

석사학위논문
Master's Thesis

동일 깊이 근접장과 비국소 정규화를 이용한
헤이즈 제거 알고리즘

Dehazing using Non-Local Regularization
with Iso-Depth Neighbor-Fields

2017

김인철 (金仁轍 Kim, Incheol)

한국과학기술원

Korea Advanced Institute of Science and Technology

석사학위논문

동일 깊이 근접장과 비국소 정규화를 이용한
헤이즈 제거 알고리즘

2017

김인철

한국과학기술원

전산학부

동일 깊이 근접장과 비국소 정규화를 이용한
헤이즈 제거 알고리즘

김 인 철

위 논문은 한국과학기술원 석사학위논문으로
학위논문 심사위원회의 심사를 통과하였음

2016년 12월 20일

심사위원장 김 민 혁

심 사 위 원 박 진 아

심 사 위 원 윤 성 의

Dehazing using Non-Local Regularization with Iso-Depth Neighbor-Fields

Incheol Kim

Advisor: Kim, Min Hyuk

A dissertation submitted to the faculty of
Korea Advanced Institute of Science and Technology in
partial fulfillment of the requirements for the degree of
Master of Science in Computer Science

Daejeon, Korea
December 20, 2016

Approved by

Kim, Min Hyuk
Professor of Computer Science

The study was conducted in accordance with Code of Research Ethics¹.

¹ Declaration of Ethical Conduct in Research: I, as a graduate student of Korea Advanced Institute of Science and Technology, hereby declare that I have not committed any act that may damage the credibility of my research. This includes, but is not limited to, falsification, thesis written by someone else, distortion of research findings, and plagiarism. I confirm that my thesis contains honest conclusions based on my own careful research under the guidance of my advisor.

MCS
20153153

김인철. 동일 깊이 근접장과 비국소 정규화를 이용한 헤이즈 제거 알고리즘. 전산학부 . 2017년. 37+iv 쪽. 지도교수: 김민혁. (영문 논문)
Incheol Kim. Dehazing using Non-Local Regularization with Iso-Depth Neighbor-Fields. School of Computing . 2017. 37+iv pages. Advisor: Kim, Min Hyuk. (Text in English)

초 록

단일 이미지에서 헤이즈를 제거하는 것은 심각한 불량조건문제(ill-posed problem)인데, 이는 정보가 매우 부족하기 때문이다. 일반적인 헤이즈 제거 알고리즘들은 대기 산란광(airlight)의 양을 자연적 이미지 통계(natural image statistics)를 이용하여 희소한 초기값을 추정하고, 이것을 전파하여 조밀한 투과율(transmission)맵을 도출하고 이것으로 헤이즈가 제거된 이미지를 복원한다. 본 학위논문에서는 불완전한 자연적 이미지 통계를 사용하는 대신 픽셀 값을 대기 벡터(atmospheric vector)로의 투영을 통해 희소 투과율 맵을 추정하는 강인한 알고리즘을 제안한다. 또한, 희소한 헤이즈를 전파하는 문제는 다른 정규화(regularization) 문제와는 다르게 이미지의 깊이 정보와 높은 상관 관계를 갖는다. 한편, 일반적인 격자형 마르코프 임의장(Markov random fields)을 이용하면 헤이즈 전파 시 이미지 내에서 깊이 값이 급격히 변하는 지점에서 아티팩트가 발생한다. 본 학위논문에서는 동일 깊이 근접장(iso-depth nearest-neighbor field)을 마르코프 임의장에 적용하여 더 정교한 헤이즈 전파를 실현하고, 이를 통해 헤이즈 제거 시 깊이 변화가 급격한 곳에서 발생하는 아티팩트를 감소시킬 수 있다. 또한, 이 방법은 다른 헤이즈 제거 알고리즘의 정규화 방법에 적용되어 아티팩트 감소를 기대할 수 있다.

핵심 낱말 헤이즈 제거, 비국소 정규화, 이미지 복원

Abstract

Removing haze from a single image is a severely ill-posed problem due to the lack of scene information. General dehazing algorithms estimate airlight initially using natural image statistics and then propagate the incompletely estimated airlight to build a dense transmission map, yielding a haze-free image. Propagating haze is different from other regularization problems, as haze is strongly correlated with depth according to the physics of light transport in participating media. However, since there is no depth information available in single-image dehazing, traditional regularization methods with a common grid random field often suffer from *haze isolation artifacts* caused by abrupt changes in scene depths. In this paper, to overcome the haze isolation problem, we propose a non-local regularization method by combining Markov random fields (MRFs) with nearest-neighbor fields (NNFs), based on our insightful observation that the NNFs searched in a hazy image associate patches at the similar depth, as local haze in the atmosphere is proportional to its depth. We validate that the proposed method can regularize haze effectively to restore a variety of natural landscape images. This proposed regularization method can be used separately with any other dehazing algorithms to enhance haze regularization.

Keywords dehazing, non-local regularization, image restoration

Contents

Contents	i
List of Tables	iii
List of Figures	iv
Chapter 1. Introduction	1
1.1 Motivation	1
1.2 Contributions	1
1.3 Thesis Outline	2
Chapter 2. Background	3
2.1 Haze Formation Model	3
2.2 Haze Regularization	4
2.2.1 Motivation	4
2.2.2 Hidden Markov Model	5
2.2.3 Markov Random Fields	7
2.2.4 MRF Estimation	7
Chapter 3. Related Work	9
3.1 Multiple Image-based Dehazing	9
3.2 Learning-based Dehazing	9
3.3 Single Image-based Dehazing	9
3.4 Haze Regularization	10
Chapter 4. Algorithm	11
4.1 Haze Estimation	11
4.1.1 Atmospheric Vector Estimation	11
4.1.2 Transmission Estimation	12
4.1.3 Removing Outliers	13
4.2 Non-Local Regularization using Iso-Depth Neighbor Fields	14
4.2.1 GMRF Model	14
4.2.2 Iso-Depth Neighbor Fields	16
Chapter 5. Results	18
5.1 External Evaluation	18
5.1.1 Qualitative Comparison	18

5.1.2	Quantitative Comparison	18
5.1.3	Regularization	20
5.2	Internal Evaluation	21
5.2.1	Running Time	21
5.2.2	Impact of Patch Size	22
5.2.3	Outlier Removal	23
Chapter 6.	Discussion and Future Work	31
6.1	Limitations	31
Chapter 7.	Conclusion	33
	Bibliography	34
	Acknowledgments in Korean	36
	Curriculum Vitae in Korean	37

List of Tables

5.1	Quantitative comparison	18
5.2	Comparison of time performance between GMRFs and GMRFs with NNFs	22

List of Figures

2.1	The process of haze removal	4
2.2	Effects of blocky artifacts in recovery of a clear image	5
2.3	Constant signals with Gaussian noise and its hidden Markov model	6
2.4	A hidden Markov model	7
2.5	A Markov random field	8
4.1	Haze estimation in a linear subspace	12
4.2	Attenuation function for compensation of non-orthogonality	14
4.3	Non-local MRF model with nearest-neighbor fields	16
4.4	Iso-depth neighbor-fields	17
5.1	Quantitative comparison	19
5.2	Comparison of regularization methods	20
5.3	Comparisons of dehazing in terms of regularization	21
5.4	Effects of nonlocal neighbors in haze regularization	22
5.5	Comparison of the impact of a patch size	23
5.6	Validation of our narrow angle outlier rejection method	24
5.7	Validation of our saturated intensity outlier rejection process	25
5.8	Comparison with other methods I	26
5.9	Comparison with other methods II	27
5.10	Comparison with other methods III	28
5.11	Comparison with other methods IV	29
5.12	Comparison with other methods V	30
6.1	Failure cases	32

Chapter 1. Introduction

1.1 Motivation

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 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 [12]. 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, which yields 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, most of propagation algorithms with a common grid random field often suffer from *haze-isolation artifacts*. The amount of haze in the atmosphere at each pixel is determined by its depth. If there is an abrupt change in scene depth, the grid random field cannot regularize a transmission map with sharp-edge discontinuity due to wrong propagation. In order to handle abrupt changes of haze density, we need a scene depth information, even though it is unavailable in single-image dehazing.

In this paper, we propose a non-local regularization for dehazing that can propagate sparse airlight estimates to yield a dense transmission map without suffering from the typical isolation problem. Our regularization approach is developed by combining Markov random fields (MRFs) with nearest-neighbor fields (NNFs) searched by PatchMatch [2]. We found that the NNFs searched in a hazy image associate patches at the similar depth. Since no depth information is available in single-image dehazing, we utilize the NNFs information to infer depth cues for propagating hidden states of scattered light, which is exponentially proportional to depth [19]. 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.

1.2 Contributions

Our **contributions** are:

- **a simple but consistent estimation of transmission** by projecting hazy signals onto an estimated atmospheric vector, accounting for piecewise consistency of airlight, and

- **a novel NNF-based non-local MRF regularization method** that refines transmission maps by adding more neighbors obtained from nearest-neighbor fields.

Most of these contributions have been presented in the following publication:

- **Incheol Kim** and Min H. Kim, *Dehazing using Non-Local Regularization with Iso-Depth Neighbor-Fields*, 12th International Conference on Computer Vision Theory and Applications (VISAPP 2017), accepted (oral presentation).

1.3 Thesis Outline

This thesis is organized as follows. Chapter 2 describes background knowledge about a haze formation model and Markov models for haze regularization. Chapter 3 introduces related works on dehazing method and haze regularization methods. Chapter 4 explains the details of our dehazing and haze regularization algorithms. Chapter 5 shows the results of our method. We share remaining issues in our method and future works in Chapter 6. Finally, Chapter 7 concludes this thesis.

Chapter 2. Background

This chapter provides a brief overview of background knowledge.

2.1 Haze Formation Model

Haze is an aerosol that consists of ashes, dust, and smoke. Haze tends to present a gray or bluish hue [19], 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. This phenomenon is mathematically defined as a *transmission* that represents the portion of light from the scene radiance that is not scattered by 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 [19, 20, 9, 10]:

$$I(x) = t(x)J(x) + (1 - t(x))A, \quad (2.1)$$

where $I(x)$ is a linearized image intensity at a pixel x , $J(x)$ is an unknown scene radiance, $t(x)$ is a transmission value, describing the portion of remaining light when the reflected light from a scene surface goes to the observer through the medium, and A is a global atmospheric vector that is unknown as well. The atmospheric vector A represents the color vector orientation and intensity of airlight in the linearized sRGB color space, and along with the interpolation term $(1 - t(x))$, the right additive term in Equation (2.1) defines the intensity of airlight at the pixel x , where the airlight is a phenomenon that acts like a light source, which is caused by scattering of participating media in the atmosphere [19]. Additionally, the atmospheric vector is independent of scene locations, i.e., the atmospheric light is globally constant, while the airlight is spatially-varying due to its dependency of scene depths.

The number 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 the scattering coefficient of the atmosphere [20] 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 transmission t and an atmospheric vector A so that scene radiance J can be recovered from the transmission t and the atmospheric vector A by the following:

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

where ϵ is a small value to prevent division by zero. The process of haze removal is shown in Figure 2.1.

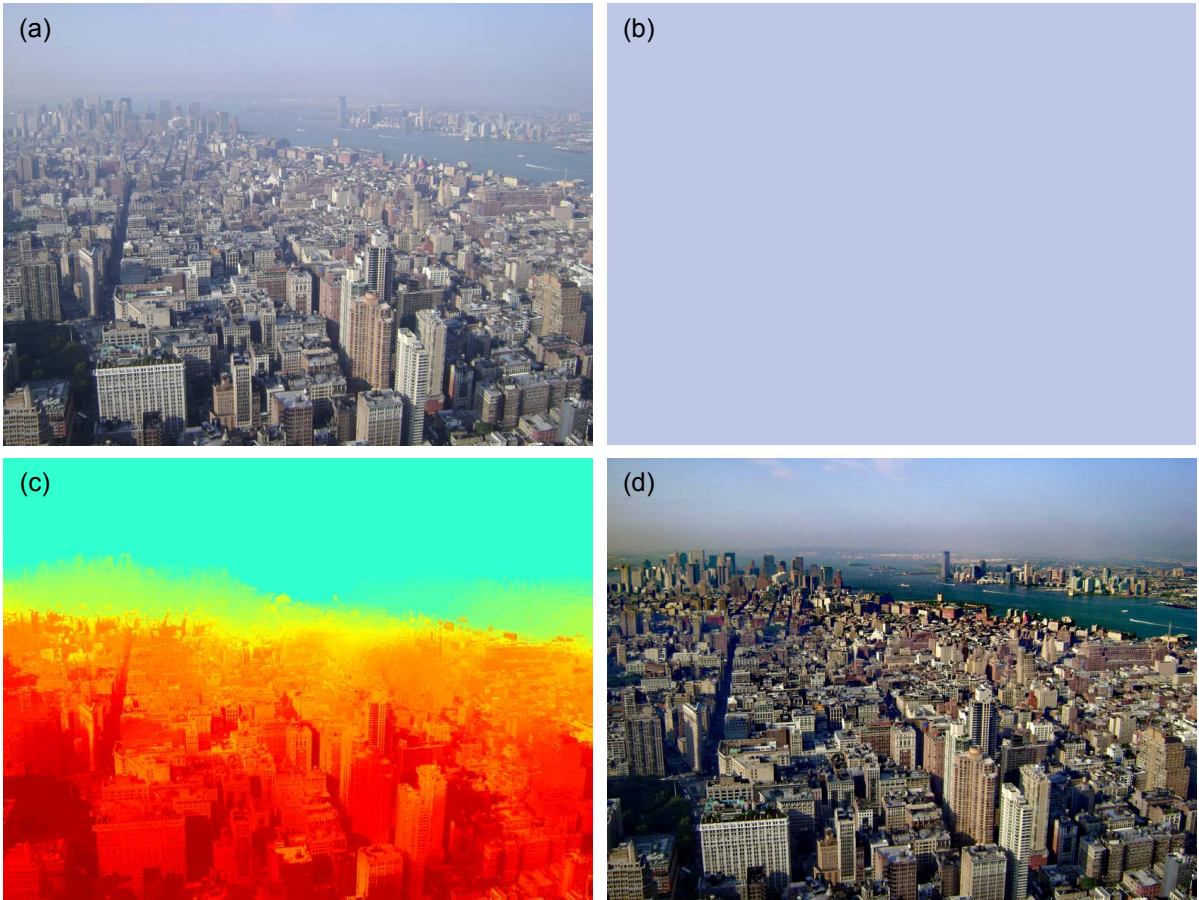


Figure 2.1: The process of haze removal. (a) The picture shows an input outdoor hazy image. (b) The atmospheric light map estimated from (a), i.e., the intensity map of haze that are full of airlight without any scene radiance. (c) The transmission map estimated from (a), which is closely related with scene depths of (a). (d) The final dehazed result that is generated by using Equation (2.1) with (a)–(c).

2.2 Haze Regularization

2.2.1 Motivation

In transmission estimation, the most widely used scheme is to compute transmission values in a patch-wise manner instead of pixel-wise computation in order to enhance the fidelity of estimation. However, the transmission values are not always constant within a patch, i.e., when there is an abrupt change in depths within a patch; therefore, the computed map may have many blocky artifacts that degrade the quality of dehazed results. Further, some methods include an outlier rejection stage to enhance the quality of haze removal, so some values may be missing after estimating transmission values. For these reasons, the necessity of regularization and propagation arises to yield a high-quality transmission map with sharp edge-discontinuities. See Figure 2.2.

We briefly describe a background knowledge of the hidden Markov model, and Markov random fields that are commonly used in haze regularization.

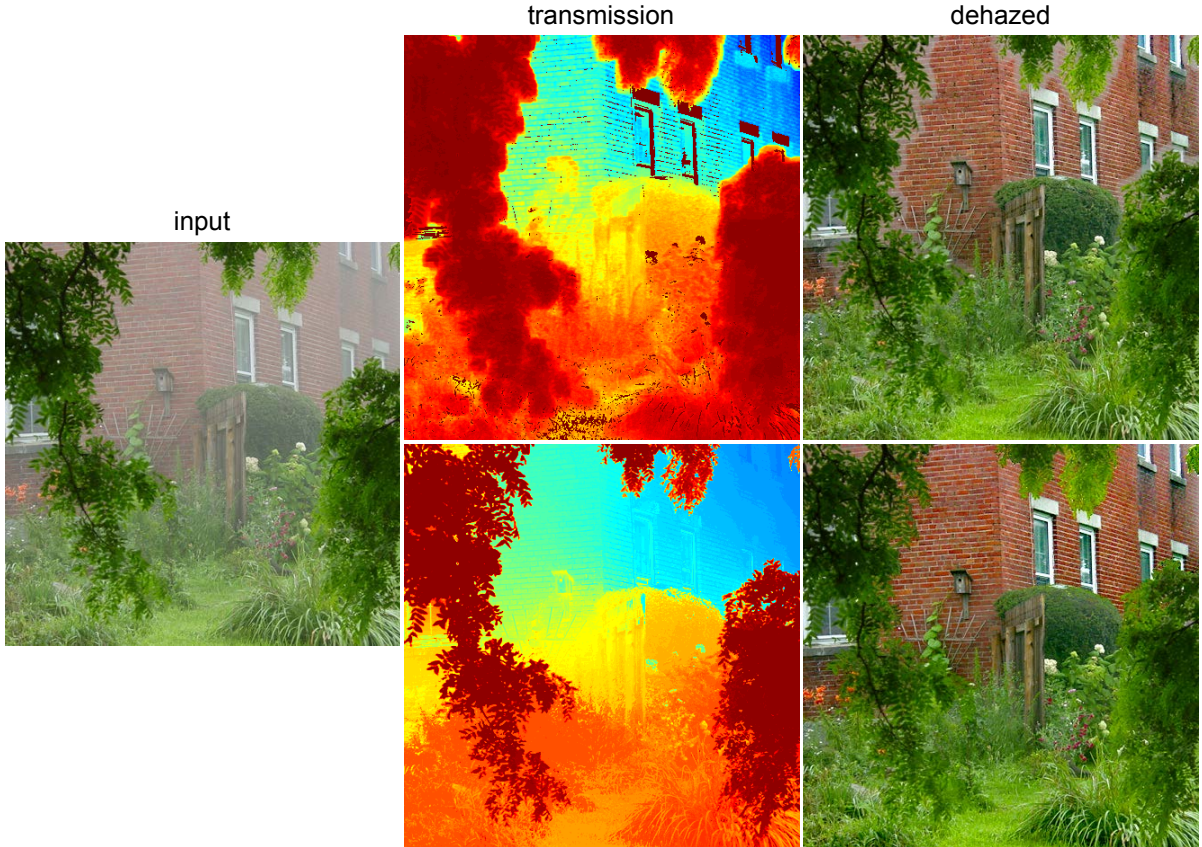


Figure 2.2: The leftmost picture is an input hazy image. The upper image in the second column shows the estimated transmission map with blocky artifacts and outliers from the input image. The lower image in the second column is the regularized transmission map from the upper one. The third column represents the dehazed results from the corresponding transmission maps in the second column.

2.2.2 Hidden Markov Model

A hidden Markov model (HMM) is a commonly-used technique to infer original signals from distorted input signals. Figure 2.3 shows an example of applying an HMM model to real data. Suppose there are N latent states, where each latent state can take K possible state instances. See Figure 2.4. The goal of inference on an HMM is to find the best sequence of the latent states given their observation values. In case of brute force estimation, we need to check $O(K^N)$ cases by putting each possible state instance and compute the probability, which is an NP-hard problem. Under Bayesian assumption, we can ignore all states that are not connected, meaning that unconnected states are independent to each other from the perspective of probability. Thus, we need to check connected states only, which is much easier than the brute force approach.

The maximum a posteriori (MAP) estimate of the HMM in Figure 2.4 is given as

$$\hat{w}_{\text{MAP}} = \arg \max_{w_{1...K}} \Pr(x_{1...N} | y_{1...N}). \quad (2.2)$$

Since

$$\Pr(x_{1...N}, y_{1...N}) = \Pr(x_{1...N} | y_{1...N}) \Pr(y_{1...N}), \quad (2.3)$$

we can decompose a joint probability into the multiplication of one probability and its conditional probability. Supposing that the observation probability $\Pr(y_{1...N})$ in Equation (2.3) is constant, we can

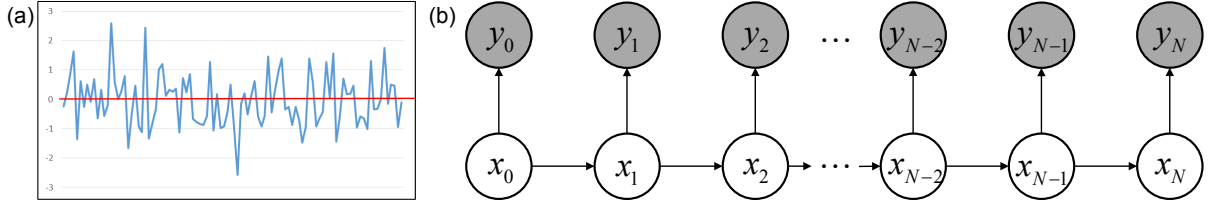


Figure 2.3: (a) The red line shows latent constant signals with value 0, which is to be estimated, and blue plot represents the input signals distorted by normally distributed noise. (b) The hidden Markov model (HMM) describes the signals shown in (a). The gray circles denote observation values that correspond to the blue spiky curve in (a). The white circles describe latent states, which is the red line in the case of (a). In HMM estimation, latent states are to be estimated from observations.

equating the joint probability and the conditional probability in MAP estimation, which is written as

$$\hat{w}_{\text{MAP}} = \arg \max_{w_{1...K}} \Pr(x_{1...N} | y_{1...N}) \equiv \arg \max_{w_{1...K}} \Pr(x_{1...N}, y_{1...N}).$$

If we take negative log function on the joint probability, this is written as

$$\arg \min_{w_{1...K}} (-\log(\Pr(x_{1...N}, y_{1...N}))).$$

Based on the HMM model in Figure 2.4, data observation probabilities are independent, so the joint probability term $\Pr(x_{1...N}, y_{1...N})$ can be factorized as

$$\Pr(x_{1...N}, y_{1...N}) = \prod_{i=1}^N \Pr(x_i, y_i).$$

Each factorized joint probability consists of likelihood and prior terms based on the HMM model in Figure 2.4, which is written as

$$\Pr(x_i, y_i) = \Pr(y_i | x_i) \Pr(x_i | x_{i-1}).$$

Based on Figure 2.4, $\Pr(y_i | x_i)$ is a likelihood term that is a connection between one latent state and its observation value, and $\Pr(x_i | x_{i-1})$ is a prior term that describes a probability of transition to one latent state from its previous latent state. Finally, the MAP estimate in a factorized form is written as

$$\begin{aligned} \hat{w}_{\text{MAP}} &= \arg \min_{w_{1...K}} (-\log(\Pr(x_{1...N}, y_{1...N}))) \\ &= \arg \min_{w_{1...K}} \left(-\sum_{n=1}^N \log[\Pr(y_i | x_i)] - \sum_{n=2}^N \log[\Pr(x_i | x_{i-1})] \right). \end{aligned}$$

Suppose we estimate the k^{th} latent state \hat{w}_k^{MAP} . Then, the MAP estimate of the k^{th} variable is

$$\begin{aligned} \hat{w}_k^{\text{MAP}} &= \Pr(x_k = \hat{w}_k^{\text{MAP}}, y_k) \\ &= \arg \min_{w_{1...K}} \left(-\sum_{i=1}^{k-1} \log[\Pr(y_i | x_i)] - \sum_{i=2}^{k-1} \log[\Pr(x_i | x_{i-1})] \right) + \Pr(y_k | x_k). \end{aligned}$$

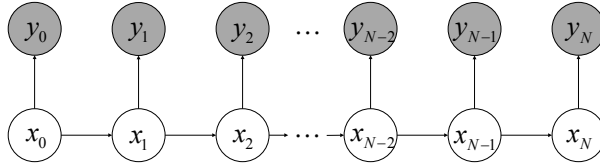


Figure 2.4: A hidden Markov model with N latent states and observations. The gray circles denote input observations values, and the white circles model latent states to be estimated. Each intermediate latent state is connected by its previous latent state, and connects its observation and next latent state. Each observation node are independent to each other since they are not connected to each other. This means that observation values do not affect the others and are not affected by the others as well.

This can be solved in a dynamic programming fashion since the previous probabilities,

$$-\sum_{i=1}^{k-1} \log [\Pr (y_i | x_i)] - \sum_{i=2}^{k-1} \log [\Pr (x_i | x_{i-1})],$$

do not affect the current latent state x_i , so we do not need to compute the previous probabilities again. This is a basic idea of Viterbi's decoding algorithm [29] that infers latent states in a first-order sequential HMM.

2.2.3 Markov Random Fields

An HMM can be used to describe sequential noisy data to infer its latent variables; however, higher-dimensional data such as images cannot be modeled with any HMM. In this case, we can establish an undirected graphical model of random variables that can represent data of a grid shape. This model is called a Markov random field (MRF). Like the case of Equation (2.2), the joint probability distribution of the MRF in Figure 2.5 is written as

$$\Pr (\mathbf{x} | \mathbf{y}) = \frac{1}{Z} \prod_{i,j} \psi_{ij} (x_i, x_j) \phi (x_i, y_i), \quad (2.4)$$

where $\Pr (\mathbf{x} | \mathbf{y})$ is the joint probability distribution of the MRF, $\psi_{ij} (x_i, x_j)$ is a prior term between two latent variables x_i and x_j , $\phi (x_i, y_i)$ is a likelihood between a latent variable x_i and its observed value y_i , and Z is a normalization factor. Inference on a joint probability of an MRF given by Equation (2.4) is accomplished by maximizing the posterior distribution function $\Pr (\mathbf{x} | \mathbf{y})$. However, unlike an HMM, the grid MRF has cycles, so obtaining the exact MAP estimate is infeasible since the number of paths from one state to another is infinite. Instead, we can obtain an approximate MAP estimate by inferring locally maximal joint probabilities so that we only need to compute suboptimal MAP estimates, whereas the globally optimal solution was computed in inferring a first-order sequential HMM.

2.2.4 MRF Estimation

There are two categories of MRF estimation methods in terms of a state space: discrete and continuous.

Discrete Space Estimation Discrete MRF solvers are basically a combinatorial optimization. They explore the discrete state instance space and compute the joint distribution for each combination of states. Then, a combination of states that locally maximizes the joint distribution is chosen as an MAP

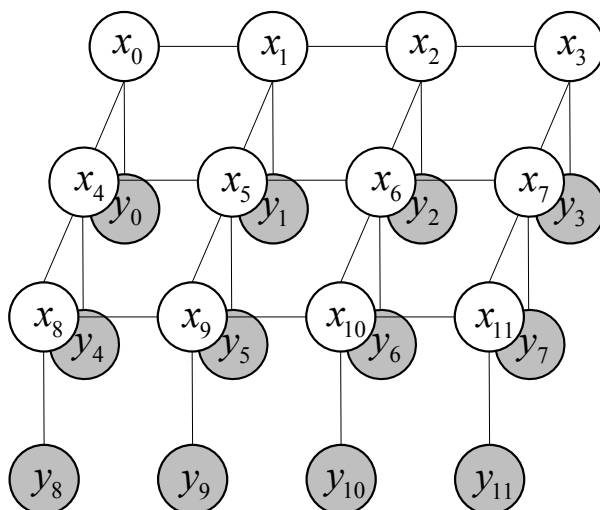


Figure 2.5: A grid MRF model. The gray nodes represent observed values, and white nodes are latent states to be estimated. Unlike the HMM shown in Figure 2.4, the MRF consists of undirected edges, so the joint probability is affected by all connected variables, whereas in an HMM only nodes that are pointed by others' pointing edges are affected.

estimate. Belief propagation [22], graph cut-based [6], iterated conditional mode [4], and tree-reweighted message passing algorithm [30] are representative discrete MRF solvers.

Continuous Space Estimation Algorithms in this category compute the optimal estimate by relaxing an original discrete state space to a continuous state space. For instance, convex relaxation of MRFs [8, 25, ?] is used to impose smoothness or continuity constraints for low-level vision problems such as denoising [23].

Chapter 3. Related Work

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

3.1 Multiple Image-based Dehazing

Since removing haze in the atmosphere is an ill-posed problem, several works have attempted to solve the problem using 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 [24]. Narasimhan et al. employ multiple images with different weather conditions to restore the degraded image using an irradiance model [19, 20]. Kopf et al. remove haze from an image with additionally known scene geometry, instead of capturing multiple images [13]. These haze formation models stand on the physics of light transport to provide sound accuracy. However, these applications could be limited at the cost to acquiring multiple input images.

3.2 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 learn a prior on natural image statistics to factorize the haze layer and the scene radiance from the hazy image. Tang et al. define haze-relevant features that are related to the properties of hazy images, and train them using the random forest regression [27]. Zhu et al. obtain the color attenuation prior using supervised learning [32]. 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. Recently, a CNN-based single image dehazing method is proposed [?]. However, no general statistical model can predict the diverse distributions of natural light environments; hence, they often fail to restore haze-free images that are not similar to the trained dataset.

3.3 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* [26, 28, 12, 21, 1, 10]. Tan and Tarel restore visibility by maximizing local contrast, assuming that clean color images have a high contrast, but this causes overly saturated results [26, 28]. He et al. exploit image statistics where a natural image in the sRGB color space should include a very low intensity within a local region [12]. However, it often overestimates the amount of haze if there is a large area having bright pixels. Nishino et al. 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 [21].

Developing the natural image prior further, Fattal 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 [10]. 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. suggest a nighttime dehazing method that removes a glow layer made by the combination of participating media and light source such as lamps [15]. Recently, a non-local transmission estimation method was proposed by Berman et al., 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 [3].

In addition, an assumption on *light transport* in natural scenes is also used. Fattal assumes that shading and transmission are statistically independent [9], and Meng et al. impose boundary conditions on light transmission [18]. In particular, our airlight estimation follows the traditional approach based on dimension-minimization approach [9], which allows for robust performance in estimating airlight.

3.4 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 (MRFs) are most commonly used in many dehazing algorithms [26, 9, 7, 21, 3], and filtering methods are also used, for instance, matting Laplacian [12], guided filtering [11], and a total variation-based approach [28, 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 attempts to mitigate this isolation problem by utilizing augmented Markov random fields, which extend connection boundaries of MRFs [10]. 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. non-locally extend the boundary in estimating haze [3]; 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 regularization of our method, we extend neighbors in MRFs with iso-depth NNFs for using additional non-local information to infer depth cues based on the physics of light transport.

Chapter 4. Algorithm

In this section, we describe our dehazing method with non-local haze regularization method. We first introduce how to estimate an atmospheric vector from a given image, and recover the clear image with transmission values. We also explain our haze estimation method with the haze formation model, which works in linear subspaces. In addition, we describe our outlier rejection method to exclude outlying pixels that severely degrade the quality of dehazing. From the previous stage, we will have an incomplete transmission map with many blocky artifacts due to the patch-wise haze estimation. We propose our non-local haze regularization method to propagate the transmission values to build a dense and sharp transmission map.

4.1 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 transmission values with certain natural image statistics within a patch based on the local smoothness assumption on scene depths.

Following this perspective of the traditional approach [9], 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 a 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 4.1 depicts the estimation process for an overview.

4.1.1 Atmospheric Vector Estimation

Airlight is a phenomenon that acts like a light source, which is caused by scattering of participating media in the atmosphere [19]. 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 [12]. 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.'s method does, so that we can estimate the atmospheric vector consistently.

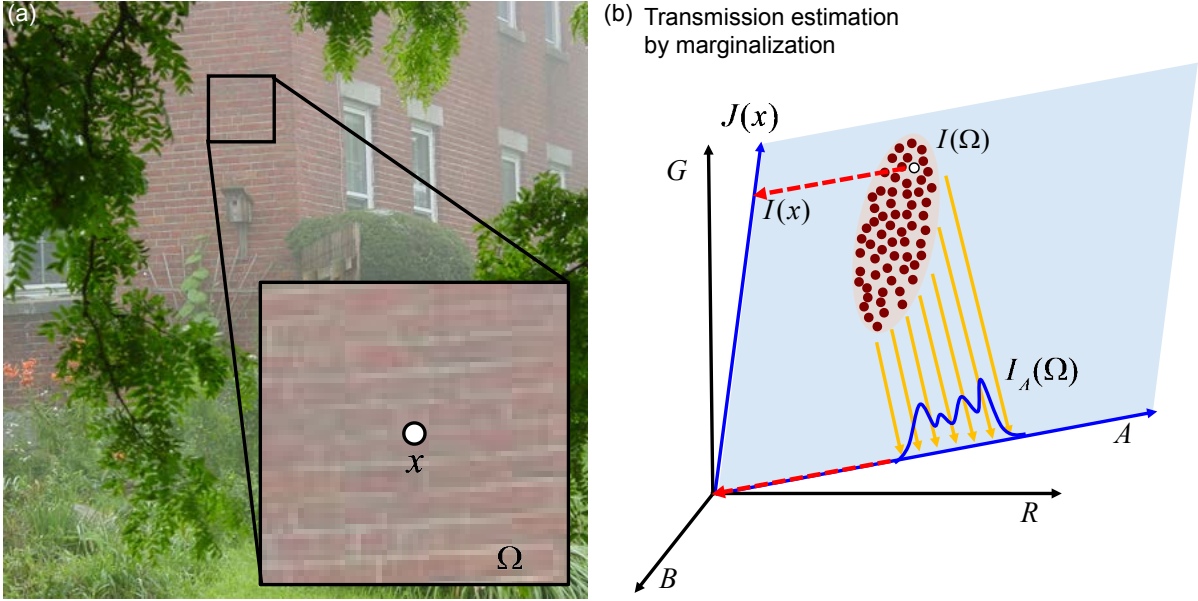


Figure 4.1: (a) Extracting a patch from a hazy image. $I(\Omega)$ is a set of linearized color pixels in the patch Ω that has a center pixel of x . The white dot indicates the center pixel x . (b) We initially estimate the amount of haze using linear subspaces [19, 9]. A is an atmospheric vector of the 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 $I(x)$. The pixel $I(x)$ 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 $I(\Omega)$. The 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.

We subsequently average the chosen pixels to reject noise.

4.1.2 Transmission Estimation

We first assume that transmission is piecewise smooth. In Equation (2.1), 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 a haze signal 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 *scalar projections* of color vectors $I(\Omega)$ onto an atmospheric vector A in a patch Ω (Figure 4.1), where the pixel x is located at the center, then it can be written as following [9]:

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

We assume the airlight within a patch to be constant while the scene radiance might vary. In order 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 Figure 4.1(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:

$$I_A^{\min}(\Omega) = P_i(I_A(k)), \quad I_A^{\min}(\Omega) \in \mathbb{R}^1,$$

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 Equation (2.1), which is

$$(1 - t(x)) \leftarrow I_A^{\min}(\Omega).$$

As a result, the transmission estimate is written as follows:

$$t(x) = 1 - I_A^{\min}(\Omega).$$

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 larger value due to the correlation between the two vectors. We attenuate the I_A^{\min} by a function with respect to the angle between a pixel vector and an atmospheric vector that is given by

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

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

$$f(\bar{\theta}) = \frac{e^{-k\bar{\theta}} - e^{-k}}{1 - e^{-k}}, \quad (4.1)$$

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 Figure 4.2(c).

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.

4.1.3 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 Section 4.2).

Narrow Angle Outliers 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

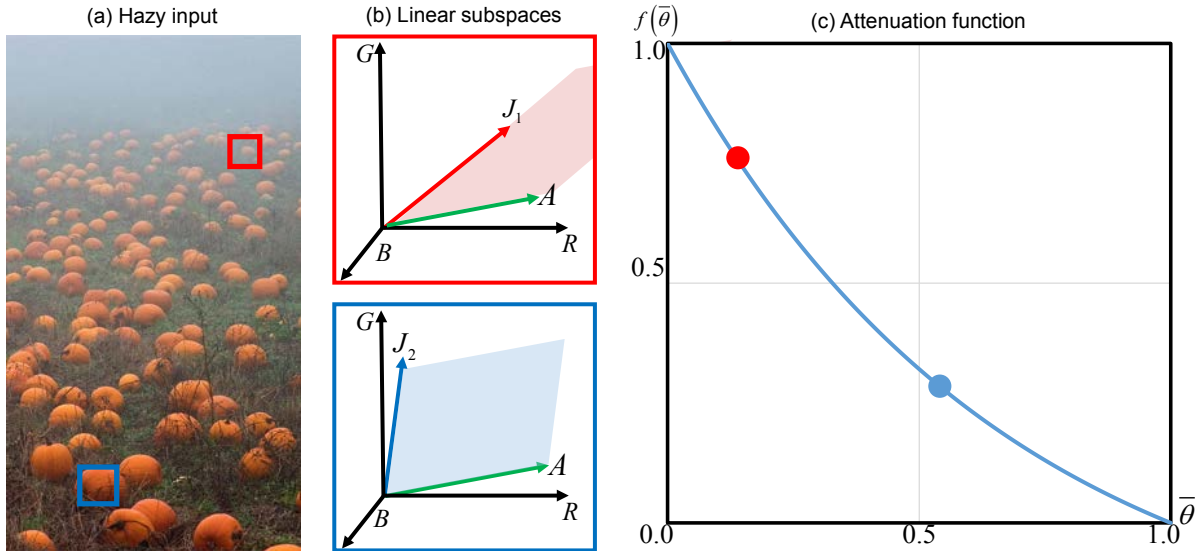


Figure 4.2: (a) A hazy input image. (b) Each single pixel from the red and blue boxes is plotted in the RGB space along with the atmospheric vector A . J_1 and J_2 in each plot correspond to the two pixels extracted. (c) The attenuation function defined as Equation (4.1) is plotted as above. The red 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.

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 radian, 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.

Saturated Intensity Outliers 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.

4.2 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 regularize valid transmission values in the image.

4.2.1 GMRF 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 regularize the incompletely estimated transmission map using Markov random fields (MRFs). The probability distribution of one node in an

MRF is given by

$$p(t(x) | \hat{t}(x)) = \phi(t(x), \hat{t}(x)) \psi(t(x), t(y)), \quad (4.2)$$

where $t(x)$ is a latent transmission variable at pixel x , $\hat{t}(x)$ is an initially estimated transmission value, $\phi()$ is a data term of the likelihood between $t(x)$ and $\hat{t}(x)$, and ψ is a smoothness prior of latent transmission $t(x)$ against neighboring transmission $t(y)$ within a patch Ω , $y \in \Omega$. While the data term $\phi()$ describes the fidelity of observations by imposing a penalty function between the latent variable and the observed value, the smoothness term $\psi()$ enforces smoothness by penalizing the errors between one latent variable and its neighboring variables.

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

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

where $\sigma_{\hat{t}}(\Omega)$ is the variance of observation values \hat{t} within patch Ω that has the center at pixel x . 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 smoothness term $\Psi()$ is written as

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

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 smoothness 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 $(t(x) - t(y))^2$ is large, then it indicates that the distance between $t(x)$ and its neighbor $t(y)$ is large, and hence the cost from the regularization term will also become large, which enforces strong smoothness between them. $\|I(x) - I(y)\|^2$ 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 a 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 [17, 9].

By combining the two terms we can obtain a posterior distribution:

$$p(t(x) | \hat{t}(x)) = \exp\left(-\frac{(t(x) - \hat{t}(x))^2}{\sigma_{\hat{t}}(\Omega)^2}\right) \prod_{y \in N_x} \exp\left(-\frac{(t(x) - t(y))^2}{\|I(x) - I(y)\|^2}\right). \quad (4.3)$$

Finally, we formulate a cost function by taking the negative log of the posterior distribution [Equa-

tion (4.3)] following [9, 10], which is written by

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

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

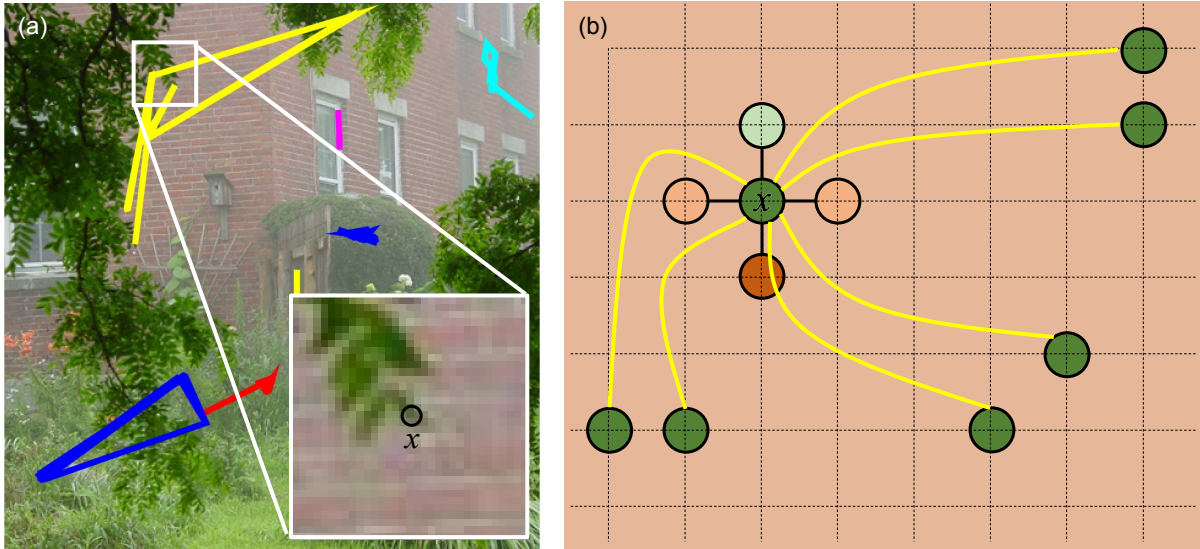


Figure 4.3: (a) The picture shows some sampled NNFs that associate pixels having similar scene depths. The line with the same color denotes association of pixels in the same NNF. (b) An MRF model of the node x from the patch in (a) associated with adjacent four neighbors and distant neighbors in the NNF. Since the node x is located in the end point of the leaf, its adjacent pixels have very different transmission values due to the 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.

4.2.2 Iso-Depth Neighbor Fields

In conventional grid MRFs, a prior term (smoothness term in this text) associates adjacent four pixels as neighbors for regularization. However, pixels in a patch lying on an edge may be isolated when the scene surface has a complicated shape. In Figure 4.3(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 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.

While Besse et al. use the PatchMatch algorithm [2] to rapidly solve non-parametric belief propagation [5], 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 as shown in Figure 4.4. This

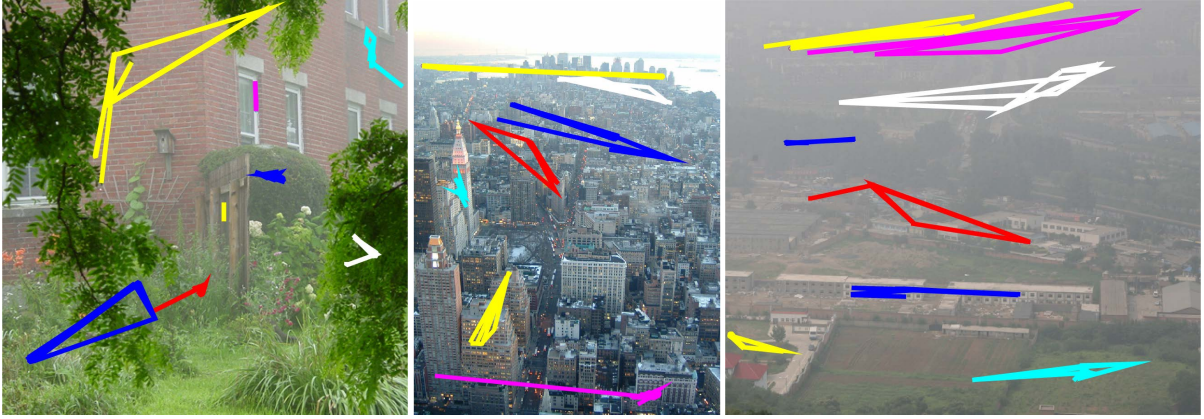


Figure 4.4: Hazy images with sampled NNFs. The line with the same color denotes association of pixels in the same NNF. As the pictures show, the NNFs associate pixels that have similar depths. By adding these neighbors to MRFs, we can connect isolated pixels shown in Figure 4.3 for more accurate regularization. As a result, the haze isolation problem can be mitigated with the insightful information.

insightful information gives a more reliable regularization penalty since the neighboring nodes in the NNF are likely to have similar transmission estimates.

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 [10, 16]. After regularization, we use the weighted median filter [31] to refine the transmission map. Algorithm 1 summarizes our dehazing algorithm as an overview.

Algorithm 1 Dehazing using Non-Local Regularization

Require: an image I

Ensure: a result image J and a transmission map t

- 1: $I_L \leftarrow \text{INVERSEGAMMACORRECTION}(I)$
 - 2: $A \leftarrow \text{ATMOSPHERICVECTORESTIMATION}(I_L)$
 - 3: **for** pixels $x = 1$ to n **do**
 - 4: $I_A(\Omega) \leftarrow I_L(\Omega) \cdot \frac{A}{\|A\|}$
 - 5: $I_A^{\min}(\Omega) \leftarrow P_i(I_A(k))_{k \in \Omega}$
 - 6: $t'(x) \leftarrow 1 - f(\theta) \cdot I_A^{\min}(\Omega)$
 - 7: $\hat{t}(x) \leftarrow \text{OUTLIERREJECTION}(t'(x), A, I_L(x))$
 - 8: **end for**
 - 9: $NNF \leftarrow \text{PATCHMATCH}(I)$
 - 10: $t \leftarrow \text{REGULARIZATION}(NNF, \hat{t}, I)$
 - 11: $J_L \leftarrow (I - (1 - t)A) / t$
 - 12: $J \leftarrow \text{GAMMACORRECTION}(J_L)$
-

Chapter 5. Results

In this chapter, we validate our dehazing method with our non-local regularization method with other existing algorithms. We evaluated our algorithm with a large number of outdoor hazy images obtained from [10] to prove robustness, and we also present comparisons with state-of-the-art dehazing methods.

5.1 External Evaluation

5.1.1 Qualitative Comparison

Figures 5.8, 5.9, 5.10, 5.11, and 5.12 qualitatively validate the robust performance in dehazing the common reference dataset of hazy scenes [10]. We compare the performance of our dehazing algorithm with three state-of-the-art methods [12, 10, 3]. We were motivated to achieve consistent performance of dehazing with less parameter controls like other image processing algorithms [?, ?]. Our method shows competitive results to other method [10] that requires manual tweaking parameters per scene to achieve plausible results.

5.1.2 Quantitative Comparison

We compare our method with the entire synthetic hazy image dataset provided by [10]. The synthetic hazy images were generated by datasets that contain clear indoor and outdoor scenes, and their corresponding depth maps. Table 5.1 reports the quantitative comparison of our method with other methods [12, 10, 3]. We also show the dehazed images used for the quantitative comparison in Figure 5.1. Our method shows competitive and consistent results particularly in dehazed images.

	He et al. [12]	Fattal14 [10]	Berman et al. [3]	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

Table 5.1: Quantitative comparisons of our method with other methods [12, 10, 3]. The error values are computed from the entire synthetic hazy image dataset provided by [10]. All figures represent mean L_1 error of the estimated transmission t (left value) and output image J (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.

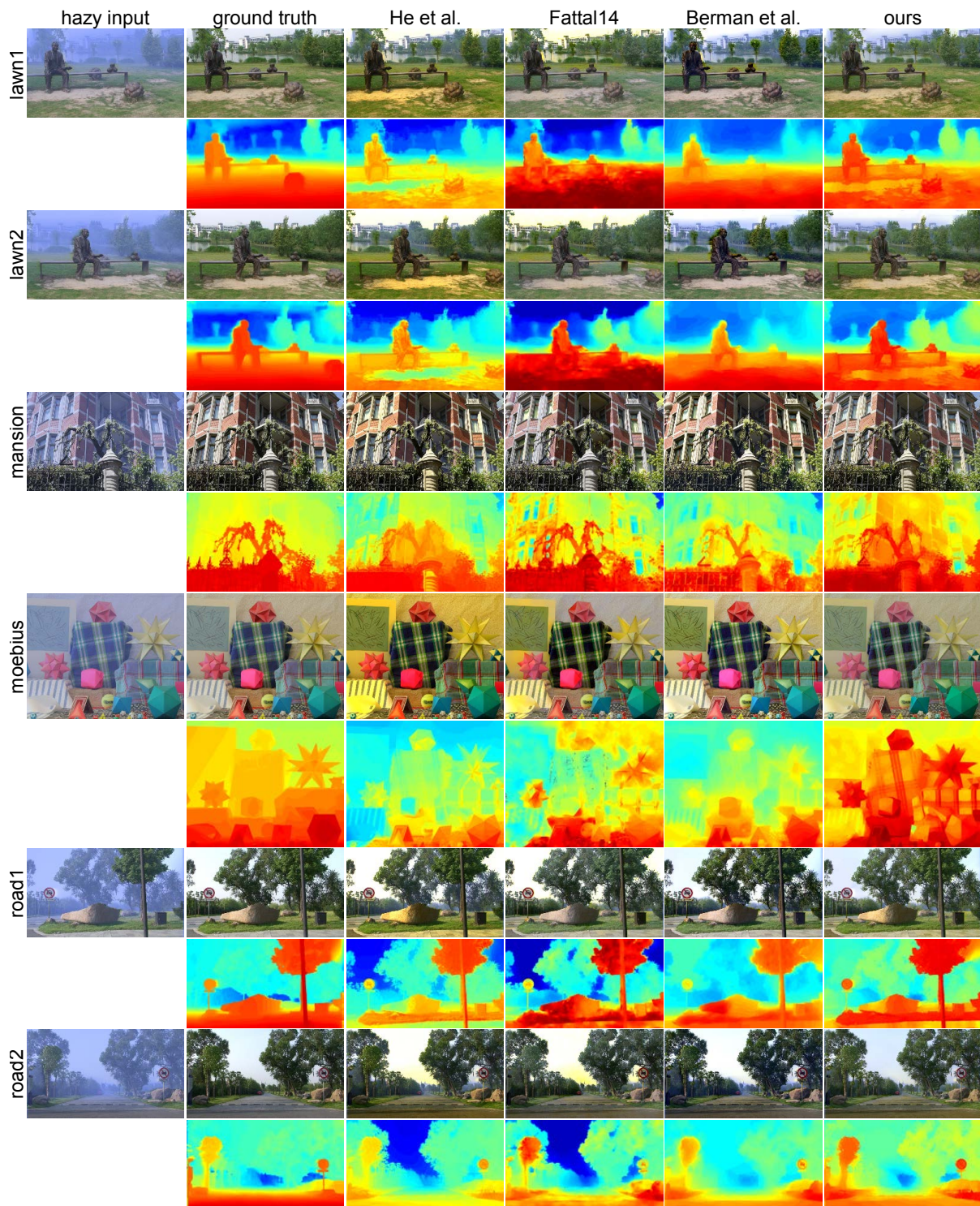


Figure 5.1: Dehazed results for the quantitative comparison shown in Table 5.1. The first column shows synthetic hazy images generated from the ground truth dataset [10] in the second column with their corresponding depth maps. The remaining columns are recovered scene radiance maps by each method. Each even row shows transmission maps corresponding to each algorithm. Our method yields consistent results compared with other methods. Parameters for each method were optimized for the highest performance for a fair comparison.

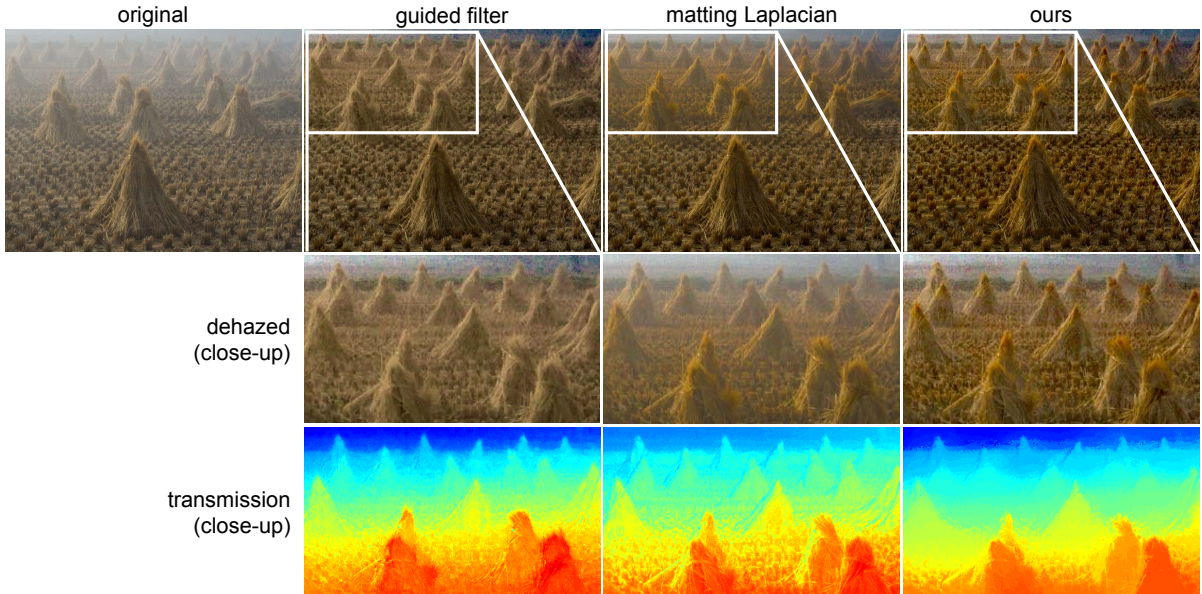


Figure 5.2: We compare our regularization with other methods. The leftmost one is the original image of cones. The first row shows dehazed results with our transmission estimation step and each regularization method written at the lower right. We cropped the dehazed images in the first row to highlight 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.

5.1.3 Regularization

We compare results of our method with those of state-of-the-art methods in terms of regularization. Berman et al. regularize initial transmission estimates with a grid GMRF as shown in third and fourth columns in Figure 5.3 [3]. Due to the lack of non-local information in regularization, some regions suffer from the haze isolation problem as mentioned above. Other than using a grid MRF, Fattal takes an augmented GMRF model for regularization, which extends neighbor fields within a local window [10]. However, it does not connect more neighbors for all pixels due to time complexity. As a result, some regions are not fully recovered from the haze isolation problem. Figure 5.3 validates that our method successfully removes haze even from a scene having abrupt depth changes with complicated patterns.

Figure 5.4 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 Figure 5.4(d) that has outliers [represented as black pixels in Figure 5.4(d)]. In particular, Figures 5.4(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 Figure 5.4(e) shows. We account for additional neighbors from NNFs to obtain a clearer transmission map shown in Figure 5.4(f). Figure 5.4(g) shows the final transmission map that we refine with a weighted median filter [31].

We also compare our regularization method with representative matting methods: the matting Laplacian method [14] and the guided filter method [11] in Figure 5.2. 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 Section 4.1, 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.

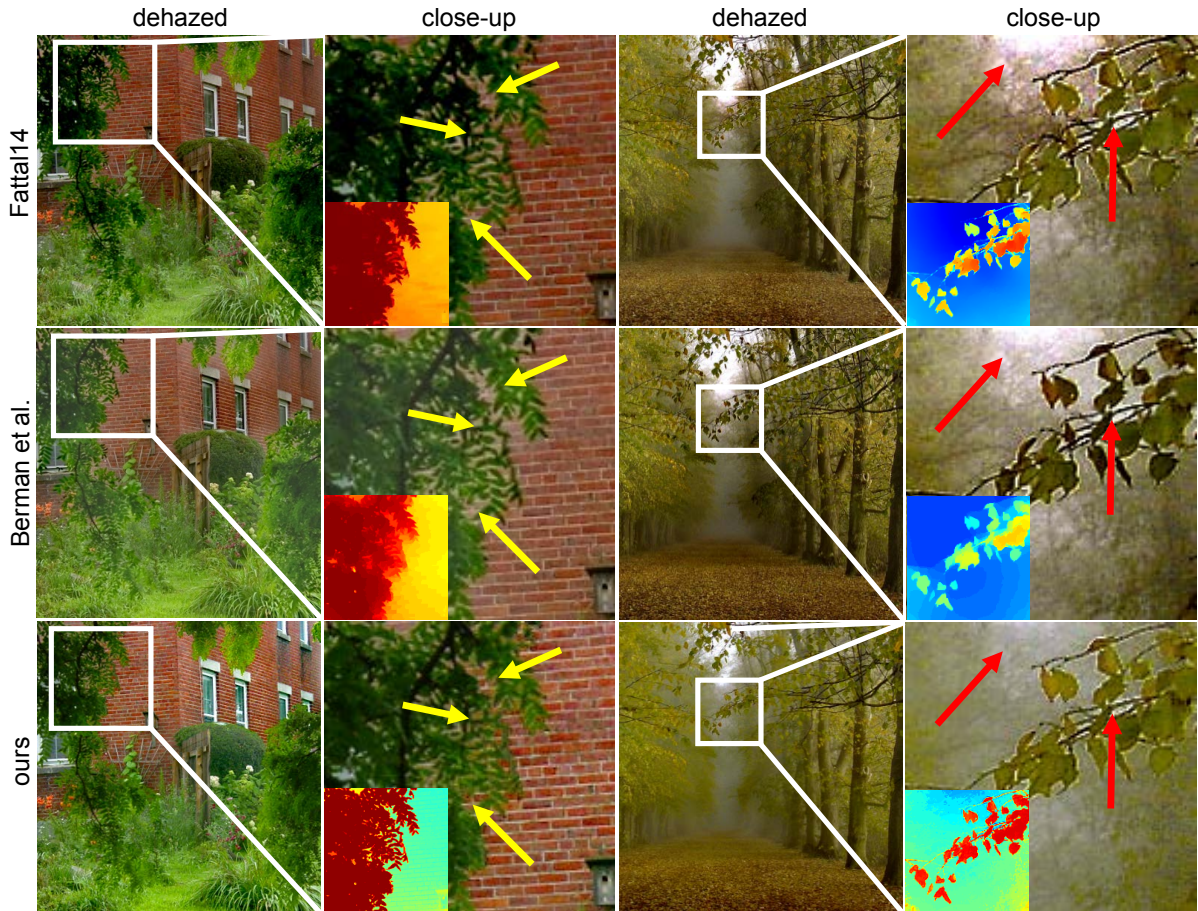


Figure 5.3: Comparisons of dehazing in terms of regularization. The two rows are results from other two methods: [10] using augmented GMRFs and [3] using traditional GMRFs, and the last row shows our results (Insets: corresponding transmission maps). While other methods often fail to obtain sharp edge-discontinuities in the images, our method yields clear recovered scene radiance maps as shown above. Some notable regions are pointed with arrows.

As a result, our regularization method yields an accurate transmission map with clear-edge discontinuities while the others overestimate the transmission estimates in turn.

5.2 Internal Evaluation

5.2.1 Running Time

We implemented our algorithm in a non-optimized MATLAB environment except the external Patch-Match algorithm [2], 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 Figure 4.1(a), our algorithm took 6.44 seconds for running the PatchMatch algorithm to seek 17 neighbors, 8.32 seconds for estimating an atmospheric vector, transmission values and rejecting outliers, 43.43 seconds for our regularization stage, and 0.65 seconds for running the weighted median filter and recovering the scene radiance, taking approximately 58.84 seconds in total.

Table 5.2 compare the computational performance of our method with traditional grid GMRFs and our iso-depth GMRFs using images shown in Figure 5.8. We also shows computational costs of obtaining

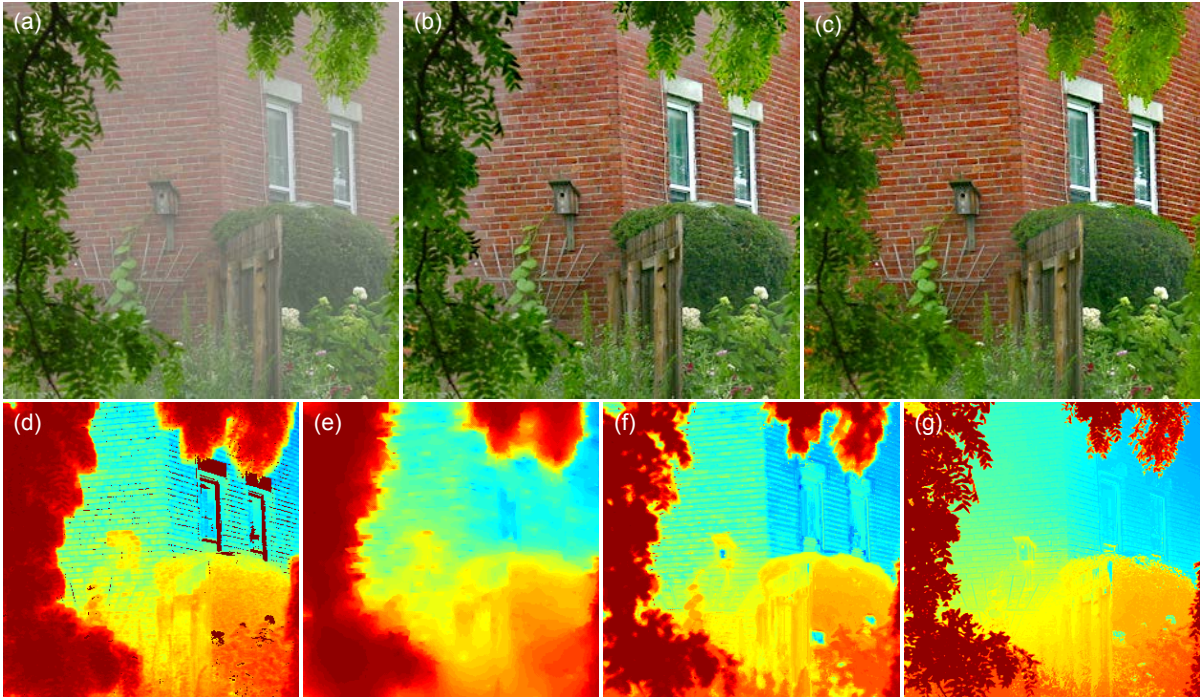


Figure 5.4: We present an example before and after applying our dehazing and regularization method. (a) The hazy input image. (b) The recovered scene radiance map with the transmission map regularized by grid MRFs (e). (c) The recovered scene radiance map with the final transmission map (g). Images (d)–(g) 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.

only NNFs with 17 neighbors using PatchMatch [2] 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.

5.2.2 Impact of Patch Size

Figure 5.5 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 is exacerbated in distant regions where their scene depth

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 (for computing NNFs only)	58.84 (6.44)	305.48 (31.82)	305.06 (28.48)	73.06 (7.15)	191.76 (18.54)	129.18 (11.01)	61.12 (7.31)	160.64 (15.82)

Table 5.2: Comparison of time performance of dehazing with the traditional grid GMRFs and our GMRFs with iso-depth NNFs (unit: second). Refer to Figures 5.8, 5.9, 5.10, 5.11, and 5.12 for processed images. The third row shows computational costs of only seeking NNFs with 17 neighbors using PatchMatch [2] in our method.



Figure 5.5: 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 , where the patch size 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.

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.

5.2.3 Outlier Removal

We validate our outlier-rejection process. Figure 5.6 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 Section 4.1.3. We estimated an atmospheric vector under the assumption that the atmospheric light is the brightest all over a scene.



Figure 5.6: Validation of our narrow angle outlier rejection method described in Section 4.1.3. 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.

As Figure 5.7 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.

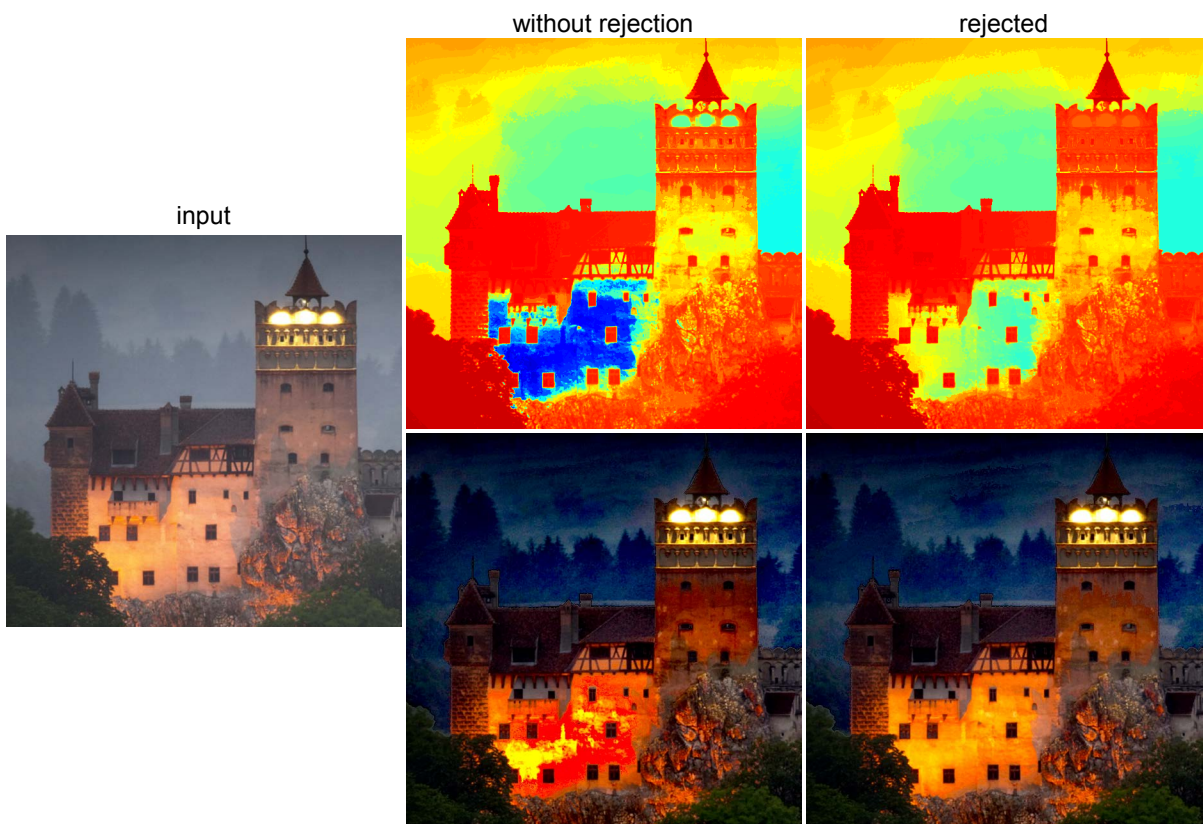


Figure 5.7: Validation of our saturated intensity outlier rejection process. In the second column, the bright regions (the direct light at the upper right and the wall of the castle in the middle) are overly saturated. Our outlier rejection succeeds to produce a consistent result by discarding those regions.

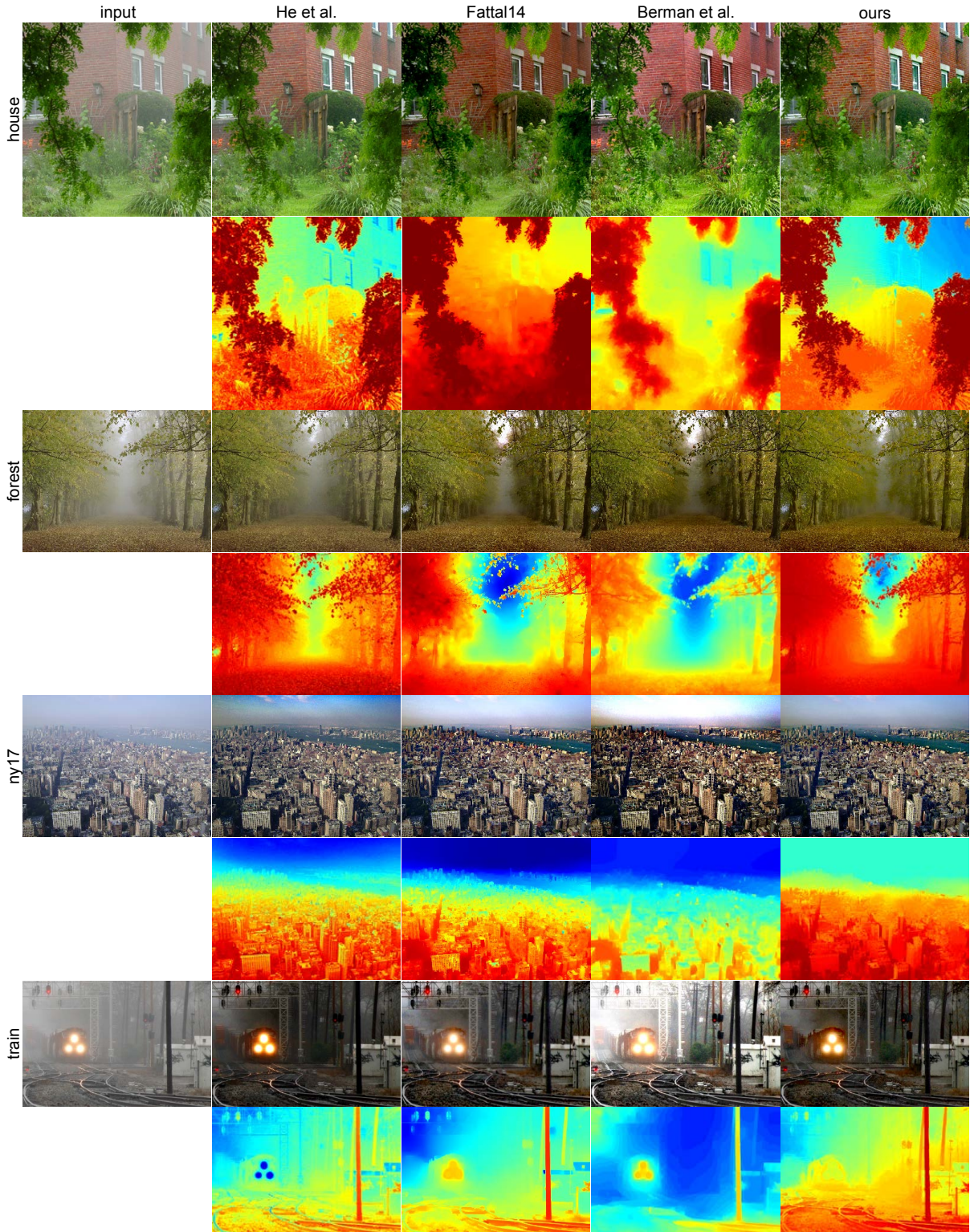


Figure 5.8: Validation of consistency of dehazing. The first column shows input images. The second, third, and fourth columns are results from [12, 10, 3], respectively. The fifth column presents our method’s results. Each odd row shows dehazed results, and each even row represents corresponding transmission maps. We use the set of parameters as described in Section 4.1.3. For the case of ny17 image, we 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 [10, 3] that requires with manual tweaking parameters to achieve plausible results.

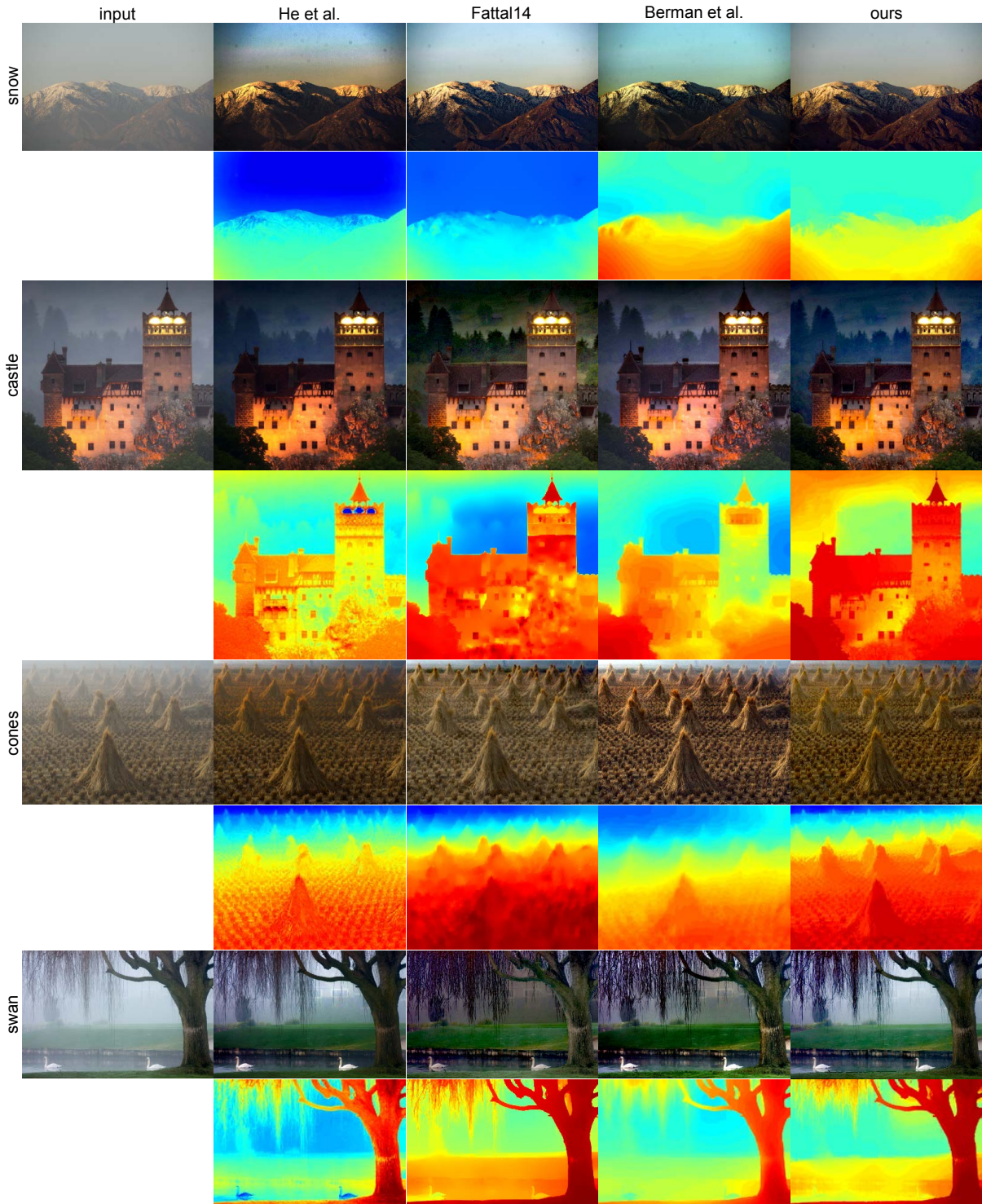


Figure 5.9: Validation of consistency of dehazing. The first column shows input images. The second, third, and fourth columns are results from [12, 10, 3], respectively. The fifth column presents our method's results. Each odd row shows dehazed results, and each even row represents corresponding transmission maps. We use the set of parameters as described in Section 4.1.3. For the case of snow image, we 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 [10, 3] that requires with manual tweaking parameters to achieve plausible results.

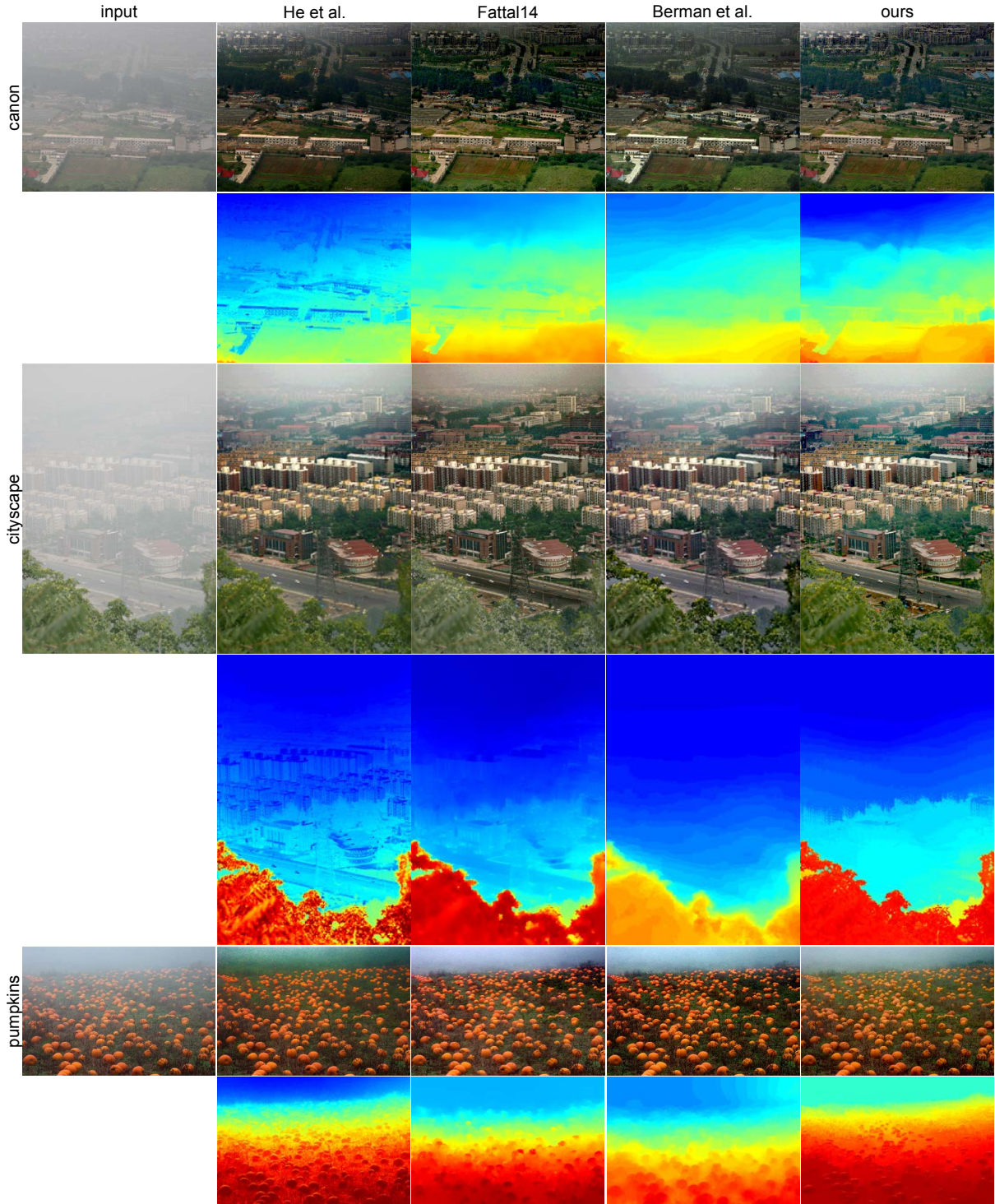


Figure 5.10: Validation of consistency of dehazing. The first column shows input images. The second, third, and fourth columns are results from [12, 10, 3], respectively. The fifth column presents our method's results. Each odd row shows dehazed results, and each even row represents corresponding transmission maps. We use the set of parameters as described in Section 4.1.3. For the case of pumpkins image, we 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 [10, 3] that requires with manual tweaking parameters to achieve plausible results.

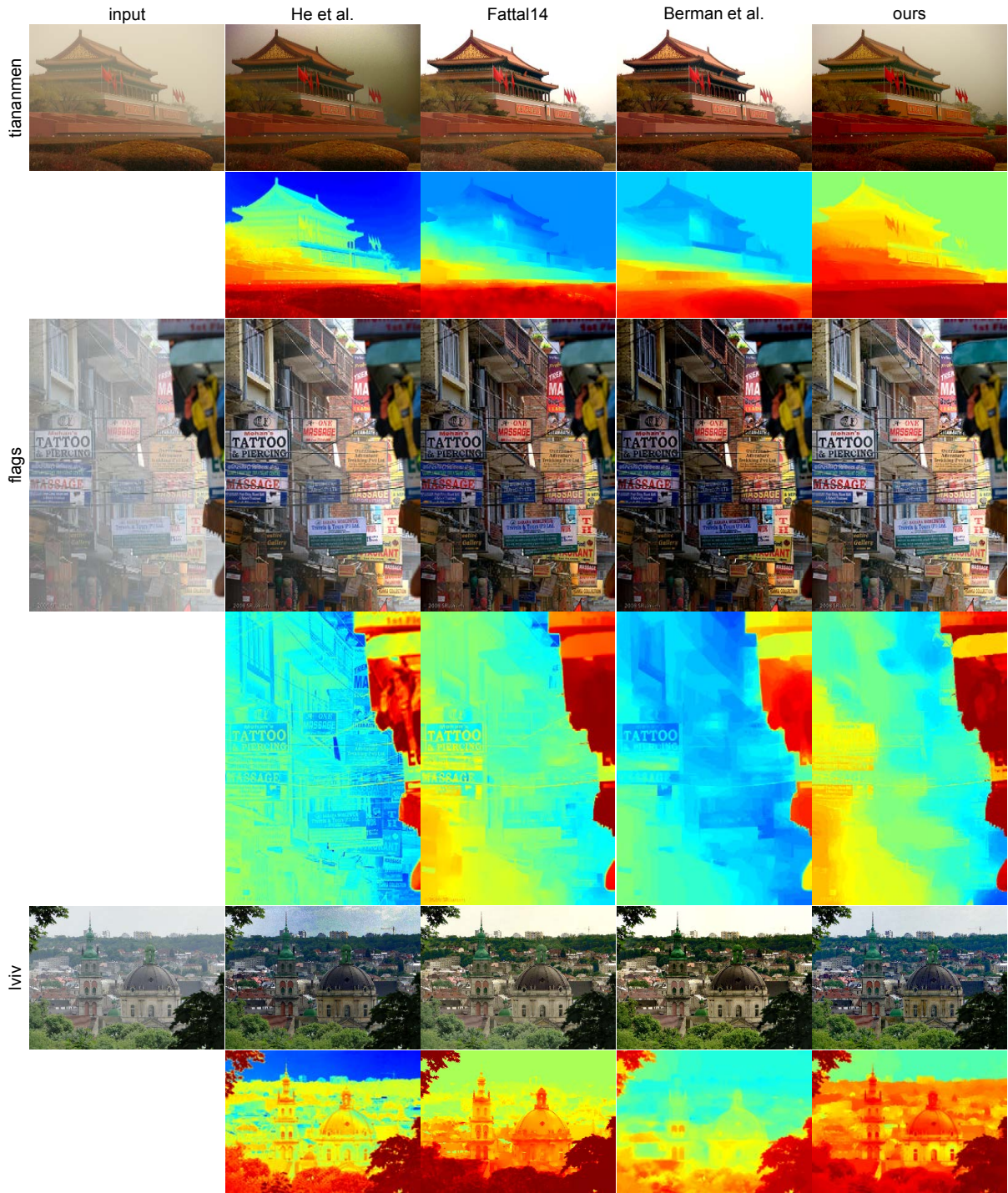


Figure 5.11: Validation of consistency of dehazing. The first column shows input images. The second, third, and fourth columns are results from [12, 10, 3], respectively. The fifth column presents our method's results. Each odd row shows dehazed results, and each even row represents corresponding transmission maps. We use the set of parameters as described in Section 4.1.3. For the cases of tiananmen and lviv images, we 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 [10, 3] that requires with manual tweaking parameters to achieve plausible results.

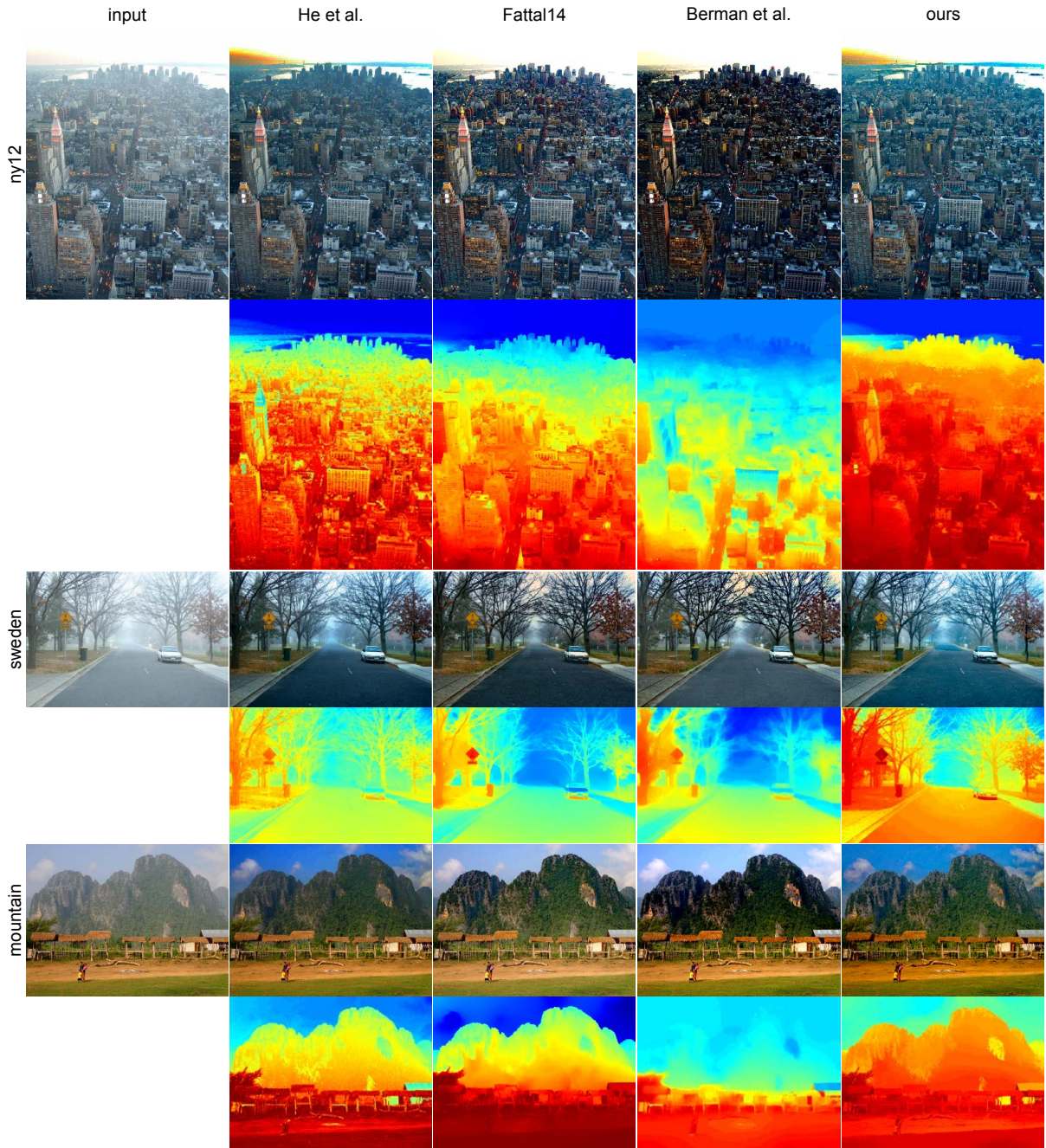


Figure 5.12: Validation of consistency of dehazing. The first column shows input images. The second, third, and fourth columns are results from [12, 10, 3], respectively. The fifth column presents our method's results. Each odd row shows dehazed results, and each even row represents corresponding transmission maps. We use the set of parameters as described in Section 4.1.3. For the cases of ny12 and mountain images, we 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 [10, 3] that requires with manual tweaking parameters to achieve plausible results.

Chapter 6. Discussion and Future Work

We discuss a range of observations and limitations made throughout the development of our method.

6.1 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 6.1 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 selected pixels in cloud regions as candidates of the atmospheric light, which is not correct. For this reason, our transmission estimation stage severely overestimated the amount of airlight, particularly in distant regions in the scene as shown in Figure 6.1(b). We validated the limitation by picking up the atmospheric vector of the image manually, and our algorithm yielded a naturally-looking result, as the Figure 6.1(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.

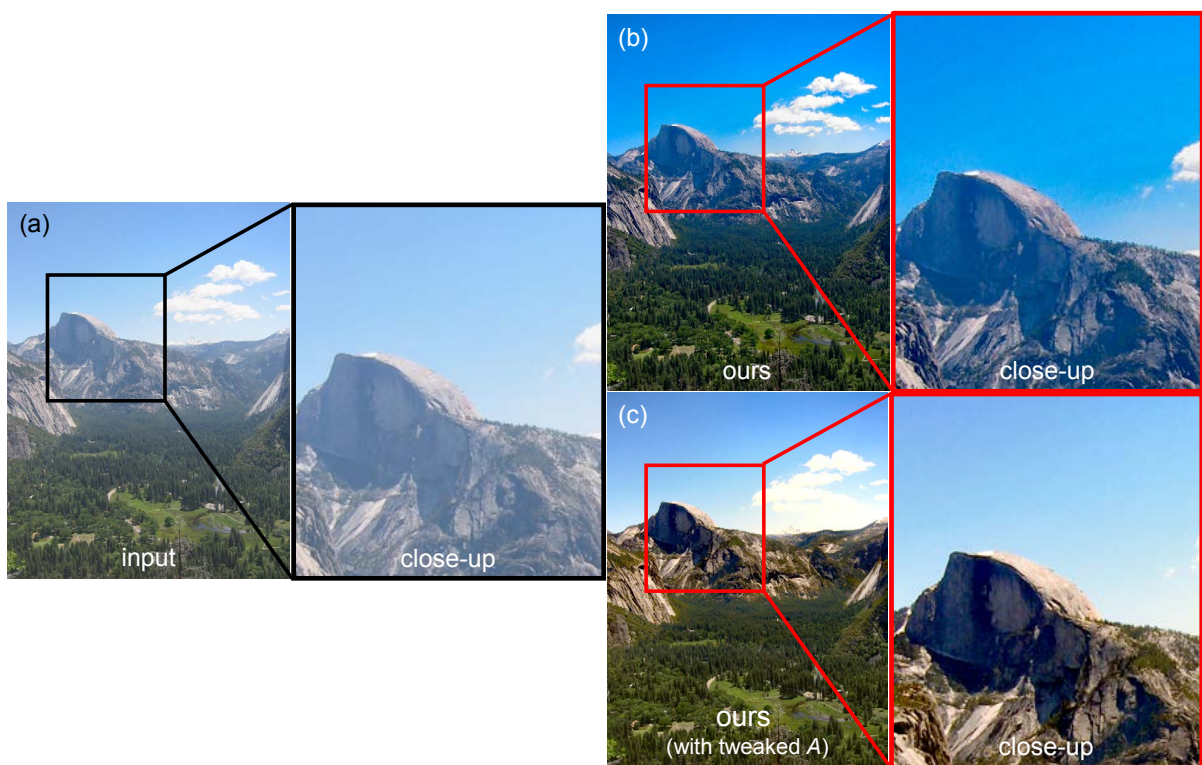


Figure 6.1: Our failure case with a landscape image. Image (a) shows the input image, and Image (b) presents our result with the same set of parameters described. Image (c) is our result produced with the manually-tweaked atmospheric vector.

Chapter 7. Conclusion

We have presented a dehazing method based on marginalization. Unlike existing single image dehazing methods, our method focuses on the airlight component in the hazy image formation model rather than relying upon scene radiance priors. Thanks to the simplicity of airlight properties, our algorithm yields consistent results. In addition, we use consistent parameters over all input images, thereby also reducing the necessity of cumbersome scene-specific parameter tweaks. We also proposed non-local regularization with iso-depth neighbor fields. Even though regularization is an essential process in dehazing, traditional GMRF-based regularization methods often fail with isolation artifacts when there is an abrupt change in depth, of which information is missing in single-image dehazing. We propose a novel non-local regularization method that utilizes NNFs searched in a hazy image to infer depth cues to obtain more reliable smoothness penalty for handling the isolation problem. We validated the robust performance of our method with extensive test images and compared it with the state-of-the-art single image-based methods. This proposed regularization method can be used separately with any other dehazing algorithms to enhance haze regularization.

Bibliography

- [1] C. O. Ancuti and C. Ancuti. Single image dehazing by multi-scale fusion. *IEEE Trans. Image Processing*, 22(8):3271–3282, 2013.
- [2] C. Barnes, E. Shechtman, A. Finkelstein, and D. B. Goldman. Patchmatch: a randomized correspondence algorithm for structural image editing. *ACM Trans. Graph*, 28(3):24:1–24:11, 2009.
- [3] D. Berman, T. Treibitz, and S. Avidan. Non-local image dehazing. In *IEEE CVPR*, pages 1674–1682, 2016.
- [4] J. Besag. On the statistical analysis of dirty pictures. *JOURNAL OF THE ROYAL STATISTICAL SOCIETY B*, 48(3):48–259, 1986.
- [5] F. Besse, C. Rother, A. W. Fitzgibbon, and J. Kautz. PMBP: Patchmatch belief propagation for correspondence field estimation. *International Journal of Computer Vision*, 110(1):2–13, 2014.
- [6] Y. Boykov, O. Veksler, and R. Zabih. Fast approximate energy minimization via graph cuts. *IEEE Trans. Pattern Anal. Mach. Intell.*, 23(11):1222–1239, 2001.
- [7] P. Carr and R. I. Hartley. Improved single image dehazing using geometry. In *DICTA 2009*, pages 103–110, 2009.
- [8] D. Cremers, T. Pock, K. Kolev, and A. Chambolle. Convex relaxation techniques for segmentation, stereo and multiview reconstruction. In *Markov Random Fields for Vision and Image Processing*. MIT Press, 2011.
- [9] R. Fattal. Single image dehazing. *ACM Trans. Graph.*, 27(3):72:1–72:9, 2008.
- [10] R. Fattal. Dehazing using color-lines. *ACM Trans. Graph.*, 34(1):13:1–13:14, 2014.
- [11] K. He, J. Sun, and X. Tang. Guided image filtering. *IEEE Trans. Pattern Anal. Mach. Intell.*, 35(6):1397–1409, 2013.
- [12] K. M. He, J. Sun, and X. Tang. Single image haze removal using dark channel prior. In *Proc. IEEE CVPR*, pages 1956–1963, 2009.
- [13] J. Kopf, B. Neubert, B. Chen, M. F. Cohen, D. Cohen-Or, O. Deussen, M. Uyttendaele, and D. Lischinski. Deep photo: model-based photograph enhancement and viewing. *ACM Trans. Graph.*, 27(5):116:1–116:10, 2008.
- [14] A. Levin, D. Lischinski, and Y. Weiss. A closed-form solution to natural image matting. *IEEE Trans. Pattern Anal. Mach. Intell.*, 30(2):228–242, 2008.
- [15] Y. Li, R. T. Tan, and M. S. Brown. Nighttime haze removal with glow and multiple light colors. In *2015 IEEE ICCV 2015, Santiago, Chile, December 7-13, 2015*, pages 226–234, 2015.
- [16] Y. P. Li and D. P. Huttenlocher. Sparse long-range random field and its application to image denoising. In *ECCV*, pages III: 344–357, 2008.

- [17] J. L. Marroquín, F. A. Velasco, M. Rivera, and M. Nakamura. Gauss-markov measure field models for low-level vision. *IEEE Trans. Pattern Anal. Mach. Intell.*, 23(4):337–348, 2001.
- [18] G. Meng, Y. Wang, J. Duan, S. Xiang, and C. Pan. Efficient image dehazing with boundary constraint and contextual regularization. In *Proc. IEEE ICCV*, pages 617–624, 2013.
- [19] S. G. Narasimhan and S. K. Nayar. Vision and the atmosphere. *Int. Journal of Computer Vision*, 48(3):233–254, 2002.
- [20] S. G. Narasimhan and S. K. Nayar. Contrast restoration of weather degraded images. *IEEE Trans. Pattern Anal. Mach. Intell.*, 25(6):713–724, 2003.
- [21] K. Nishino, L. Kratz, and S. Lombardi. Bayesian defogging. *Int. Journal of Computer Vision*, 98(3):263–278, 2012.
- [22] J. Pearl. Reverend bayes on inference engines: A distributed hierarchical approach. In *Proceedings of the National Conference on Artificial Intelligence. Pittsburgh, PA, August 18-20, 1982.*, pages 133–136, 1982.
- [23] K. G. G. Samuel and M. F. Tappen. Learning optimized MAP estimates in continuously-valued MRF models. In *2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2009), 20-25 June 2009, Miami, Florida, USA*, pages 477–484, 2009.
- [24] Y. Y. Schechner, S. G. Narasimhan, and S. K. Nayar. Instant dehazing of images using polarization. In *Proc. IEEE CVPR*, pages I:325–332, 2001.
- [25] R. Szeliski. Bayesian modeling of uncertainty in low-level vision. *International Journal of Computer Vision*, 5(3):271–301, 1990.
- [26] R. T. Tan. Visibility in bad weather from a single image. In *Proc. IEEE CVPR*, pages 1–8, 2008.
- [27] K. Tang, J. Yang, and J. Wang. Investigating haze-relevant features in a learning framework for image dehazing. In *Proc. IEEE CVPR*, pages 2995–3002, 2014.
- [28] J. Tarel and N. Hautière. Fast visibility restoration from a single color or gray level image. In *Proc. IEEE ICCV*, pages 2201–2208, 2009.
- [29] A. J. Viterbi. Error bounds for convolutional codes and an asymptotically optimum decoding algorithm. *IEEE Trans. Information Theory*, 13(2):260–269, 1967.
- [30] M. J. Wainwright, T. S. Jaakkola, and A. S. Willsky. MAP estimation via agreement on trees: message-passing and linear programming. *IEEE Trans. Information Theory*, 51(11):3697–3717, 2005.
- [31] Q. Zhang, L. Xu, and J. Jia. 100+ times faster weighted median filter (WMF). In *CVPR*, pages 2830–2837. IEEE, 2014.
- [32] Q. Zhu, J. Mai, and L. Shao. A fast single image haze removal algorithm using color attenuation prior. *IEEE Trans. Image Processing*, 24(11):3522–3533, 2015.

Acknowledgments in Korean

감사의 글을 쓴다. 감사의 글 두 번째 문장이다.

Curriculum Vitae in Korean

이 름: 김 인 철
생 년 월 일: 1989년 12월 15일
출 생 지: 울산광역시 남구 신정동 700-9
주 소: 울산광역시 중구 태화동 동부아파트 105동 1601호
전 자 주 소: kimic89@gmail.com

학 력

2005. 3. – 2008. 2. 현대청운고등학교
2009. 3. – 2015. 2. 한양대학교 컴퓨터공학부 (B.S.)

경 력

2010. 2. – 2012.3. 공군 병장 만기 전역
2015. 3. – 2016. 12. 한국과학기술원 전산학부 일반조교

학 회 활 동

1. Incheol Kim and Min H. Kim, *Dehazing using Non-Local Regularization with Iso-Depth Neighbor-Fields*, 12th International Conference on Computer Vision Theory and Applications (VISAPP 2017), accepted (oral presentation).