehazing method can clarify hazy images robustly thanks to our iso-depth regularization approach. Our non-local regularization method infers nonlocal iso-depth cues to obtain more reliable smoothness penalty for better handling the isolation problem even with blunt changes of depth. The proposed iso-depth regularization method is independent of haze-component estimation so that it is directly applicable to any state-of-the-art dehazing methods.

Acknowledgments. Min H. Kim, the corresponding author, acknowledges Korea NRF grants (2016R1-A2B2013031, 2013M3A6A6073718) and additional support by KOCCA in MCST of Korea, Cross-Ministry Giga KOREA Project (GK17P0200), Samsung Electronics (SRFC-IT1402-02), and an ICT R&D program of MSIT/IITP of Korea (2017-0-00072, 2016-0-00018). We also would like to appreciate Seung-Hwan Baek's helpful comments.

References

1. Kim, I., Kim, M.H.: Dehazing using non-local regularization with Iso-depth neighbor-fields. In: Proceedings of International Conference Computer Vision, Theory and Applications (VISAPP 2017), Porto, Portugal (2017)
2. He, K.M., Sun, J., Tang, X.: Single image haze removal using dark channel prior. In: Proceedings of IEEE CVPR, pp. 1956–1963 (2009)
3. Fattal, R.: Single image dehazing. ACM Trans. Graph. **27**, 72:1–72:9 (2008)
4. Fattal, R.: Dehazing using color-lines. ACM Trans. Graph. **34**, 13:1–13:14 (2014)
5. Berman, D., Treibitz, T., Avidan, S.: Non-local image dehazing. In: IEEE CVPR, pp. 1674–1682 (2016)
6. Barnes, C., Shechtman, E., Finkelstein, A., Goldman, D.B.: Patchmatch: a randomized correspondence algorithm for structural image editing. ACM Trans. Graph. **28**, 24:1–24:11 (2009)
7. Narasimhan, S.G., Nayar, S.K.: Vision and the atmosphere. Int. J. Comput. Vis. **48**, 233–254 (2002)
8. Schechner, Y.Y., Narasimhan, S.G., Nayar, S.K.: Instant dehazing of images using polarization. In: Proceedings of IEEE CVPR, pp. I:325–I:332 (2001)
9. Narasimhan, S.G., Nayar, S.K.: Contrast restoration of weather degraded images. IEEE Trans. Pattern Anal. Mach. Intell. **25**, 713–724 (2003)
10. Kopf, J., et al.: Deep photo: model-based photograph enhancement and viewing. ACM Trans. Graph. **27**, 116:1–116:10 (2008)
11. Tang, K., Yang, J., Wang, J.: Investigating haze-relevant features in a learning framework for image dehazing. In: Proceedings of IEEE CVPR, pp. 2995–3002 (2014)
12. Zhu, Q., Mai, J., Shao, L.: A fast single image haze removal algorithm using color attenuation prior. IEEE Trans. Image Process. **24**, 3522–3533 (2015)
13. Tan, R.T.: Visibility in bad weather from a single image. In: Proceedings of IEEE CVPR, pp. 1–8 (2008)
14. Tarel, J., Hautière, N.: Fast visibility restoration from a single color or gray level image. In: Proceedings of IEEE ICCV, pp. 2201–2208 (2009)
15. Nishino, K., Kratz, L., Lombardi, S.: Bayesian defogging. Int. J. Comput. Vis. **98**, 263–278 (2012)
16. Ancuti, C.O., Ancuti, C.: Single image dehazing by multi-scale fusion. IEEE Trans. Image Process. **22**, 3271–3282 (2013)
17. Li, Y., Tan, R.T., Brown, M.S.: Nighttime haze removal with glow and multiple light colors. In: 2015 IEEE ICCV 2015, Santiago, Chile, 7–13 December 2015, pp. 226–234 (2015)
18. Meng, G., Wang, Y., Duan, J., Xiang, S., Pan, C.: Efficient image dehazing with boundary constraint and contextual regularization. In: Proceedings of IEEE ICCV, pp. 617–624 (2013)
19. Carr, P., Hartley, R.I.: Improved single image dehazing using geometry. In: DICTA 2009, pp. 103–110 (2009)
20. He, K., Sun, J., Tang, X.: Guided image filtering. IEEE Trans. Pattern Anal. Mach. Intell. **35**, 1397–1409 (2013)
21. Marroquín, J.L., Velasco, F.A., Rivera, M., Nakamura, M.: Gauss-Markov measure field models for low-level vision. IEEE Trans. Pattern Anal. Mach. Intell. **23**, 337–348 (2001)

22. Li, Y.P., Huttenlocher, D.P.: Sparse long-range random field and its application to image denoising. In: Forsyth, D., Torr, P., Zisserman, A. (eds.) ECCV 2008. LNCS, vol. 5304, pp. 344–357. Springer, Heidelberg (2008). https://doi.org/10.1007/978-3-540-88690-7_26
23. Zhang, Q., Xu, L., Jia, J.: 100+ times faster weighted median filter (WMF). In: CVPR, pp. 2830–2837. IEEE (2014)
24. Levin, A., Lischinski, D., Weiss, Y.: A closed-form solution to natural image matting. IEEE Trans. Pattern Anal. Mach. Intell. **30**, 228–242 (2008)
25. Kim, M.H., Kautz, J.: Consistent tone reproduction. In: Proceedings of the IASTED International Conference on Computer Graphics and Imaging (CGIM 2008), Innsbruck, Austria, pp. 152–159. IASTED/ACTA Press (2008)
26. Kim, M.H., Kautz, J.: Consistent scene illumination using a chromatic flash. In: Proceedings of Eurographics Workshop on Computational Aesthetics (CAe 2009), British Columbia, Canada, Eurographics, pp. 83–89 (2009)