



A Novel Total Generalized Variation Model for Image Dehazing

Yanan Gu¹ · Xiaoping Yang² · Yiming Gao¹

Received: 23 October 2018 / Accepted: 14 September 2019 / Published online: 24 September 2019
© Springer Science+Business Media, LLC, part of Springer Nature 2019

Abstract

In this paper, we propose a new variational model for removing haze from a single input image. The proposed model combines two total generalized variation (TGV) regularizations, which are related to the image intensity and the transmission map, respectively, to build an optimization problem. Actually, TGV functionals are more appropriate for describing a natural color image and its transmission map with slanted plane. By minimizing the energy functional with double-TGV regularizations, we obtain the final haze-free image and the refined transmission map simultaneously instead of the general two-step framework. The existence and uniqueness of solutions to the proposed variational model are further obtained. Moreover, the variational model can be solved in a unified way by realizing a primal–dual method for associated saddle-point problems. A number of experimental results on natural hazy images are presented to demonstrate our superior performance, in comparison with some state-of-the-art methods in terms of the subjective and objective visual quality assessments. Compared with the total variation-based models, the proposed model can generate a haze-free image with less staircasing artifacts in the slanted plane and more details in the remote scene of an input image.

Keywords Dehaze · Atmospheric scattering model · TGV · Dark channel prior

1 Introduction

The phenomenon of haze, due to the presence of airborne particulate matter, absorbs the reflected light of the scene and scatters some atmospheric light to the camera [1]. As such, the visibility, contrast and vividness of the scene are dramatically degraded, which is a major threat to subsequent high-level computer vision tasks, such as object detection and recognition. Hence, it is important to design an image dehazing method to improve the environmental adaptability of the visual system.

In the literature, various methods have been proposed to handle the image dehazing problem. We can roughly classify these methods into two categories: multiple image dehazing

and single image dehazing. However, the haze removal from one single image has now gained the dominant popularity, since it is more practical for realistic settings. This paper focuses on the problem of single image dehazing. In computer vision, the formation of hazy image is usually described by the atmospheric scattering model [2]:

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

where \mathbf{I} is the observed hazy image intensity defined on the image domain $\Omega \subset \mathbf{R}^2$, \mathbf{J} is the scene radiance to be recovered, \mathbf{A} is the atmospheric light and assumed to be constant over the whole image, t is the medium transmission and x denotes the image coordinates. The transmission describes the portion of the light reaching the camera without being scattered. In a homogeneous atmosphere, the transmission t can be further expressed as $t(x) = e^{-\eta d(x)}$, where $d(x)$ is the distance from the scene point to the camera and η is the scattering coefficient of the atmosphere. The task of dehazing is to estimate \mathbf{J} (with \mathbf{A} and t as by-products) from the input image \mathbf{I} , which is a severely ill-posed problem. Therefore, haze removal is a challenging problem, especially for single image dehazing. Various prior knowledges have been proposed to capture deterministic (e.g., physical rules and variational energy) or statistical properties of the hazy image.

✉ Yanan Gu
guyanan0519@gmail.com

Xiaoping Yang
xpyang@nju.edu.cn

Yiming Gao
gao_yiming@hotmail.com

¹ School of Science, Nanjing University of Science and Technology, Nanjing 210094, People's Republic of China

² Department of Mathematics, Nanjing University, Nanjing 210093, People's Republic of China

For instance, Tan et al. [3] proposed a dehazing model based on the observation that the clear or enhanced images usually have higher contrast than hazy images. After that, Fattal [4] estimated the transmission through a refined image formation model, in which the cost functions of surface shading and transmission are considered to be statistically uncorrelated. He et al. [5] leveraged an empirical observation that the local patches in haze-free images often contain some low albedo values in at least one color channel. Inspired by this observation, they proposed a dark channel prior (DCP) to estimate the haze thickness. One may have the initial estimation of the transmission t (denoted t_0) from He et al.'s model:

$$t_0(x) = 1 - \mu \min_{c \in \{R, G, B\}} \left(\min_{y \in N(x)} \frac{I^c(y)}{A^c} \right),$$

where μ is used to keep some depth information of a natural image and is commonly set as $\mu = 0.95$, I^c is a color channel of the hazy image \mathbf{I} , the atmospheric light A_c can be estimated as the brightest pixel color in the dark channel (c denotes a color channel). This coarse transmission map is computed locally and thus often needs to be refined. He et al. utilized the soft-matting technique to reduce the haze effect. As the soft-matting algorithm is too complex, Tarel et al. [6] proposed a fast image restoration algorithm by using a median filtering and its variant to replace the soft-matting algorithm. Nowadays, many imaging problems are solved by variational methods. Meng et al. [7] reformulated DCP as a boundary constraint and incorporated a contextual regularization to model transmissions of hazy images. They found an optimal transmission function $t(x)$ by minimizing the following objective function:

$$\frac{\lambda}{2} \|t - t_0\|_2^2 + \sum_{j \in E} \|W_j \circ (D_j \otimes t)\|_1,$$

where λ is the regularization parameter for balancing the two terms, E is an index set, \circ represents the element-wise multiplication operator, \otimes stands for the convolution operator, D_j is a first-order differential operator and W_j ($j \in \omega$) is a weighting matrix. Later, Zhu et al. [8] created a linear model to estimate the scene depth under a novel color attenuation prior and learned the model parameters in HSV space. Although this approach can produce visually pleasing performance, it often fails to remove the dense haze in the remote scenes. The total variation (TV) and its adaptive term, such as weighted TV or vectorial TV, have been proved to be effective for edge preservation, and they have been applied in many image processing tasks, such as image denoising, image restoration, deblurring, filtering, and so on. Fang et al. [9] proposed a new variational model with weighted vectorial total variation (WVTV) to dehaze and denoise simultaneously. The minimizing problem is given as follows:

$$\min_{\mathbf{g}, d} E(\mathbf{g}, d) = \|h \nabla \mathbf{g}\|_{TV} + \lambda \|\nabla d\|_{TV} + \frac{1}{2} \|\mathbf{g} - \mathbf{f} - \mathbf{d}\|^2 + \frac{\gamma}{2} \|d - d_0\|^2,$$

where $d = -\frac{1}{\eta} \log t(x)$, $d_0 = -\frac{1}{\eta} \log t_0(x)$, $\mathbf{g} = \frac{1}{\eta} \log(\mathbf{A} - \mathbf{J})$, $\mathbf{f} = \frac{1}{\eta} \log(\mathbf{A} - \mathbf{I})$. The TV approach, however, also has some shortcomings, most notably the staircasing phenomenon. If one natural image contains not only flat but also slanted regions, then the image dehazing on the basis of the total variation tends to be piecewise constant (staircasing). The work in [10] aimed to solve a variational model based on image-guided total generalized variation (TGV) regularization to suppress artifacts in hazy images. They refined the transmission by using a global method based on image-guided total generalized variation (TGV) regularization:

$$\min_{t, \omega} \left\{ \alpha_1 \int |D^{1/2}(\nabla t - \omega)| dx + \alpha_0 \int |\nabla \omega| dx + \int |t - t_0| dx \right\},$$

where $D^{1/2}$ is the anisotropic diffusion tensor defined as:

$$D^{1/2} = \exp(-\gamma |\nabla I|^\beta) nn^T + n^\perp n^{\perp T},$$

where n is the direction of the gradient of the guided image $n = \frac{\nabla I}{|\nabla I|}$ and n^\perp is the perpendicular direction, γ , β are parameters to adjust the sharpness and magnitude of the tensor and ω is an auxiliary variable. After the transmission map was refined, they recovered the scene radiance \mathbf{J} by gradient residual minimization (GRM):

$$\min_{\mathbf{J}} \left\{ \frac{1}{2} \int \|\mathbf{J}t - (\mathbf{I} - \mathbf{A} + \mathbf{A}t)\|_2^2 dx + \eta \int \|\nabla \mathbf{J} - \nabla \mathbf{I}\|_0 dx \right\},$$

where the l_0 norm counts the number of nonzero elements and η is a weighting parameter. However, for very faraway objects, their method cannot significantly increase their contrast, which is due to the ambiguity between the artifacts and true objects covered by very thick haze, and mistakenly remove the distant objects to produce over-smoothed results. Lately, as convolutional neural networks (CNNs) have witnessed prevailing success in computer vision tasks, they have been introduced to image dehazing as well. Cai et al. [11] proposed a trainable model to estimate the transmission matrix from a hazy image called DehazeNet. Ren et al. [12] further exploited a multiscale CNN (MSCNN) that first generated a coarse-scale transmission matrix and later refined it.

In this paper, we propose a novel variational model with double-TGV regularizations related to $\mathbf{J}(x)$ and $t(x)$ for image dehazing and prove the existence and uniqueness of solutions of the proposed model. TGV regularization can

suppress artifacts in hazy images. We do not need to estimate the transmission in advance as the haze estimation step presented in [10]. We develop a primal–dual algorithm for optimal solutions by alternating updating the transmission map $t(x)$ and the scene radiance $\mathbf{J}(x)$ step-by-step. Moreover, the two variables affect each other in the calculating process.

The rest of our paper is organized as follows. In Sect. 2, we deduce the proposed model and prove the existence and uniqueness of the solution for the model. In Sect. 3, we present the primal–dual algorithm to solve our model. In Sect. 4, we give some experimental results on natural images and illustrate the superior performance of our model. In Sect. 5, we provide concluding remarks.

2 The Proposed Model

In this section, we introduce our variational model for recovering the haze-free image without estimating the refined transmission in advance.

(1) can be rewritten as

$$A_c - I_c(x) = t(x)(A_c - J_c(x)), \tag{2}$$

where c denotes the color channel. In order to handle the product expression in (2), we convert it into the logarithmic domain, i.e.,

$$\log(A_c - I_c(x)) = \log t(x) + \log(A_c - J_c(x)).$$

As we know, t is an exponential function of the depth map d , so we have $d(x) = -\frac{1}{\eta} \log t(x)$, where again η is the scattering coefficient and, for simplicity, we set it to be 1 for our experiments. We further set

$$f_c(x) = \frac{1}{\eta} \log(A_c - I_c(x)), \quad g_c(x) = \frac{1}{\eta} \log(A_c - J_c(x)).$$

Then, Eq. (2) is equivalent to

$$\mathbf{g}(x) = \mathbf{f}(x) + \mathbf{d}(x), \quad x \in \Omega$$

where $\mathbf{g} = \{g_R, g_G, g_B\}$, $\mathbf{f} = \{f_R, f_G, f_B\}$, $\mathbf{d}(x) = \{d, d, d\}$.

In order to deal with color images, it is convenient to have a notion of TGV for vector-valued images $\mathbf{g} : \Omega \subset \mathbf{R}^2 \rightarrow \mathbf{R}^L$ for some $L \geq 1$. Here, $L = 3$ for the RGB color space. A vector-valued version of $\text{TGV}_\alpha^k(\mathbf{g})$ can then be defined as

$$\begin{aligned} \text{TGV}_\alpha^k(\mathbf{g}) = \sup \left\{ \int_\Omega \sum_{l=1}^3 \mathbf{g}_l (\text{div}^k \mathbf{v}_l) \, dx \mid \mathbf{v} \right. \\ \left. \in \mathcal{C}_c^k(\Omega, \text{Sym}^k(\mathbf{R}^2)^3), \right. \\ \left. \|\text{div}^k \mathbf{v}\|_\infty \leq \alpha_\kappa, \kappa = 0, \dots, k-1 \right\}, \end{aligned}$$

where \mathbf{g}_l is a color channel of the image \mathbf{g} , \mathbf{v}_l is a component of \mathbf{v} , $\text{Sym}^k(\mathbf{R}^2) = \{\xi : \underbrace{\mathbf{R}^2 \times \dots \times \mathbf{R}^2}_{k \text{ times}} \rightarrow \mathbf{R} \mid \xi \text{ multilinear and symmetric}\}$ denotes the space of symmetric tensors on \mathbf{R}^2 (for $k = 1$, $\text{Sym}^1(\mathbf{R}^2) = \mathbf{R}^2$; and for $k = 2$, $\text{Sym}^2(\mathbf{R}^2) = S^{2 \times 2}$) and α_κ are fixed positive parameters.

The vector-valued space $\text{BGV}_\alpha^k(\Omega, \mathbf{R}^3)$ is defined as follows:

$$\text{BGV}_\alpha^k(\Omega, \mathbf{R}^3) = \{\mathbf{g} \in L^1(\Omega, \mathbf{R}^3) \mid \text{TGV}_\alpha^k(\mathbf{g}) < \infty\}.$$

The scalar version of TGV_β^k for d can be defined as same as the vector-valued version of TGV_α^k with $l = 1$, and the scalar space BGV is defined as $\text{BGV}_\beta^k(\Omega)$.

For $k = 1$, $\alpha_0 = 1$, TGV_α^k coincides with TV and the space BGV_α^k coincides with space BV. Indeed, solutions of variational problems with TV regularization admit many desirable properties, most notably the appearance of sharp edges. Unfortunately, one can also observe typical artifacts which are associated with the regularization with TV. The most prominent of these artifacts are the so-called staircasing effect and blocky effect. This is a side effect of the model assumption that an image consists of piecewise constant. Natural images are, however, often piecewise smooth due to shading, for instance. Compared with the conventional TV, total generalized variation (TGV) prefers piecewise smooth images. This is also a desired property for the transmission, because we may have a slanted plane (e.g., road, bridge) whose transmission varies smoothly along with the change of depth. As a consequence, we use TGV regularization terms for a hazy image and its transmission map both.

We thus make the following assumptions:

- We suppose functions $d(x)$ and $\mathbf{g}(x)$ are piecewise smooth in Ω . Thus, $\text{TGV}_\beta^k(d)$ and $\text{TGV}_\alpha^k(\mathbf{g})$ can be set as the regularization terms of d and \mathbf{g} .
- The sum of $\mathbf{d}(x)$ and $\mathbf{f}(x)$ is close to $\mathbf{g}(x)$ in the sense of square norm. Thus, the fidelity terms are given by $\int_\Omega (\mathbf{g} - \mathbf{d} - \mathbf{f})^2 dx$ and $\int_\Omega (d - d_0)^2 dx$.

Combining the above two assumptions, we propose the following energy functional for recovering a haze-free image:

$$E(\mathbf{g}, d) = \text{TGV}_\alpha^k(\mathbf{g}) + \frac{1}{2} \int_\Omega (\mathbf{g} - \mathbf{f} - \mathbf{d})^2 dx + \text{TGV}_\beta^k(d) + \frac{\gamma}{2} \int_\Omega (d - d_0)^2 dx, \tag{3}$$

where $d_0 = -\log t_0$, and γ is a positive parameter.

Then, the total space Λ of our energy functional can be defined as

$$\Lambda := \{(\mathbf{g}, d) | (\mathbf{g}, d) \in \text{BGV}_\alpha^k(\Omega, \mathbf{R}^3) \times \text{BGV}_\beta^k(\Omega)\}.$$

After the functional space has been chosen, the minimizing problem of (3) can be written as the following form:

$$\inf_{(\mathbf{g}, d) \in \Lambda} E(\mathbf{g}, d). \tag{4}$$

In the following, we will discuss the existence and uniqueness of the global minimizer for the proposed model. It is easy to show the lower semicontinuity of TGV, and the proof can be seen in [13].

Lemma 1 (Lower semicontinuity) *If the sequences $\mathbf{g}^n \subset \text{BGV}_\alpha^k(\Omega, \mathbf{R}^3)$ and $\mathbf{g}^* \in \text{BGV}_\alpha^k(\Omega, \mathbf{R}^3)$ satisfy that $\mathbf{g}^n \rightarrow \mathbf{g}^*$ in $L^1(\Omega, \mathbf{R}^3)$, i.e.,*

$$\lim_{n \rightarrow +\infty} \int_\Omega |\mathbf{g}^n - \mathbf{g}^*| dx = 0,$$

then

$$\text{TGV}_\alpha^k(\mathbf{g}^*) \leq \liminf_{n \rightarrow +\infty} \text{TGV}_\alpha^k(\mathbf{g}^n).$$

Theorem 1 *Assume that $\mathbf{f} \in L^2(\Omega, \mathbf{R}^3)$ and that $d_0 \in L^2(\Omega)$ is nonnegative, then the minimization problem (4) admits a unique solution $(\mathbf{g}^*, d^*) \in \Lambda$.*

Proof First, the energy functional $E(\mathbf{g}, d)$ is clearly non-negative and proper because $E(0, 0)$ is a finite value. Let $\{(\mathbf{g}^n, d^n)\}$ be a minimizing sequences of problem (4). Then, there is a constant $M > 0$, such that

$$E(\mathbf{g}^n, d^n) = \text{TGV}_\alpha^k(\mathbf{g}^n) + \frac{1}{2} \int_\Omega (\mathbf{g}^n - \mathbf{f} - \mathbf{d}^n)^2 dx + \text{TGV}_\beta^k(d^n) + \frac{\gamma}{2} \int_\Omega (d^n - d_0)^2 dx \leq M.$$

Thus,

$$\text{TGV}_\alpha^k(\mathbf{g}^n) \leq M, \quad \text{TGV}_\beta^k(d^n) \leq M. \tag{5}$$

Since $d_0 \in L^2(\Omega)$, $\mathbf{f} \in L^2(\Omega, \mathbf{R}^3)$, we can assume that

$$\|d_0\|_2 \leq M, \quad \|\mathbf{f}\|_2 \leq M.$$

Observing that

$$\frac{\gamma}{2} \int_\Omega |d^n - d_0|^2 dx \leq M \quad \text{and} \\ \frac{1}{2} \int_\Omega |\mathbf{g}^n - \mathbf{f} - \mathbf{d}^n|^2 dx \leq M,$$

we have

$$\|d^n\|_2^2 = \|(d^n - d_0) + d_0\|_2^2 \\ \leq 2\|d^n - d_0\|_2^2 + 2\|d_0\|_2^2 \\ \leq \frac{4}{\gamma}M + 2M^2$$

and

$$\|\mathbf{g}^n\|_2^2 = \|(\mathbf{g}^n - \mathbf{f} - \mathbf{d}^n) + \mathbf{f} + \mathbf{d}^n\|_2^2 \\ \leq 3\|\mathbf{g}^n - \mathbf{f} - \mathbf{d}^n\|_2^2 + 3\|\mathbf{f}\|_2^2 + 3\|\mathbf{d}^n\|_2^2 \\ \leq 6M + \frac{12}{\gamma}M + 9M^2.$$

Then, the boundedness of $\|d^n\|_1$ and $\|\mathbf{g}^n\|_1$ is automatically obtained by Holder’s inequality:

$$\|d^n\|_1^2 \leq |\Omega| \|d^n\|_2^2 \leq |\Omega| \left(\frac{4}{\gamma}M + 2M^2 \right), \\ \|\mathbf{g}^n\|_1^2 \leq |\Omega| \|\mathbf{g}^n\|_2^2 \leq |\Omega| \left(6M + \frac{12}{\gamma}M + 9M^2 \right),$$

where γ is a positive parameter. Namely, the sequences $\{\mathbf{g}^n\}$ and $\{d^n\}$ are bounded in $L^1(\Omega)$ and $L^1(\Omega, \mathbf{R}^3)$, respectively. Combining it with (5), we can conclude that the sequences $\{\mathbf{g}^n\}$ and $\{d^n\}$ are bounded in $\text{BGV}_\alpha^k(\Omega, \mathbf{R}^3)$ and $\text{BGV}_\beta^k(\Omega)$, respectively. By Sobolev imbedding theorems, there exist subsequences (\mathbf{g}^n, d^n) such that

$$d^n \rightarrow d^* \quad \text{and} \quad \mathbf{g}^n \rightarrow \mathbf{g}^* \quad \text{strongly in } L^1(\Omega), \\ d^n \rightarrow d^* \quad \text{and} \quad \mathbf{g}^n \rightarrow \mathbf{g}^* \quad \text{weakly in } L^2(\Omega), \\ d^n \xrightarrow{\omega^*} d^* \quad \text{and} \quad \mathbf{g}^n \xrightarrow{\omega^*} \mathbf{g}^* \quad \text{in BGV},$$

where $(\mathbf{g}^*, d^*) \in \text{BGV}_\alpha^k(\Omega, \mathbf{R}^3) \times \text{BGV}_\beta^k(\Omega)$. By the weakly lower semicontinuity of the L^2 and Lemma 1, a standard process can prove that

$$\min_{(\mathbf{g}, d) \in \Lambda} E(\mathbf{g}, d) = \liminf_{n \rightarrow +\infty} E(\mathbf{g}^n, d^n) \geq E(\mathbf{g}^*, d^*).$$

Hence, (\mathbf{g}^*, d^*) is a minimal point of $E(\mathbf{g}, d)$.

All the terms in (3) are convex, so E is convex. (Here, \mathbf{g} is a three-dimensional vector value and d is one-dimensional value.) Moreover, the Hessian matrix of the functional $E_0(\mathbf{g}, d) := \frac{1}{2} \int_\Omega (\mathbf{g} - \mathbf{f} - \mathbf{d})^2 dx + \frac{\gamma}{2} \int_\Omega (d - d_0)^2 dx$ is

$$H = \begin{bmatrix} \frac{\partial^2 E_0}{\partial \mathbf{g}^2} & \frac{\partial^2 E_0}{\partial \mathbf{g} \partial d} \\ \frac{\partial^2 E_0}{\partial d \partial \mathbf{g}} & \frac{\partial^2 E_0}{\partial d^2} \end{bmatrix} = \begin{bmatrix} 1 & -3 \\ -1 & 3 + \gamma \end{bmatrix}.$$

H is positive definite; thus, E_0 is strictly convex. As a result, E is strictly convex, too.

Therefore, we can conclude that the variational problem (4) admits a unique solution. \square

3 The Algorithm

To solve the minimizing problem (4), we apply the primal–dual algorithm. The present paper addresses this issue by analyzing the case of $k = 2$, i.e., total generalized variation of second order.

With the constraint $\nabla^T q = p$, we can now deduce a discrete version of $\text{TGV}_\alpha^2(u)$:

$$\text{TGV}_\alpha^2(u) = \max \left\{ \langle u, \nabla^T p \rangle_{\mathbf{R}^L} \mid (p, q) \in (\mathbf{R}^2)^L \times \text{Sym}^2(\mathbf{R}^2)^L, \nabla^T q = p, \|q\|_\infty \leq \alpha_0, \|p\|_\infty \leq \alpha_1 \right\},$$

The inner products we used here are defined as follows:

$$\text{for } u, r \in \mathbf{R}^L : \langle u, r \rangle_{\mathbf{R}^L} = \sum_{l=1}^L \langle u^l, r^l \rangle,$$

$$\begin{aligned} \text{for } u, r \in (\mathbf{R}^2)^L : \langle u, r \rangle_{(\mathbf{R}^2)^L} \\ = \sum_{l=1}^L (\langle (u^l)^1, (r^l)^1 \rangle + \langle (u^l)^2, (r^l)^2 \rangle), \end{aligned}$$

$$\begin{aligned} \text{for } u, r \in \text{Sym}^2(\mathbf{R}^2)^L : \\ \langle u, r \rangle_{\text{Sym}^2(\mathbf{R}^2)^L} = \sum_{l=1}^L (\langle (u^l)^{11}, (r^l)^{11} \rangle \\ + \langle (u^l)^{22}, (r^l)^{22} \rangle + 2\langle (u^l)^{12}, (r^l)^{12} \rangle), \end{aligned}$$

and $\langle \cdot, \cdot \rangle$ denotes the Euclidean inner product.

Introducing indicator functionals, i.e.,

$$\mathcal{I}_K(x) = \begin{cases} 0 & \text{if } x \in K \\ \infty & \text{else} \end{cases}$$

and observing that

$$-\mathcal{I}_{\{0\}}(\nabla^T q - p) = \min_{\omega \in (\mathbf{R}^2)^L} \langle \omega, \nabla^T q - p \rangle_{(\mathbf{R}^2)^L},$$

the discrete functional can be rewritten as

$$\begin{aligned} \text{TGV}_\alpha^2(u) = \max_{p, q} \min_{\omega \in (\mathbf{R}^2)^L} \langle u, \nabla^T p \rangle_{\mathbf{R}^L} + \langle \omega, \nabla^T q - p \rangle_{(\mathbf{R}^2)^L} \\ - \mathcal{I}_{\{\|\cdot\| \leq \alpha_0\}}(q) - \mathcal{I}_{\{\|\cdot\| \leq \alpha_1\}}(p). \end{aligned}$$

One can show that the maximum and minimum can be interchanged. Moreover, the constraints are symmetric around 0, so we have

$$\begin{aligned} \text{TGV}_\alpha^2(u) = \min_{\omega} \max_{p, q} \langle \nabla u - \omega, p \rangle_{(\mathbf{R}^2)^L} + \langle \mathcal{E}\omega, q \rangle \\ - \mathcal{I}_{\{\|\cdot\| \leq \alpha_0\}}(q) - \mathcal{I}_{\{\|\cdot\| \leq \alpha_1\}}(p), \end{aligned}$$

where \mathcal{E} is the symmetric derivative $\mathcal{E}\omega = \frac{1}{2}(\nabla\omega + \nabla\omega^T)$.

For $L = 1$, the discrete scalar version of $\text{TGV}_\beta^2(d)$ can be defined; for $L = 3$, the discrete vector-valued version of $\text{TGV}_\alpha^2(\mathbf{g})$ can be defined.

Then, a discretization of the variational problem (4) is given by the saddle-point problem

$$\begin{aligned} \min_{\mathbf{g}, d, \omega_1, \omega_2, \mathbf{p}_1, p_2, \mathbf{q}_1, q_2} \max \langle \nabla \mathbf{g} - \omega_1, \mathbf{p}_1 \rangle_{(\mathbf{R}^2)^3} + \langle \mathcal{E}\omega_1, \mathbf{q}_1 \rangle_{\text{Sym}^2(\mathbf{R}^2)^3} \\ + \frac{1}{2} \int_{\Omega} (\mathbf{g} - \mathbf{f} - \mathbf{d})^2 dx + \langle \nabla d - \omega_2, p_2 \rangle_{\mathbf{R}^2} \\ + \langle \mathcal{E}\omega_2, q_2 \rangle_{\text{Sym}^2(\mathbf{R}^2)} + \frac{\gamma}{2} \int_{\Omega} (d - d_0)^2 dx \\ - I_{\{\|\cdot\| \leq \alpha_1\}}(\mathbf{p}_1) - I_{\{\|\cdot\| \leq \alpha_2\}}(\mathbf{q}_1) - I_{\{\|\cdot\| \leq \beta_1\}}(p_2) \\ - I_{\{\|\cdot\| \leq \beta_2\}}(q_2). \end{aligned} \tag{6}$$

Note that (6) can be solved by many efficient methods, such as the augmented Lagrangian method and split Bregman iteration. Here, we chose to employ the primal–dual method due to its simplicity which is summarized in the following:

Algorithm. The computational procedure of the algorithm for the proposed model.

Initialization : $\mathbf{g}^0 = \mathbf{g}^0 = \mathbf{f} + \mathbf{d}_0, \bar{d}^0 = d^0 = d_0, \bar{\omega}_1^0 = \omega_1^0 = \nabla \mathbf{g}^0,$
 $\bar{\omega}_2^0 = \omega_2^0 = \nabla d^0, \mathbf{p}_1^0, p_2^0, \mathbf{q}_1^0, q_2^0 = 0$

for $k = 0$ **to** Maxiteration **do**

$\mathbf{p}_1^{k+1} = \mathcal{P}_{\alpha_1}[\mathbf{p}_1^k + \tau(\nabla \mathbf{g}^k - \bar{\omega}_1^k)],$
 $p_2^{k+1} = \mathcal{P}_{\beta_1}[p_2^k + \tau(\nabla \bar{d}^k - \bar{\omega}_2^k)],$
 $\mathbf{q}_1^{k+1} = \mathcal{P}_{\alpha_2}[\mathbf{q}_1^k + \tau \mathcal{E} \bar{\omega}_1^k],$
 $q_2^{k+1} = \mathcal{P}_{\beta_2}[q_2^k + \tau \mathcal{E} \bar{\omega}_2^k],$
 $d^{k+1} = \frac{(\sigma+1)(-\Sigma f_c + \gamma d_0 + \frac{1}{\sigma} d^k + \nabla^T p_2^{k+1}) + \Sigma(\sigma f_c + g_c^k + \sigma \nabla^T p_1^{k+1})}{\gamma(1+\sigma) + \frac{1}{\sigma} + 4},$
 $\mathbf{g}^{k+1} = \frac{\sigma(\mathbf{f} + \mathbf{d}^{k+1} + \nabla^T \mathbf{p}_1^{k+1}) + \mathbf{g}^k}{1 + \sigma},$
 $\omega_1^{k+1} = \omega_1^k + \sigma(\nabla^T \mathbf{q}_1^{k+1} + \mathbf{p}_1^{k+1}),$
 $\omega_2^{k+1} = \omega_2^k + \sigma(\nabla^T q_2^{k+1} + p_2^{k+1}),$
 $\bar{\mathbf{g}}^{k+1} = 2\mathbf{g}^{k+1} - \mathbf{g}^k,$
 $\bar{d}^{k+1} = 2d^{k+1} - d^k,$
 $\bar{\omega}_1^{k+1} = 2\omega_1^{k+1} - \omega_1^k,$
 $\bar{\omega}_2^{k+1} = 2\omega_2^{k+1} - \omega_2^k,$

end for

In the algorithm, $\mathbf{d}^{k+1} = (d^{k+1}, d^{k+1}, d^{k+1})$, $\tau, \sigma > 0$ are the step sizes and k is the iteration number. The project operators \mathcal{P} is defined:

$$\mathcal{P}[x] = \frac{x}{\max\{1, |x|\}}.$$

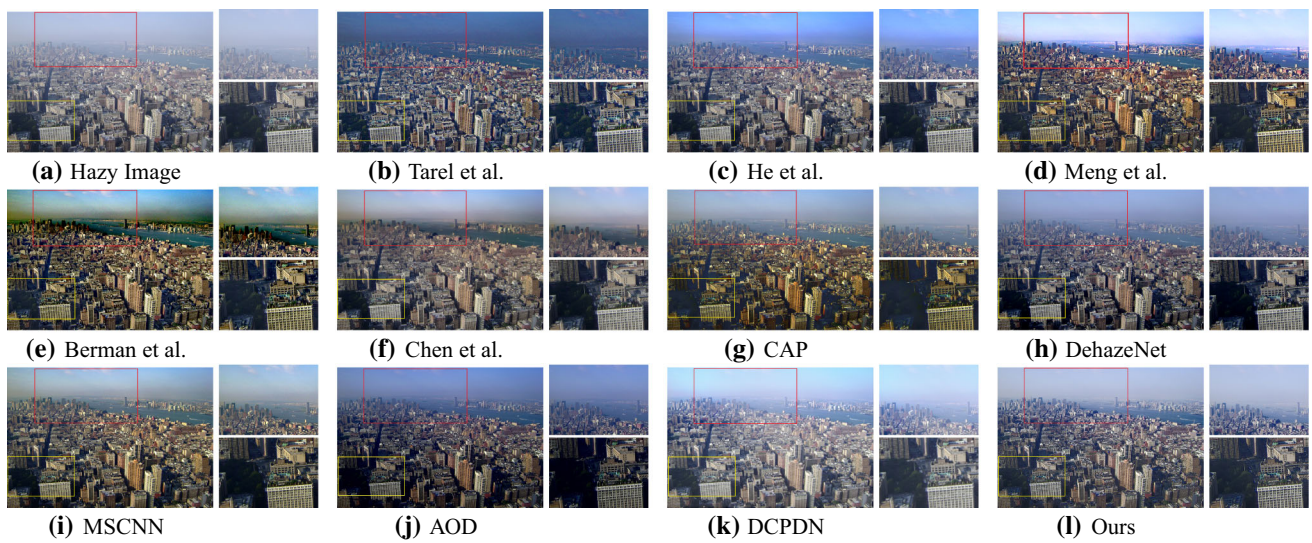


Fig. 1 Thin haze removal

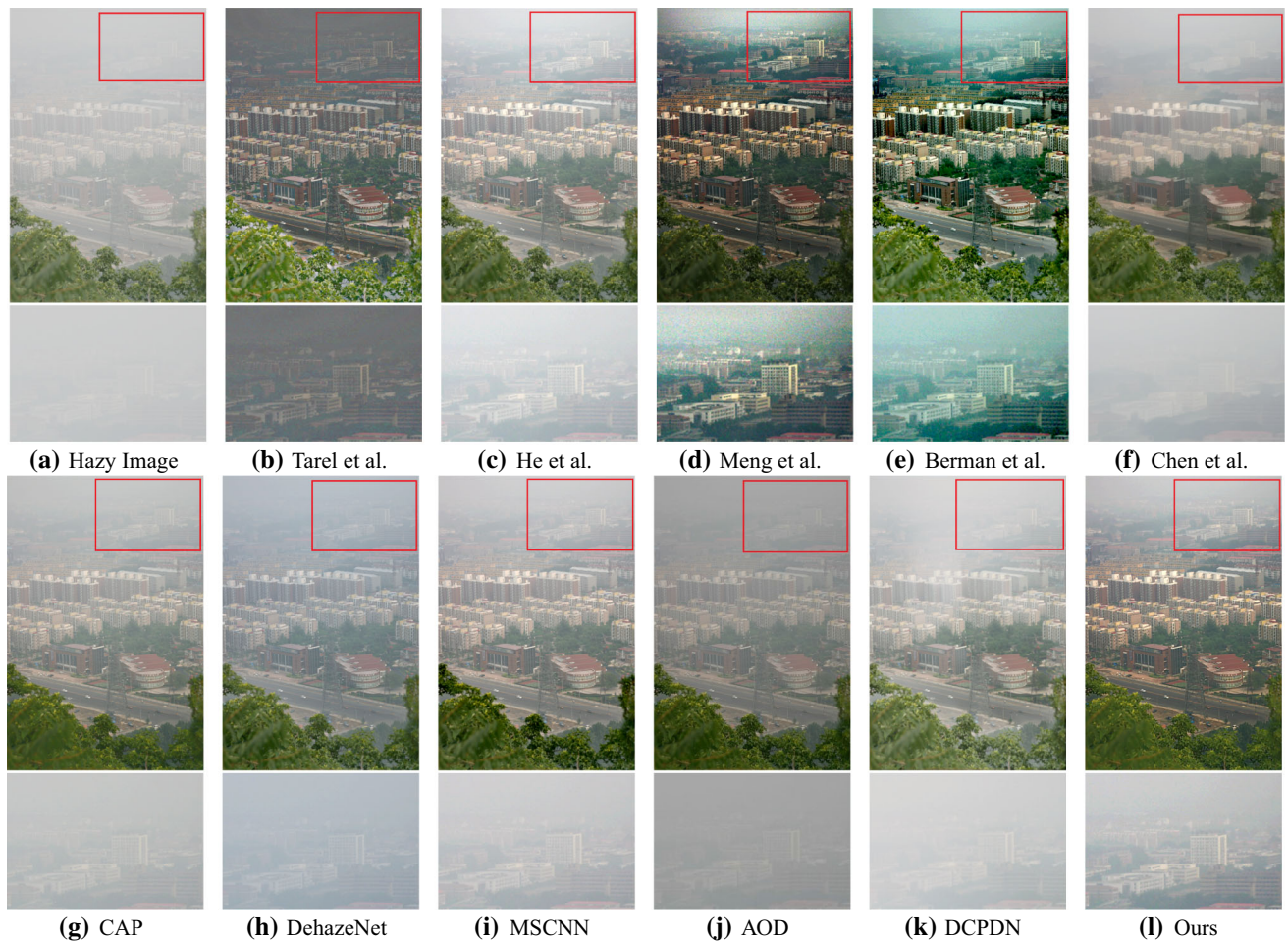


Fig. 2 Dense haze removal

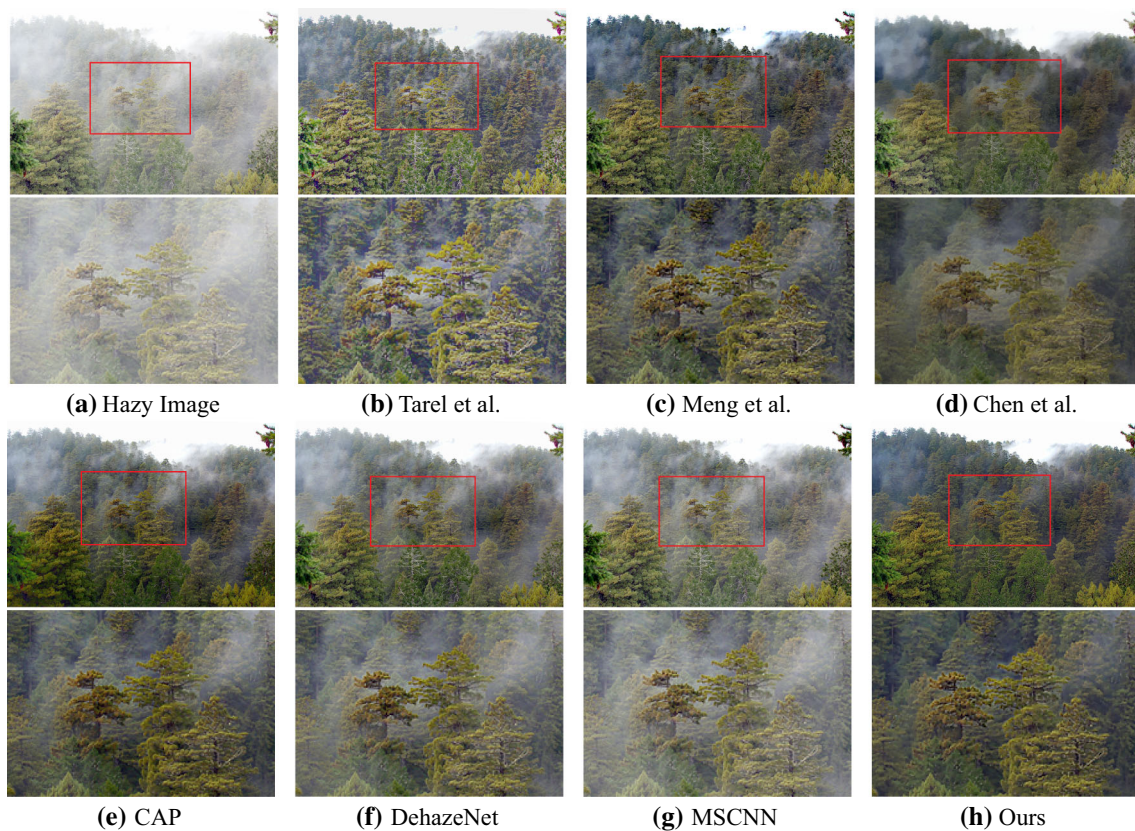


Fig. 3 Inhomogeneous haze removal

The divergence and gradient operators in the optimization are approximated by using standard finite difference. More details are defined as in [13].

4 Experimental Results

In this section, we present some experimental results to show the performance and the effectiveness of our model, in comparison with Tarel et al.'s model [6], DCP model [5], Meng et al.'s model [7], Berman et al.'s model [14], Chen et al.'s model [10], CAP model [8], DehazeNet [11], MSCNN [12], AOD-Net [15] and DCPDN [16]. The first five conventional methods are proposed to refine the coarse transmission map obtained by dark channel prior, and the others are some state-of-the-art of learning methods nowadays. Note that the most of the following experiments are implemented in MATLAB R2015b on an Intel 2.50 GHz PC with 8 GB RAM; especially, AOD-Net and DCPDN are implemented in Pycaffe on an Intel PC with GTX1064 GPU. In our model, the parameters α and β adjust the importance of the two TGV regularization terms. We observe that the results of the proposed model change very little as the parameters α and β vary in ranges [0.01, 0.1] and

[0.2, 0.8]. The parameter γ determines the fidelity term; as a result, we cannot set γ to be too small. In our experiment, we fix γ in [2, 4]. We choose the parameters $\alpha = 0.02$, $\beta = 0.5$ and $\gamma = 4$ in our manuscript by inferring on best visual appearance after taking many times experiments. For other comparison models, we use the parameter ranges given in their corresponding papers (or programs) and choose the values to give the best visual effect.

4.1 Comparison to the State-of-Art Methods

The haze of the input image in Fig. 1 is thin, and the scene depth of it is varying continuously. The dehazed result of Tarel et al.'s model [6] has oversaturated color in the sky region, and we can see this unpleasant effect in the close-up view. To some extent, the dehazed result of Meng et al.'s model [7] can remove the haze of the image, but it suffers from color distortion owing to the error estimation of the atmospheric light. Besides, the results of Berman et al.'s model [14], Chen et al.'s model [10], CAP model [8] and MSCNN [12] also suffer from unpleasing color shifting. The result of DCP model [5] still contains haze in the distant horizon. The dehazed results of CAP model [8], DehazeNet [11],

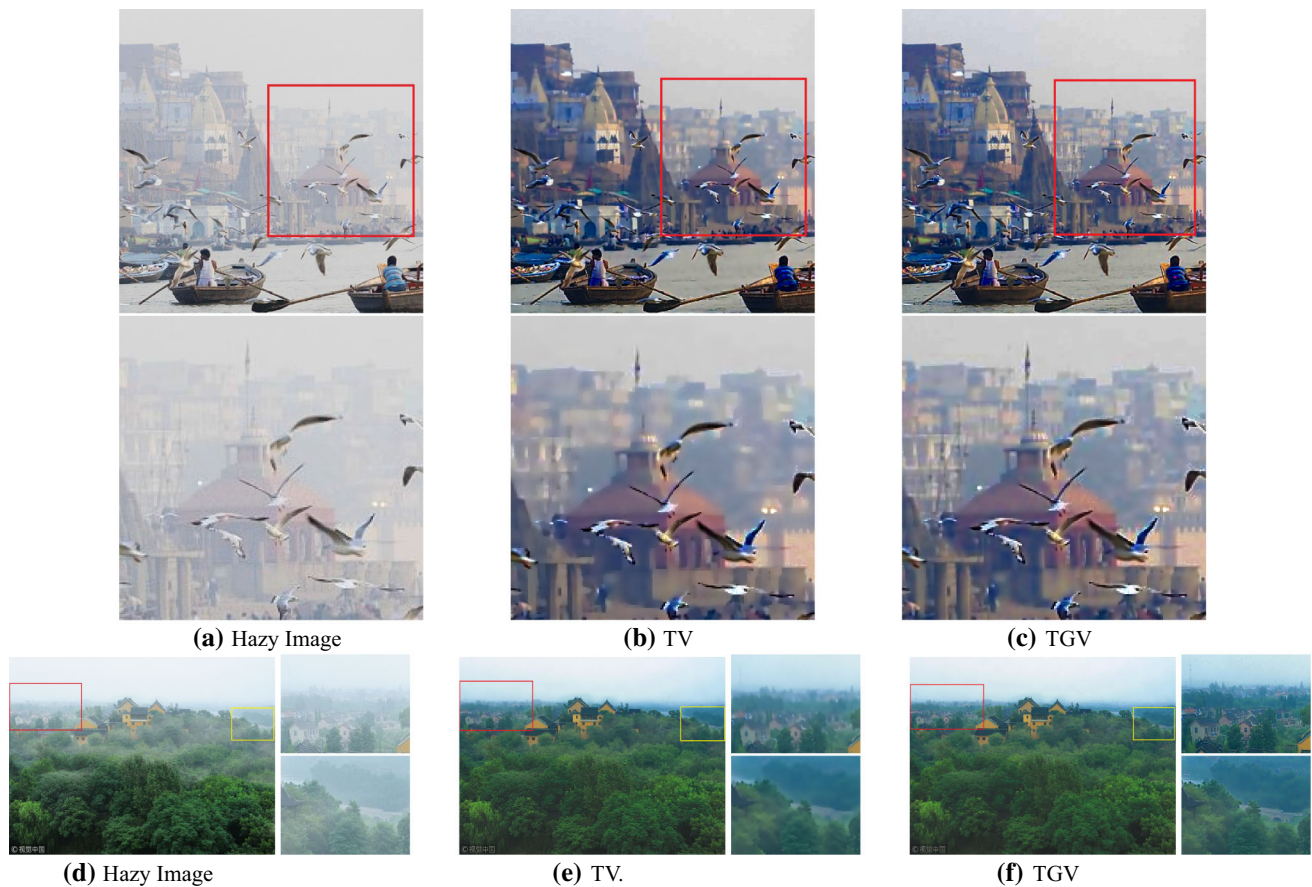


Fig. 4 A comparison between the TV model and the TGV model

MSCNN [12] and AOD-Net [15] are a little dark and lack of enhancement in some areas (see the close-up view in the yellow rectangle). DCPDN's [11] result and our result contain much less visual artifacts and appear to be more natural.

In Fig. 2, the image suffers from dense haze and has discontinuous scene depth. We can observe that edge artifacts around scene depth become severe in Tarel et al. model [6]'s result. The result of Chen et al.'s model [10] is quite over-smoothed for the distant objects. The results of Meng et al.'s model [7] and Berman et al.'s model [14] have the compelling visibility, but they suffer from color distortion which makes the image appear to be unnatural. The dehazed results of the learning-based methods are enhanced, but not as clear as ours. The result of our method is comparable to the DCP [5] method's, but our result has a higher image contrast. Compared with the other results, our dehazed image is more balanced in the visibility and the naturalness.

In Fig. 3, the image is full of inhomogeneous haze and discontinuous depth. For this input, most existing dehazing methods' results are visual unnoticeable, especially in heavy haze regions. The haze cannot be effectively removed using the previous methods, even the learning methods nowadays, such as CAP [8], DehazeNet [11] and MSCNN [12]. This is

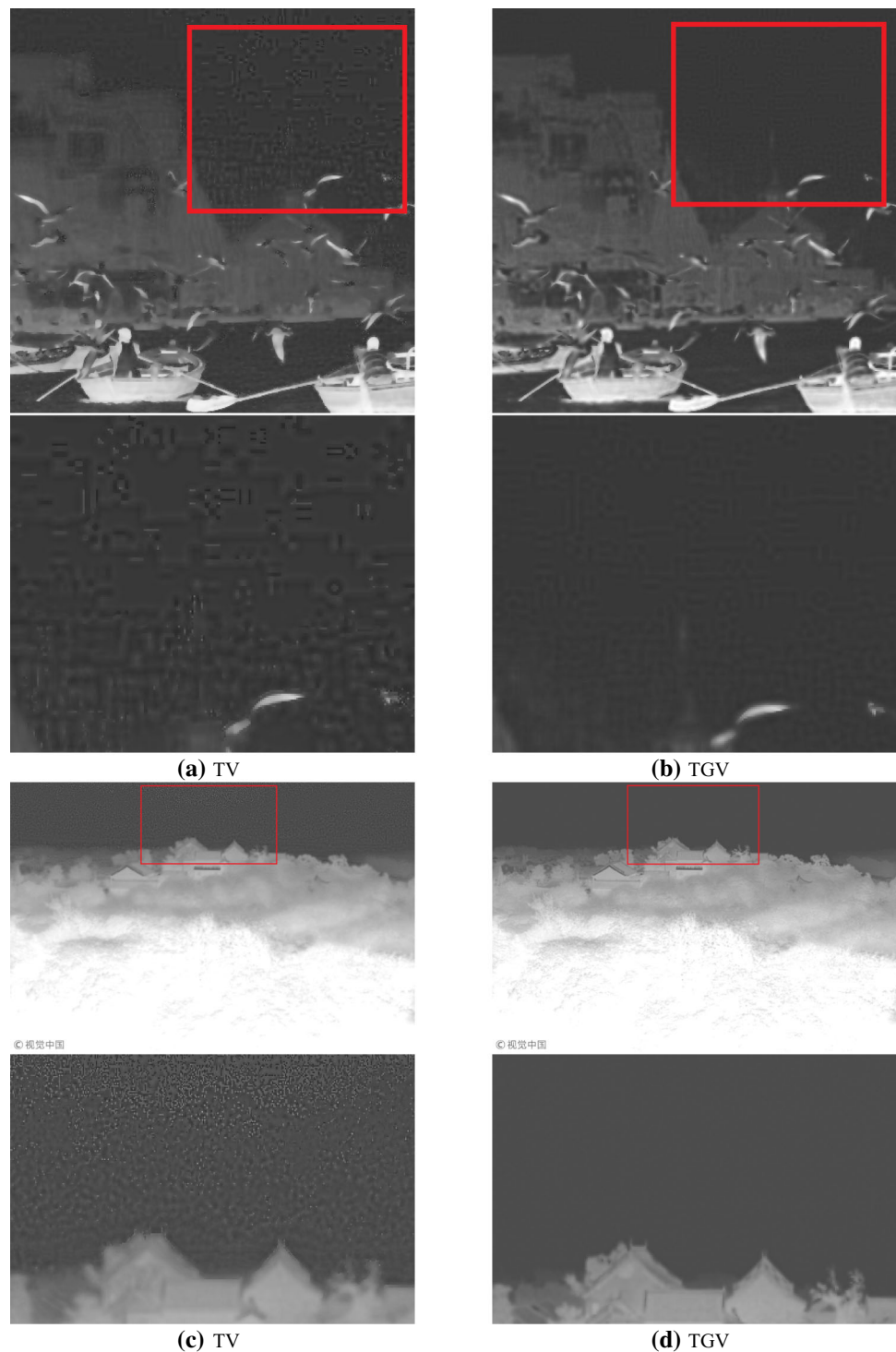
because, during the dataset training process, the haze image is synthesized homogeneously. Our result can remove more haze in the image with less visual artifact, but the color of our image is prone to dark. However, a better result may be obtained by further enhancing its brightness via exposure compensation. We can see the close-ups from the boxed areas in the bottom.

4.2 Comparison to TV-Based Models

We further compare our method with the TV-based dehazing model [9]. Two examples are shown in Fig. 4, where the input images have lower quality in the remote scenes. We can see that fine details in the remote scenes are lost in TV model's results such as the buildings in the red square, the trees, houses and telegraph pole in the red rectangle. More apparently, the bridge (the close-up view in the yellow rectangle) in the dehazed image using TV model is missing. As we know, the kernel of TGV_{α}^2 reads as

$$\ker(TGV_{\alpha}^2) = \{u(x) = Ax + b \mid A \in \mathbf{R}^{3 \times 2}, b \in \mathbf{R}^3\},$$

Fig. 5 A comparison between the TV model's transmission map and the TGV model's transmission map

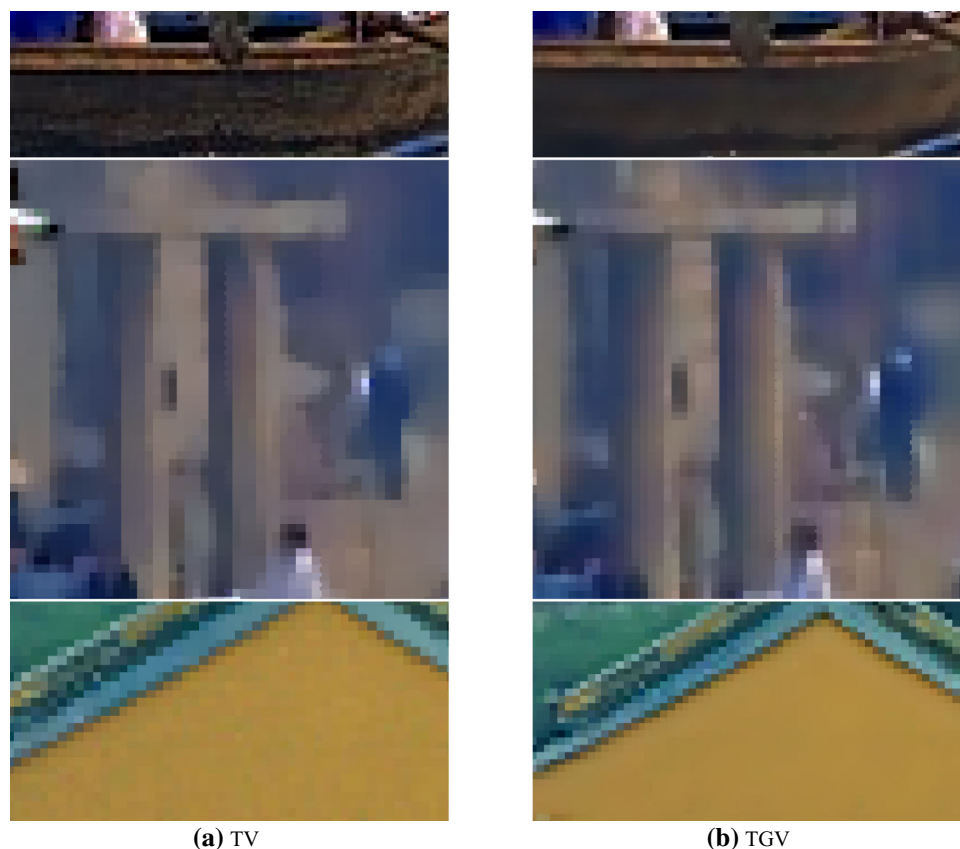


but the kernel of TV consists of constant functions. For the TGV model, the dehazed image can be represented as e^{Ax+b} roughly in a local region, while e^c for the TV model, which implies that our TGV model allows the image change linearly in this local region. This reveals that the range of the color value is larger in the TGV model's result than in the

TV model's; in other words, more color and content will be preserved.

In Fig. 5, we compare the transmission maps obtained by the TV model and our model. We can see that there are strong visible blocky artifacts appearing in the sky region of the TV model's results which are alleviated after dehazing by our TGV model. As we know, the transmission cannot be well

Fig. 6 A close-up comparison between the TV model and the TGV model



approximated by piecewise constant functions when it varies smoothly along with the change in depth.

The superiority of the TGV model is more apparent in Fig. 6, in which close-up views of Fig. 4 are displayed. From the images, we observe that the hull, the brown pillar and the yellow wall recovered by the TV model suffer more staircasing effect than by our model. Meanwhile, the top of the brown building and the water surface of the TV model's result in Fig. 4 are over-smoothed. This eliminates the possibility that the weights of the TV regularization terms are too small. The staircasing effect is produced by the TV model itself due to its inaccurate assumption of scene depth and surface radiance being piecewise constant.

4.3 Objective Quality Evaluations

A good image quality assessment method needs to compare the effect of visibility, color restoration and image structure similarity of different dehazing algorithms. As shown in Table 1, no-reference image quality assessments [17] are used here, because the full-reference and reduced-reference image quality assessments need a clear image corresponding to the hazy image to act as the reference image. This is hard to be satisfied in real applications.

In Table 1, we compare the quality assessments of the dehazing methods with those of our method, such as IVM, VCM, Cg, UQI and so on. The definitions of these objective image quality assessment are described in detail in the appendix. According to the subjective quality assessments in Sect. 4.14.1, we can find that DehazeNet, MSCNN and CAP obtain the satisfactory visual results with respect to Figs. 1, 2 and 3, respectively, so in Table 1 we take the three methods to compare with our method. As shown in the table, our model achieves better values of most visibility assessment criteria than the other methods. In some sense, when all the other values of assessments are equal, a smaller SSIM indicates less haze in the dehazed image. Thus, in this case, the SSIM here is only a reference for the readers. (There is no truth image in real-world applications of image dehazing and the original hazy image is always chosen as the reference image, so large SSIM and UQI do not mean the image is of better dehazed result. For example, the SSIM index of two identical hazy images must be larger than the SSIM index of the hazy image and the dehazed image.) Compared with the result of CAP, the color of our result is a little dark. So the HCC values of ours shown in the table are smaller than CAP's. Next, our algorithm achieves better values, on average, than MSCNN and DehazeNet algorithms. There is no dehazing method that has very good performance under all

Table 1 Blind quantitative analysis comparison of the results

Figure	Method	IVM	VCM	SSIM	Cg	UQI	HCC
Figure 1	DehazeNet [11]	8.0128	70.1923	0.8996	0.3402	0.8229	0.2022
	Ours	8.3739	75.1923	0.8457	0.2861	0.8805	0.6304
Figure 2	MSCNN [12]	5.2842	45.8333	0.8299	0.2123	0.9174	0.3749
	Ours	6.5800	44.5000	0.6875	0.2429	0.8488	0.1474
Figure 3	CAP [8]	8.9442	50	0.7345	0.3630	0.6854	− 0.1434
	Ours	10.9083	71.9642	0.4067	1.1900	0.3374	−0.2016

Bold values indicate the better result

Table 2 Comparison of average model running time (in s)

Image size	DCP [5]	Meng [7]	Berman [14]	CAP [8]	DehazeNet [11]	AOD-Net [15]	TV [9]	Chen [10]	Ours
465*384	1.7197	4.4014	6.4785	2.2900	3.8172	0.6218	26.9408	163.0120	38.3732
400*600	2.1226	5.1096	7.9396	2.3397	4.5488	0.5215	31.3883	196.5083	53.7700
1024*768	7.4820	10.8102	16.6591	4.4108	15.5929	0.5879	99.6759	493.7204	172.2233

Table 3 Quantitative analysis comparison of synthetic images

	Image	DCP [5]	Meng et al. [7]	Berman et al. [14]	MSCNN [12]	Ours
SSIM	Baby	0.8426	0.8215	0.8811	0.9048	0.9002
	Aloe	0.9500	0.8030	0.8171	0.8993	0.9335
	Toy	0.8703	0.8710	0.8734	0.8774	0.8508
PSNR	Baby	11.6593	17.7222	16.5943	17.6007	18.7031
	Aloe	18.4094	17.2590	18.6789	21.1392	29.4708
	Toy	14.4741	20.5046	19.1675	19.5174	20.7632
MSE	Baby	0.0682	0.0169	0.0219	0.0174	0.0135
	Aloe	0.0144	0.0188	0.0136	0.0077	0.0011
	Toy	0.0357	0.0089	0.0121	0.0112	0.0084

Bold values indicate the best result

conditions, so the quality assessment indexes can be used as references for the subjective assessment.

The runtimes of DCP model, Meng et al.'s model, Berman et al.'s model, CAP model, DehazeNet model, AOD model, TV model, Chen et al.'s model and our model are listed in Table 2. Furthermore, we have compared runtimes of three image sizes of 465*384, 400*600 and 1024*768. Most of the methods are implemented in MATLAB R2015b but AOD-Net and DCPDN are implemented in Pycaffe. CAP model, DehazeNet model and AOD model are learning methods using trained model parameters; as a result, the runtimes of them are very fast. DCP method and Berman et al.'s method do not use the variational model, so they take little time in iterative calculation. TV model, Chen et al.'s model and our model are variational models using two variables while Meng et al. use one variable to optimize their model; therefore, their method is faster than ours. We want to accelerate our model with a more efficient fast algorithm in the future.

To verify the effectiveness on complete images, we test our method on synthesized hazy images from stereo images with a known depth map $d(x)$. And our method is com-

pared with DCP, Meng et al., Berman et al. and MSCNN in Fig. 7. To quantitatively assess these methods, we use a series of evaluation criteria in terms of the difference between each pair of ground-truth image and dehazed result. Apart from the widely used structural similarity (SSIM) indices, we use additional evaluation indexes, namely mean square error (MSE) and peak signal-to-noise ratio (PSNR). In Table 3, our method achieves the best performance in PSNR and MSE. On the whole, our method achieves a satisfactory performance on synthesized images.

5 Conclusion

In this paper, we have proposed a variational model with TGV regularizations to remove the haze of a single image. The proposed dehazing method obtains the optimal solutions through a primal–dual algorithm. We have validated the superior performance of our method over some state-of-the-art methods. Compared with the TV dehazing model [7], the proposed method can remove the visible blocky artifacts appearing in the transmission maps and recover haze-free images with less

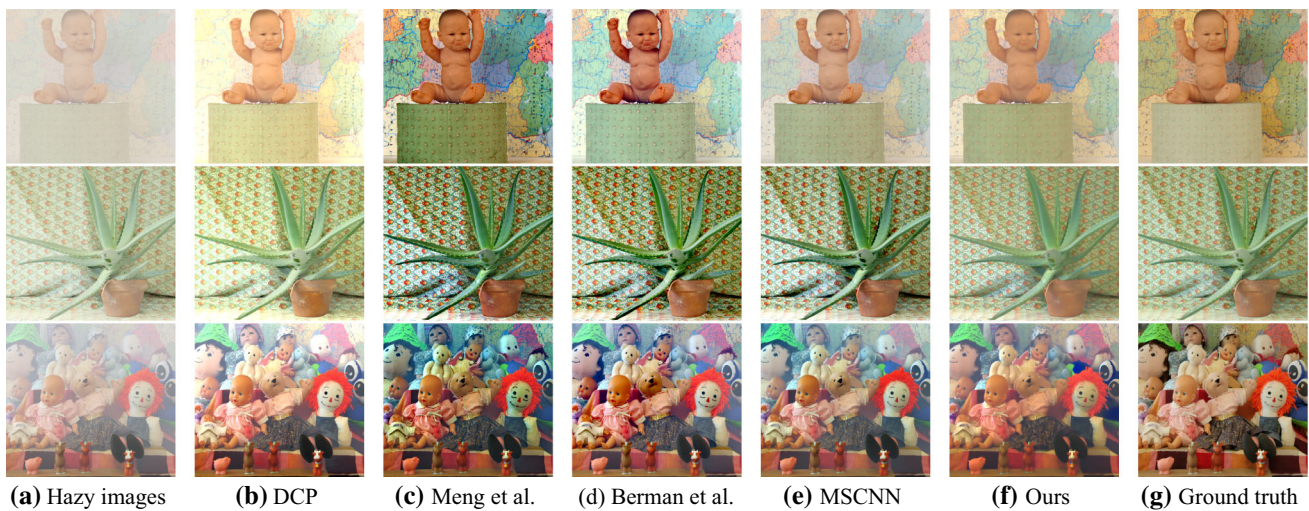


Fig. 7 Dehazing results on synthetic images

visual staircasing artifacts in the slanted plane, more details in the remote regions. But the color of our output image is a little dim; it still needs to be improved. In the future, we will consider how to enhance its brightness adaptively.

Acknowledgements The funding was provided by National Natural Science Foundation of China (Grant Nos. 11531005, 91330101).

Appendix: Image Quality Assessment

The criteria of objective image quality assessment we used are described in detail in [17].

Yu et al. [18] presented an image visibility measurement method based on the visible edge segmentation. The IVM is defined as

$$\text{IVM} = \frac{n_r}{n_{\text{total}}} \log \sum_{x \in A} C(x).$$

where n_r is the number of visible edges, n_{total} is the number of edges, $C(x)$ is the mean contrast and A denotes the image area of visible edges.

Jobson et al. [19] proposed a visual contrast measure (VCM) to quantify the degree of visibility of the image which is calculated as follows:

$$\text{VCM} = 100 * \frac{R_v}{R_t},$$

where R_v is the number of local areas, the standard deviation of which is larger than the given threshold and R_t is the total number of local areas. The VCM uses the local standard deviation which denotes the contrast of the image to measure the visibility. In general, the higher the VCM, the clearer the enhanced image.

The contrast of a clear image is usually much higher than that of a hazy image, so image contrast can be used to com-

pare different dehazing algorithms. The higher the contrast of the enhanced image, the better the dehazing algorithm. Ma and Wen [20] used the image global contrast to compare the performance of different dehazing algorithms. Tripathi et al. [21] used contrast gain to compare different dehazing algorithms. Contrast gain denotes the mean contrast difference between the enhanced image and original hazy image and is calculated by

$$C_{\text{gain}} = \bar{C}_J - \bar{C}_I,$$

where \bar{C}_J and \bar{C}_I represent the mean contrast of the enhanced image and hazy image, respectively. And C is the local contrast of the image in a small window and is calculated by

$$C(x, y) = \frac{S(x, y)}{m(x, y)},$$

where

$$S(x, y) = \frac{1}{(2r+1)^2} \sum_j \sum_i (I(x+i, y+j) - m(x, y))^2,$$

$$m(x, y) = \frac{1}{(2r+1)^2} \sum_j \sum_i I(x+i, y+j),$$

and r is the radius of the local area. The larger the contrast gain, the better the result of the dehazing algorithm.

Yu et al. thought that a good dehazing algorithm should allow the original hazy image and enhanced image to have similar histogram distributions. They used the histogram correlation coefficient (HCC) of the two color images as a criterion to assess the performance of color restoration.

Wu and Zhu used the image structural similarity (SSIM) and universal quality index (UQI) to assess the performance of the structural similarity between the original hazy image and the enhanced image.

References

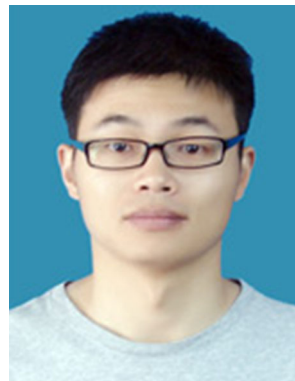
- Nayar, S.K., Narasimhan, S.G.: Vision in bad weather. *Proc. IEEE Conf. Comput. Vis.* **2**, 820–827 (1999)
- Narasimhan, S.G., Nayar, S.K.: Contrast restoration of weather degraded images. *IEEE Trans. Pattern Anal. Mach. Learn.* **25**(6), 713–724 (2003)
- Tan, R.T.: Visibility in bad weather from a single image. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–8 (2008)
- Fattal, R.: Single image dehazing. *ACM Trans. Graph.* **27**(3), 72:1–72:10 (2008)
- He, K., Sun, J., Tang, X.: Single image haze removal using dark channel prior. *IEEE Trans. Pattern Anal. Mach. Intell.* **33**(12), 2341–2353 (2011)
- Tarel, J.-P., Hautière, N.: Fast visibility restoration from a single color or gray level image. In: *Proceedings of IEEE 12th International Conference on Computer Vision*, pp. 2201–2208 (2009)
- Meng, G., Wang, Y., Duan, J., Xiang, S., Pan, C.: Efficient image dehazing with boundary constraint and contextual regularization. In: *Proceedings of IEEE International Conference on Computer Vision*, pp. 617–624 (2013)
- Zhu, Q., Mai, J., Shao, L.: A fast single image haze removal algorithm using color attenuation prior. *IEEE Trans. Image Process.* **24**(11), 3522–3533 (2015)
- Fang, F., Li, F., Zeng, T.: Single image dehazing and denoising: a fast variational approach. *SIAM J. Imaging Sci.* **7**(2), 969–996 (2014)
- Chen, C., Do, Minh N., Wang, J.: Robust image and video dehazing with visual artifact suppression via gradient residual minimization. *ECCV* **9906**, 576–591 (2016)
- Cai, B., Xu, X., Jia, K., Qing, C., Tao, D.: DehazeNet: an end-to-end system for single image haze removal. *IEEE Trans. Image Process.* **25**(11), 5187–5198 (2016)
- Ren, W., Liu, S., Zhang, H., Pan, J., Cao, X., Yang, M.: Single image dehazing via multi-scale convolutional neural networks. *Comput. Vision ECCV* **9906**, 154–169 (2016)
- Bredies, K.: Recovering piecewise smooth multichannel images by minimization of convex functionals with total generalized variation penalty. *Lecture Notes Comput. Sci.* **8293**, 44–77 (2014)
- Berman, D., Avidan, S., Treibitz, T.: Non-local image dehazing. In: *CVPR*, pp. 1674–1682 (2016)
- Li, B., Peng, X., Wang, Z., Xu, J., Feng, D.: An all-in-one network for dehazing and beyond. In: *ICCV* (2017)
- Zhang, H., Patel, V.M.: Densely connected pyramid dehazing network. In: *CVPR* (2018)
- Xu, Y., Wen, J., Fei, L., Zhang, Z.: Review of video and image defogging algorithms and related studies on image restoration and enhancement. *IEEE Access* **4**, 165–188 (2017)
- Yu, X., Xiao, C., Deng, M., Peng, L.: A classification algorithm to distinguish image as haze or non-haze. In: *Proceedings of IEEE International Conference on Image Graphics*, pp. 286–289 (2011)
- Jobson, D.J., Rahman, Z.-U., Woodell, G.A., Hines, G.D.: A comparison of visual statistics for the image enhancement of FORESITE aerial images with those of major image classes. In: *Proceedings of SPIE*, pp. 624601-1–624601-8 (2006)
- Ma, Z., Jie, W.: Single-scale retinex sea fog removal algorithm fused the edge information. *J. Comput. Aided Des. Comput. Graph.* **27**(2), 217–225 (2015)
- Tripathi, A.K., Mukhopadhyay, S.: Removal of fog from images: a review. *IETE Tech. Rev.* **29**(2), 148–156 (2012)



Yanan Gu received the Masters degree in mathematics from Guangxi University for Nationalities, China, in 2013. Since 2013, she has been a Ph.D. student in Nanjing University of Science and Technology, China. Her research interests are image denoising, image dehazing and image inpainting using the partial differential equations.



Xiaoping Yang received the Ph.D. degree in mathematics from Hunan University, China, in 1992. He is currently a professor in the Department of Mathematics, Nanjing University. His research interests include partial differential equations, geometric analysis and image processing.



Yiming Gao received the Ph.D. degree from the Nanjing University of Science and Technology, Nanjing, China, in 2018. He holds a postdoctoral position with the School of Mathematics, Nankai University, China. His research interests are variational methods in imaging, high-order regularization and texture modeling.