이학박사 학위논문

Regularization methods for image denoising and underwater image dehazing

(영상 잡음 제거와 수중 영상 복원을 위한 정규화 방법)

2020년 2월

서울대학교 대학원

수리과학부

조 중 희

Regularization methods for image denoising and underwater image dehazing

(영상 잡음 제거와 수중 영상 복원을 위한 정규화 방법)

지도교수 강 명 주

이 논문을 이학박사 학위논문으로 제출함

2019년 10월

서울대학교 대학원

수리과학부

조 중 희

조 중 희의 이학박사 학위논문을 인준함

2019년 12월

위 원 장 _____ (인)

부 위 원 장 _____ (인)

위 원 _____ (인)

위 원 _____ (인)

위 원 _____ (인)

Regularization methods for image denoising and underwater image dehazing

A dissertation
submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
to the faculty of the Graduate School of
Seoul National University

by

Junghee Cho

Dissertation Director : Professor Myungjoo Kang

Department of Mathematical Science
Seoul National University

February 2020

Abstract

Regularization methods for image denoising and underwater image dehazing

In this thesis, we discuss regularization methods for denoising images corrupted by Gaussian or Cauchy noise and image dehazing in underwater. In image denoising, we introduce the second-order extension of structure tensor total variation and propose a hybrid method for additive Gaussian noise. Furthermore, we apply the weighted nuclear norm under nonlocal framework to remove additive Cauchy noise in images. We adopt the nonconvex alternating direction method of multiplier to solve the problem iteratively. Subsequently, based on the color ellipsoid prior which is effective for restoring hazy image in the atmosphere, we suggest novel dehazing method adapted for underwater condition. Because attenuation rate of light varies depending on wavelength of light in water, we apply the color ellipsoid prior only for green and blue channels and combine it with intensity map of red channel to refine the obtained depth map further. Numerical experiments show that our proposed methods show superior results compared with other methods both in quantitative and qualitative aspects.

Key words: Denoising, Dehazing in underwater, Gaussian noise, Cauchy noise, Structure tensor, Weighted nuclear norm, Color ellipsoid prior
Student Number: 2013-22914

Contents

Abstract **i**

1 Introduction **1**

1.1 Image denoising for Gaussian and Cauchy noise 2

1.2 Underwater image dehazing 5

2 Preliminaries **9**

2.1 Variational models for image denoising 9

2.1.1 Data-fidelity 9

2.1.2 Regularization 11

2.1.3 Optimization algorithm 14

2.2 Methods for image dehazing in the air 15

2.2.1 Dark channel prior 16

2.2.2 Color ellipsoid prior 19

3 Image denoising for Gaussian and Cauchy noise **23**

3.1 Second-order structure tensor and hybrid STV 23

3.1.1 Structure tensor total variation 24

3.1.2 Proposed model 28

3.1.3 Discretization of the model 31

3.1.4 Numerical algorithm 35

3.1.5 Experimental results 37

3.2 Weighted nuclear norm minimization for Cauchy noise 46
3.2.1 Variational models for Cauchy noise 46
3.2.2 Low rank minimization by weighted nuclear norm . . . 52
3.2.3 Proposed method 55
3.2.4 ADMM algorithm 56
3.2.5 Numerical method and experimental results 58

4 Image restoration in underwater 71
4.1 Scientific background 72
4.2 Proposed method 73
4.2.1 Color ellipsoid prior on underwater 74
4.2.2 Background light estimation 78
4.3 Experimental results 80

5 Conclusion 87

Appendices 89

A Proofs and supplementary material of Section 3.1 89
A.1 Proof of theorem 3.1.9 89
A.2 Finding the gradient step size for Algorithm 1 93

B Proofs and supplementary material of Section 3.2 95
B.1 Proof of theorem 3.2.3 95

Abstract (in Korean) 111

Acknowledgement (in Korean) 113

CONTENTS

Chapter 1

Introduction

Image processing is an inter-disciplinary field crossing boundaries between mathematics, statistics, computer science, and cognitive science. It aims at having computers deal with and understand a digitized image as human eyes can do and more. Research in this field involves object recognition, segmentation, motion estimation, image restoration and so on. It can be used in various areas ranging from face recognition and autonomous driving to analysis of medical data and defect detection in semiconductor manufacturing. Among diverse research subjects mentioned above, image restoration is a fundamental job in image processing and it seeks to restore an original image from a given image which is distorted by several factors. It covers a wide variety of tasks such as denoising, deblurring, inpainting and dehazing.

Image restoration is an inverse problem which consists of using a set of observations to infer the inputs or cause that produce the measurements. A solution of inverse problem is non-unique because there are many possible cases of input parameters that give the same observations. Therefore, it is required to make prior knowledge to extract a feasible solution out of numerous candidates. An introduction of a prior assumption to the inverse problem, which is called regularization, is achieved explicitly by insertion of additional mathematical term to the model or implicitly by optimization algo-

rithm. Because information which is contained in the regularization heavily affect the properties of acquired solution, it requires that we need meticulously designed regularization to solve the inverse problem. In this thesis, we study various regularization methods of mathematical models for denoising and underwater dehazing which are subjects of image restoration.

1.1 Image denoising for Gaussian and Cauchy noise

Image denoising aims at recovering image degraded by noise which occurs necessarily by the environmental effect on image sensor during acquisition process and instability of medium in the transmission of signal. It has been widely studied when the noise is additive and follows Gaussian distribution. Gaussian noise arises in electronic amplifier or detectors and also caused by natural sources such as thermal noise which comes from reset operation of image sensor [48]. Another type of noise, Cauchy noise, have received less attention compared with Gaussian noise despite its frequent occurrence in various scientific and engineering situations. Cauchy noise follows Cauchy distribution which are heavy-tailed and it shows impulsive characteristic in comparison to Gaussian noise. It appeared in atmospheric and underwater acoustic signals that are used in radar and sonar applications, multi-access networks, and power line communication channels [19, 23, 31, 33, 36, 58, 59, 71].

Following [55], we give basic model of the situation as image is composed of discrete rectangular pixel grid and also assume that noise acts on each pixel with independent and identically distributed manner.

Let $h > 0$ and $n_x, n_y \in \mathbb{N}$. Discrete image of size $n_x \times n_y$ is given as matrix $\mathbf{u} = (u_{ij})$ where $u_{ij} \in \mathbb{R}$ (or $\mathbb{R}^3$) indicates intensity value (or vector) at pixel $x_{ij} = (ih, jh)$ and $(i, j) \in \mathcal{I} := \{1, \cdots, n_x\} \times \{1, \cdots, n_y\}$. The parameter

h determines a distance between each pixel, in other words, a resolution of image. Then an additive noise $\mathbf{n} = (n_{ij})$ degrades the image $\mathbf{u}$ as follows:

$$\mathbf{f} = \mathbf{u} + \mathbf{n} \tag{1.1}$$

where $\mathbf{n}$ is a matrix with same size of $\mathbf{u}$ and n_{ij}'s are samples from i.i.d probability distributions. Given that we know $\mathbf{f}$ by observation, we study method of finding $\mathbf{u}$ by variational model which is adequately adapted to the property of noise applied to the image.

Generally, denoising techniques can be largely classified into three categories, which are filtering, variational model and deep learning-based method. Filter-based methods [2, 17, 21, 35, 37, 38, 52] introduce a carefully designed filters under reasonable *a priori* assumptions and apply them to a given image. Minimizing mean squared error in local window gives adaptive filtering to a noisy image [38]. In [2, 21, 52], the authors apply Fourier or wavelet transform to an image so that one can conduct filtering in the frequency domain. Laus *et al.* [37] introduced a generalized myriad filter which is adapted for Cauchy noise and apply it nonlocally for estimating the parameter of the noise distribution. Deep learning-based methods [41, 44, 68, 69] show impressive performance recently due to the huge amount of training data and deep, complex architecture of neural network. However, to the best of our knowledge, it involves considerable computational cost for training and capacity of the model is dependent on features of training data. Furthermore, mathematical analysis is deficient that explains how the model actually works.

In variational framework, we define an energy functional which reflects the properties of given noise and obtain a solution that minimizes the functional. The energy functional usually consists of a fidelity term that is derived from probability distribution of noise and regularization term that contains *a priori* information and guides a solution to have specific properties. Based on the maximum a posteriori (MAP) estimate, fidelity terms for Gaus-

sian and Cauchy noise can be derived. It is in the form of L_2-norm for Gaussian and nonconvex for Cauchy noise respectively.

Total variation (TV) [54] is one of the most widely used regularization because of its effectiveness and low computational cost. It is capable of preserving edges and discontinuities sharply while smoothing noisy parts. Subsequently, there follow various extensions of TV and they are applied to restoring an image corrupted by Gaussian noise. These models include total generalized variation [8], higher-order TV [16, 49], fractional-order TV [67] and structure tensor total variation (STV) [39]. In particular, STV generalizes TV in that it maintains all the advantageous properties of TV and utilizes additional information from the neighborhood of each point. It calculates eigenvalues of structure tensor which is defined on every point of image and contains information about local variation of image intensity. Although it shows better denoising performance compared with other TV-based methods, it still exhibits cartoon-like feature called staircase artifacts which can be seen in the results of TV-based methods.

Note that TV regularization is one of local methods which consider information of adjacent regions of interested point. In contrast, there exist nonlocal methods [10, 21, 27, 29, 34] which utilize information of every point in image. They show superior performance over local methods because of differences in the amount of information used for each type of methods. Among nonlocal methods, weighted nuclear norm minimization (WNNM) [29] assumes a low rank property of matrix whose column consists of vectorized similar patches in image. The rank of a matrix can be approximated by nuclear norm of a matrix [12–14], i.e., the sum of singular values of a matrix and it is usually used in the area of matrix completion. In accordance with previous works, the low rank assumption is put into practice by using weighted nuclear norm of the patch-based matrix as a regularization term.

When it comes to restoring an image corrupted by Cauchy noise, there exist only a few variational models [45, 56] until recently. Both are based on

TV regularization and the difference is that the model becomes convexified by inserting additional quadratic term in [56] while the authors of [45] solve the nonconvex model directly by applying nonconvex alternating direction method of multipliers (ADMM).

In this thesis, in Section 3.1, we suggest a new regularization called second-order structure tensor total variation (STV2) which is a higher order extension of STV for denoising Gaussian noise. For defining STV2, we construct second-order structure tensor from vectorial form of hessian as an analogue of the structure tensor. We propose the hybrid STV which is a convex combination of STV and STV2. In Section 3.2, we propose to use the weighted nuclear norm for denoising Cauchy noise. Because the data-fidelity term of the model is nonconvex, we apply the nonconvex ADMM to solve the problem iteratively.

1.2 Underwater image dehazing

Image dehazing aims to restore outdoor images which are often degraded by haze, fog and smoke in the atmosphere. The air becomes turbid when it is mixed with solid and liquid particles like dust or water-droplets. In the presence of haze in the air, scene radiance from the object in visual range undergoes degradation by absorption and scattering. As light travels through the haze, it is attenuated along the line of sight. Furthermore, background light from the scattering of ambient light due to haze is added to the directly attenuated light. As a result of the above effects, contrast and color of an image are distorted and a visibility of the scene is limited. As the distance between camera and the scene point is farther, the amount of degradation by haze increases. It is represented by transmission (visibility) of the scene point. The image formation model for hazy scene is given as a convex combination of the scene radiance and global atmospheric light (background light) where the transmission is a balancing coefficient between two terms.

If we take an image in underwater instead of air, degradation of the image is also caused by absorption and scattering of light due to medium, but exhibits different aspects from the degradation by haze in the air. First of all, in underwater, the amount of attenuation of light depends on the wavelength of light as it travels through the medium while, in atmosphere, we deal with the attenuation process independently of wavelength of light. Additionally, there are several types of water that exhibits different degree of attenuation with respect to wavelength. In optical oceanography, water is categorized into 10 classes such as five oceanic and five coastal types based on the diffuse downwelling coefficients [57]. Accordingly, transmission map of the scene point depends on the color channel of image and type of water in which image is taken. The image formation model for underwater scene is a convex combination of scene radiance and global background light where the transmissions playing a role of balancing coefficient in the model vary depending on the color channel of the image.

For restoration of hazy outdoor images, many methods have been suggested [5, 11, 24, 26, 30, 61]. In [30], the authors propose a simple yet effective image prior called dark channel prior. They observe that most local patches in clean outdoor images possess some pixels which have very low intensities at least one color channel. From this observation, they estimate transmission map of the hazy image and recover a hazy-free image of good quality. But it has some problems such as blocky artifacts and invalidity to large white region. After that, in [11], a method to supplement the dark channel prior is suggested. By statistically removing noisy pixels which are unrelated to most of pixels in a patch, they estimate the transmission map robustly. In [61], the authors propose an algorithm that estimates global atmospheric light by continuously decomposing the image into quarters and selecting one region of them according to intensity and variation of the region.

For underwater image restoration problems, there also exist various meth-

ods [3, 4, 18, 22, 42, 50, 70]. Applying the dark channel prior directly to the underwater image is not effective because rapid attenuation of red light in underwater disturb an accurate estimation of transmission map. To resolve this drawback, the authors of [22] propose to use information of blue and green channels when applying the dark channel prior. This method brings an improvement over original dark channel prior method, but they assume the uniform transmission across color channels. In [42], the authors recover blue and green channels following the algorithm in [22]. Then, red channel is corrected under the Gray-World assumption. In [50], rather than transmission, depth estimation method for underwater scenes is proposed. It is based on the fact that the intensity of blurriness and attenuation of red light gets stronger as the scene depth becomes deeper. In [3, 4], a revised underwater image formation model is proposed so that it reflects the physical properties of attenuation in underwater more accurately. Specifically, the authors suggest that direct and back-scattered signals are governed by distinct coefficients and each of these coefficients depend on factors other than optical properties of the water. Although the proposed model in [4] is more accurate than previous one, it is demanding to solve the new model because it has far more unknowns than given number of equations. Therefore, one needs more prior assumptions and additional information such as depth to carry out restoration task.

In this thesis, in Chapter 4, we introduce a novel method for restoring images taken in underwater. We estimate a depth map of image utilizing two kinds of information. First, we apply the color ellipsoid prior [11] to blue and green channels to get a raw depth map. Then, convex combination of the obtained depth map with normalized red channel gives the refined depth map. To avoid artifacts caused by the use of patch-based method, we adopt the superpixel segmentation method [1] when applying the color ellipsoid prior.

The rest of this thesis is organized as follows: in Chapter 2, we recall previous works related to regularization methods and optimization algorithms

for image denoising and dehazing. In Chapter 3, we propose our regularization methods for denoising an image corrupted by Gaussian and Cauchy noise respectively. For each type of noise, we study the numerical algorithms for solving the proposed models, present experimental results and compare them with other denoising methods. In Chapter 4, a novel method for restoring underwater image is proposed. We present experimental results of the proposed method and compare them with other existing methods. In Chapter 5, we conclude the thesis with a few remarks.

Chapter 2

Preliminaries

In this chapter, we investigate several existing image restoration models with regularization methods and numerical algorithms for solving them, which would be helpful for grasping the materials in the following chapters.

2.1 Variational models for image denoising

Let $\Omega \subset \mathbb{R}^2$ be an image domain. Suppose we have a noisy image $f : \Omega \to \mathbb{R}$ corrupted by additive noise n and it is represented as follows:

$$f = u + n \tag{2.1}$$

where $u : \Omega \to \mathbb{R}$ is a latent clean image. Unlike (1.1), we assume (2.1) to be defined on continuous domain Ω. At each point x of Ω, the noise n is realized from a specific probability distribution with i.i.d. manner.

2.1.1 Data-fidelity

Variational model for image denoising mostly consists of a data-fidelity term and regularization term. A form of data-fidelity term depends on the

property of given noise and it is derived from maximum a priori (MAP) estimate. Let us assume that the noise n follows Gaussian distribution $N(0, \sigma^2)$. From MAP principle, we expect that the probability of finding latent parameter u should maximize among possible parameters given the observation of noisy image f. In other words,

$$u^* = \arg\max_u P(u|f).$$

Then, from Bayes' rule,

$$u^* = \arg\max_u \frac{P(f|u)P(u)}{P(f)} = \arg\max_u P(f|u)P(u). \tag{2.2}$$

Now, we take minus logarithm of right side of (2.2) to have the following:

$$\begin{aligned} u^* &= \arg\min_u -\log P(f|u) - \log P(u) \\ &= \arg\min_u \int_\Omega -\log P(f(x)|u(x))dx - \log P(u). \end{aligned} \tag{2.3}$$

The first term in (2.3) indicates the occurrence probability of intensity $f(x)$ of the noisy image given intensity $u(x)$ of the latent image at point $x \in \Omega$. Hence, because the noise follows Gaussian distribution, we have

$$\log P(f(x)|u(x)) = -\frac{1}{2}\log(2\pi) - \log\sigma - \frac{(u(x) - f(x))^2}{2\sigma^2},$$

Furthermore, *a priori* knowledge $P(u)$ on the distribution of latent image u is required and it is represented by a regularization term. After omitting the constant terms, the resulting variational model is

$$\arg\min_u -\log P(u) - \int_\Omega \frac{(u(x) - y(x))^2}{2\sigma^2}dx.$$

We can identify that the data-fidelity term of a variational model for Gaussian noise is given in the form of L_2-norm. In general, a variational model for denoising Gaussian noise is given as follows:

$$\arg\min_u \mathcal{R}(u) + \frac{\lambda}{2}\int_\Omega |u-f|^2\,dx, \quad \lambda > 0, \tag{2.4}$$

where f is an observed noisy image, $\mathcal{R}(u)$ is a regularization term and λ is a parameter balancing two terms.

2.1.2 Regularization

Regularization term of a variational model possesses information on the characteristics of a solution that is acquired from the model and it is important that we introduce a regularization which is appropriate for a given problem and reflects our intention.

Total variation (TV)

We recall that total variation (TV) [54] has been in great use since its appearance. A variational model adopting TV is given as follows:

$$\arg\min_u |u|_{\mathrm{BV}(\Omega)} + \frac{\lambda}{2}\int_\Omega |u-f|^2\,dx, \quad \lambda > 0. \tag{2.5}$$

Here, $\mathrm{BV}(\Omega)$ is a space of bounded variation functions and is a subspace of $L^1(\Omega)$. Note that

$$|u|_{\mathrm{BV}(\Omega)} := \sup\left\{-\int_\Omega u\,\mathrm{div}\phi\,dx \;\mid\; \phi \in C_c^\infty(\Omega,\mathbb{R}^C),\ \|\phi(x)\|_\infty \le 1\ \forall x \in \Omega\right\},$$

$$\mathrm{BV}(\Omega) := \left\{u \in L^1(\Omega) \,:\, |u|_{\mathrm{BV}(\Omega)} < +\infty\right\},$$

and when u is smooth, i.e., $u \in C^1(\Omega)$ (or $W^{1,1}(\Omega)$), $\mathrm{TV}(u)$ becomes

$$|u|_{\mathrm{BV}(\Omega)} = \int_\Omega |\nabla u| \, dx.$$

It is well known that TV is convex and lower semi-continuous. Subsequently, the existence and uniqueness of the solution of the problem (2.5) are obtained.

Theorem 2.1.1. ([15]) Let f be in $L^2(\Omega)$. Then the problem (2.5) has a solution in $BV(\Omega)$ and it is unique.

Proof. Take a minimizing sequence $\{u_n\}_{n\in\mathbb{N}}$ for the energy functional in (2.5). Because $\{u_n\}_{n\in\mathbb{N}}$ is bounded in $L^2(\Omega)$, there exists a subsequence $\{u_{n_k}\}_{n_k\in\mathbb{N}}$ converging to some u in $L^2(\Omega)$. By lower semi-continuity of TV and Fatou's lemma, $u \in BV(\Omega)$ and it minimizes the energy functional. The uniqueness follows from the convexity of TV. □

For denoising task, TV performs very well in that it makes homogeneous region clear and preserves edge like boundary of object or textures. However, the solution becomes approximately piecewise constant function and stain-like feature called staircase artifact appear which is caused by discontinuities between constant regions.

Combined first and second order approach

A combined first and second order approach [49] is also one of the straightforward higher-order extension of TV. It combines total variation of an image and total variation of gradient of an image. A variational model adopting combined approach is given as follows:

$$\arg\min_u \alpha \int_\Omega |\nabla u| \, dx + \beta \int_\Omega |\nabla^2 u| \, dx + \frac{1}{2} \int_\Omega |u - f|^2 \, dx, \quad \alpha, \beta > 0. \tag{2.6}$$

It is known that the model (2.6) has the unique minimizer. The additional second order term acts as an assistant to remove artifacts caused by first order term without creating any serious blur.

Total generalized variation (TGV)

Another regularization method is total generalized variation (TGV) [8] which is a higher-order extension of TV. It is of the form:

$$\mathrm{TGV}_\alpha^k(u) = \sup\Big\{ \int_\Omega u \mathrm{div}^k \phi \, dx \;\mid\; \phi \in C_c^k(\Omega, \mathrm{Sym}^k(\mathbb{R}^2)), \\ \|\mathrm{div}^\ell \phi\|_\infty \le \alpha_\ell, , \ell = 0, 1, \cdots, k-1 \Big\},$$

where $\mathrm{Sym}^k(\mathbb{R}^2))$ denotes the space of symmetric tensors of order k with arguments in $\mathbb{R}^2$, and α_ℓ are fixed positive parameters. The second order TGV, i.e. TGV_α^2, is alternatively expressed as [9]:

$$\mathrm{TGV}_\alpha^2(u) = \min_{w \in \mathrm{BD}(\Omega)} \alpha_1 \int_\Omega |\nabla u - w| \, dx + \alpha_0 \int_\Omega |\mathcal{E} w| \, dx \tag{2.7}$$

where $\mathrm{BD}(\Omega)$ denotes the space of vector fields of bounded deformation, i.e., $w \in L^1(\Omega, \mathbb{R}^2)$ such that the distributional symmetrized derivative $\mathcal{E}w = \frac{1}{2}(\nabla w + \nabla w^T)$ is a 2×2 symmetric matrix-valued Radon measure.

The use of second order TGV as a regularization to denoising shows a smoothed result without staircase artifact and the solution leads to a piecewise affine function. Intuitively, it is explained as follows. In smooth regions of an image u, second derivative of u is locally small. It leads that w in (2.7) is chosen locally as ∇u and TGV_α^2 penalizes second derivative of u. Around edges, on the other hand, second derivative of u is relatively larger than ∇u. It makes w close to 0 locally and TGV_α^2 penalizes total variation of u. In this way, TGV becomes locally affine in smooth regions while preserving edges of an image.

2.1.3 Optimization algorithm

It needs that we find the numerical solution minimizing the given energy functional. There are various optimization algorithms for obtaining the solution. Here, we introduce the algorithms which is later related to solving the proposed models.

Alternating direction method of multipliers (ADMM)

The alternating direction method of multipliers (ADMM) transforms contrained problem into unconstrained form and minimize each variable independently. Suppose we are given the problem:

$$\begin{aligned} &\min_{x,y} f(x) + g(y), \\ &\text{subject to } Ax + By = c, \end{aligned}$$

where f, g satisfy some conditions.

We define the augmented lagrangian $\mathcal{L}_\beta(x, y, w)$ as follows:

$$\mathcal{L}_\beta(x, y, w) = f(x) + g(y) + \langle w, Ax + By - c\rangle + \frac{\beta}{2}\|Ax + By - c\|_2^2,$$

where $\beta > 0$ and w is auxiliary variable.

Then, the ADMM minimizes each variable separately and update auxiliary variable:

$$\begin{aligned} x^{k+1} &= \arg\min_x \mathcal{L}_\beta(x, y^k, w^k), \\ y^{k+1} &= \arg\min_y \mathcal{L}_\beta(x^{k+1}, y, w^k), \\ w^{k+1} &= w^k + \beta(Ax^{k+1} + By^{k+1} - c). \end{aligned}$$

It is known [62] that the sequence $\{(x^k, y^k, w^k)\}$ converges if f, g are convex

and possibly for some nonconvex, nonsmooth cases.

Nesterov's accelerated gradient

Suppose we have a smooth convex function f on $\mathbb{R}^n$ which is β-Lipschitz gradient continuous, i.e., for any $x, y \in \mathbb{R}^n$,

$$\|\nabla f(x) - \nabla f(y)\| \leq \beta \|x - y\|.$$

Then, it is well known that the gradient descent algorithm $x_{n+1} = x_n - \eta \nabla f(x_n)$ converges to minimum of f with convergence rate of order $\mathcal{O}(1/n)$ if $\eta \leq \frac{1}{\beta}$. Subsequently, Nesterov [46] introduce the accelerated gradient descent algorithm which attains the convergence rate of order $\mathcal{O}(1/n^2)$.

If we define $t_1 = 0$, $y_1 = x_1$ for arbitrary initial point x_1, the algorithm is simply given as follows: for $n \geq 1$,

$$\begin{aligned} y_{n+1} &= x_n - \eta \nabla f(x_n), \\ t_{n+1} &= \frac{1 + \sqrt{1 + 4t_n^2}}{2}, \\ x_{n+1} &= y_{n+1} + \frac{t_n - 1}{t_{n+1}}(y_{n+1} - y_n). \end{aligned}$$

2.2 Methods for image dehazing in the air

In the presence of haze in the atmosphere, an image we take is represented by sum of attenuated direct signal from the scene point and global atmospheric light from the scattering of background. Then, the model for the formation of hazy image is given as follows:

$$\mathbf{I}(x) = \mathbf{J}(x)t(x) + \mathbf{A}(1 - t(x)), \tag{2.8}$$

where **I** is an observed hazy image, **J** is the scene radiance, **A** is the global atmospheric light and t is the transmission of scene point describing the portion of light that is not attenuated and reaches the observer. The first term in the right hand side of (2.8) account for the direct signal attenuated by ratio of t and the second term explains airlight by the scattering of light due to medium.

We assume that the atmosphere is homogeneous so that the transmission t is expressed as

$$t(x) = \exp(-\beta d(x)), \tag{2.9}$$

where β is a scattering coefficient of the atmosphere and $d(x)$ is a distance from the observer to the scene point x. In pure air, the amount of absorption and scattering of light is a function of wavelength. But, when air contains solid and liquid particles floating around the neighborhood of surface of the earth, multiple scattering effects become significant and effect of wavelength on degradation process diminishes [3].

Usually, the given image possess three color channels. To restore scene radiance **J** given hazy image **I** in (2.8), we need to know the transmission t and global atmospheric light **A** which is more than the number of given equations. It requires that we introduce some regularization to make up the shortage of given information.

2.2.1 Dark channel prior

In [30], the authors observe that most of local patches in clean image contain pixels with very low intensities at least one color channel. Under the observation, they assume that such low intensity is zero.

We can verify the validity of the above assumption if we think on the contrary. Let us assume that for a local patch of haze-free image, all pixels in the patch has high intensities at all color channels. Then, the patch would be very bright and close to white color. Such patch rarely exists in

clean image except the sky region or large object with color close to white.

The dark channel of an image is defined as each local patch of the dark channel has the intensity which is smallest among all pixels and color channels of the corresponding patch of the image. An example of the dark channel of clean outdoor image is presented in Figure 2.1.

(a) **(b)**

Figure 2.1: Illustration of validity of the dark channel prior. (a) Clean outdoor image, (b) The result of identifying dark channel value at each patch of the image.

By applying the dark channel prior, we can restore the hazy image. At first, assume that global atmospheric light **A** is known. For a local patch $\Omega(x)$ around x, take minimum over pixels in the patch on (2.8). Then, we have:

$$\min_{y\in\Omega(x)}(I^c(y)) = \tilde{t}(x)\min_{y\in\Omega(x)}(J^c(x)) + (1-\tilde{t}(x))A^c,$$

where c indicates each color channel and we assume that transmission in the local patch $\Omega(x)$ is the same as $\tilde{t}(x)$. Dividing both sides by A^c and taking minimum over color channels. Then, it leads to:

$$\min_c(\min_{y\in\Omega(x)}(\frac{I^c(y)}{A^c})) = \tilde{t}(x)\min_c(\min_{y\in\Omega(x)}(\frac{J^c(y)}{A^c})) + (1-\tilde{t}(x)). \tag{2.10}$$

By the dark channel prior, the first term on the right-hand side of (2.10) becomes zero because A^c is always positive. Then, we earn the transmission map as follows:

$$\tilde{t}(x) = 1 - \min_c(\min_{y\in\Omega(x)}(\frac{I^c(y)}{A^c})).$$

Global atmospheric light $\mathbf{A}$ is also obtained by the use of dark channel. Therefore, from t and $\mathbf{A}$, we can recover scene radiance $\mathbf{J}$ as:

$$\mathbf{J}(x) = \frac{\mathbf{I}(x) - \mathbf{A}}{\max(t(x), t_0)} + \mathbf{A},$$

where the constant t_0 is adopted to prevent the transmission map from being close to zero.

In Figure 2.2, we present the exemplar dehazing result using the dark channel prior. We can identify that visibility of the image is improved, but there are block artifacts in the background due to the use of patch-based method.

(a) **(b)**

Figure 2.2: Example of dehazing using the dark channel prior. (a) Hazy outdoor image, (b) The result of dehazing using the dark channel prior.

2.2.2 Color ellipsoid prior

Although the dark channel prior performs fairly well for dehazing task, there are still some drawbacks. These include blocky artifacts in the background, invalidity to large white region and vulnerability to noise in a local patch.

The color ellipsoid prior, introduced in [11], tries to make improvement so that the dark channel method is more robust to irregular pixels in a local patch. If pixels in a local patch are scattered in RGB space, the cluster region is approximated by ellipsoid while excluding statistically deviated pixels from the most pixels in the patch. Regarding the constructed ellipsoid as valid data of the patch, the dark channel prior is applied. Under this process called the color ellipsoid prior method, random and irregular pixels in the patch can be statistically ruled out. In Figure 2.3. we present an illustration of constructing the color ellipsoid for data of a local patch in RGB space.

Let us normalize pixels in a local patch Ω by global atmospheric light $\mathbf{A} = [A^r, A^g, A^b]^T$ at each channel:

$$\bar{I} = \left[\frac{I^r}{A^r}, \frac{I^g}{A^g}, \frac{I^b}{A^b} \right]^T,$$

and construct the ellipsoid E approximating the normalized pixels $\bar{I}$:

$$E = \{z \mid (z-\mu)^T \Sigma^{-1} (z-\mu) \leq 1\},$$

where μ is a mean vector of normalized pixels $\bar{I}$ and Σ is a covariance matrix

$$\Sigma = \frac{1}{|\Omega|} \sum_{\bar{I} \in \Omega} (\bar{I} - \mu)(\bar{I} - \mu)^T = \begin{pmatrix} \sigma_r^2 & \sigma_{rg} & \sigma_{rb} \\ \sigma_{gr} & \sigma_g^2 & \sigma_{gb} \\ \sigma_{br} & \sigma_{bg} & \sigma_b^2 \end{pmatrix}.$$

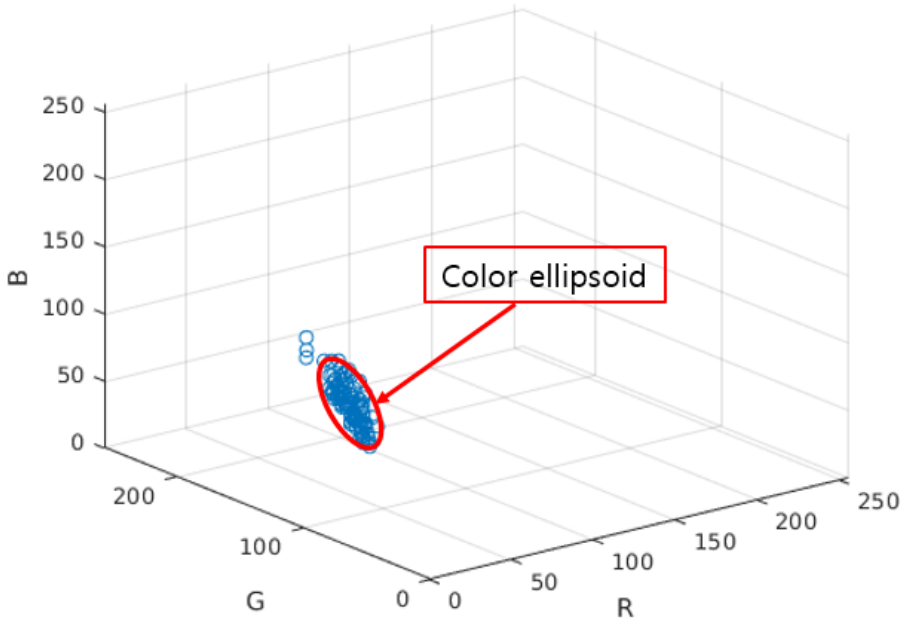


Figure 2.3: Illustration of distribution of image patch in RGB space and color ellipsoid approximating the cluster of pixels.

Subsequently, applying the dark channel prior to the ellipsoid E requires to find minimum among distances from surface of the ellipsoid to RG, RB, BG planes. The minimum occurs when normal vector of point on the ellipsoid is parallel with one of the normal vectors of RG, RB, BG planes. Then, it leads that the transmission $\tilde{t}$ can be obtained as follows:

$$\tilde{t}(x) = 1 - \min_{c}(\mu_c - \sigma_c),$$

where c indicates color channels of the image.

Remark 2.2.1. There is a possibility that if we scatter pixels of a local

patch in RGB space, there exist two distinct clusters which are not random noise but indication of different objects. Then, it is impossible to approximate the patch by a single ellipsoid. It causes the occurrence of halo artifacts and requires additional post-processing.

Chapter 3

Image denoising for Gaussian and Cauchy noise

3.1 Second-order structure tensor and hybrid STV

Structure tensor total variation (STV) [39] regularization is a generalization of the total variation (TV) or ROF model [54]. STV is different from TV in that at each point of image, STV adaptively takes the directions of highest and smallest intensity variation of neighborhood of a point while TV always considers vertical and horizontal directions. Thus, STV shows better ability to detect edge and smooth region than TV. Nevertheless, STV still produces staircase artifacts because it is a TV based regularization.

In this chapter, to remove staircase artifacts, we introduce a second-order structure tensor total variation (STV2), which is a higher order extension of structure tensor total variation. For STV2, we use the eigenvalues of weighted averaged matrix whose column is a vectorial form of Hessian of each channel of image. However, the valuable edges of an image can be attenuated when using only second order STV. Therefore, we suggest a hybrid STV which is a

convex combination of STV and STV2. It improves image restoration tasks such as denoising and deblurring by sufficiently smoothing homogeneous regions while preserving edge parts.

3.1.1 Structure tensor total variation

Structure tensor

Let us denote an image domain by $\Omega \subset \mathbb{R}^2$ and assume $\mathbf{u} = [u_1, \cdots, u_M] : \Omega \to \mathbb{R}^M$ to be an image with M channels. For example, if we have a color image with RGB channels, M equals 3.

At a point $x \in \Omega$, TV or vectorial TV (VTV) [7] considers a gradient vector $\nabla u_i(x)$ $(1 \leq i \leq M)$ of an image $\mathbf{u}$. It can be acquired by concentrating on the targeted point, thus limiting the amount of information we can get. Accordingly, a new concept called structure tensor is introduced to take into account information from neighborhood of the point.

Let $\mathbf{n}$ be a unit normal vector at a point $x \in \Omega$. Then intensity variation of each channel of image $\mathbf{u}$ in the direction $\mathbf{n}$ is given as follows:

$$\frac{\partial \mathbf{u}}{\partial \mathbf{n}} = (\mathbf{J}\mathbf{u}(x))\mathbf{n},$$

where $\mathbf{J}\mathbf{u}$ is the Jacobian matrix of $\mathbf{u}$ and it is defined as

$$\mathbf{J}\mathbf{u}(x) = [\nabla u_1(x), \cdots, \nabla u_M(x)]^T.$$

Then, the magnitude of intensity variation at x is given by

$$\left\| \frac{\partial \mathbf{u}}{\partial \mathbf{n}} \right\|_2 = \sqrt{\mathbf{n}^T (\mathbf{J}\mathbf{u}(x))^T (\mathbf{J}\mathbf{u}(x)) \mathbf{n}}.$$

Because $\mathbf{n}$ is an arbitrary directional vector, $(\mathbf{J}\mathbf{u}(x))^T(\mathbf{J}\mathbf{u}(x))$ determines the magnitude of intensity variation at x. We convolve the above quantity

with the nonnegative, rotationally symmetric kernel K so that information of neighborhood of a point is additionally considered [39].

Definition 3.1.1. Let the image $\mathbf{u}$ is given. *The structure tensor* $S_K\mathbf{u}(x)$ of $\mathbf{u}$ at point x is

$$S_K\mathbf{u}(x) = K * [\mathbf{J}\mathbf{u}^T\mathbf{J}\mathbf{u}](x),$$

where $\mathbf{J}$ is the Jacobian operator and K is a nonnegative, rotationally symmetric kernel.

Note that structure tensor $S_K\mathbf{u}(x)$ is a 2×2 symmetric, positive semi-definite matrix and it has two nonnegative eigenvalues $\lambda^+(x)$, $\lambda^-(x)$ ($\lambda^+(x) \geq \lambda^-(x)$). The root mean square of weighted magnitude of intensity variation at x is represented by eigenvalues of structure tensor:

Let $w \in (-\pi, \pi]$ is an angle between the directional vector $\mathbf{n}$ and unit-eigenvector θ^+ corresponding to λ^+, then we have

$$\sqrt{K * \left\| \frac{\partial \mathbf{u}}{\partial \mathbf{n}} \right\|_2^2} = \sqrt{\mathbf{n}^T(S_K\mathbf{u})\mathbf{n}} = \sqrt{\lambda^+ \cos^2 w + \lambda^- \sin^2 w}. \tag{3.1}$$

It informs that $\sqrt{\lambda^+}$ and $\sqrt{\lambda^-}$ represents maximum and minimum weighted magnitude of intensity variation with respect to direction, respectively.

If we consider an ellipse $P(w)$ whose axis are θ^+ and θ^- with radius $\sqrt{\lambda^+}$ and $\sqrt{\lambda^-}$ respectively,

$$P(w) = (\sqrt{\lambda^+}\cos w)\theta^+ + (\sqrt{\lambda^-}\sin w)\theta^-, \;\; w \in (-\pi, \pi],$$

then (3.1) can be also interpreted as a distance from a point on ellipse to the center of an ellipse where a line connecting center to the point makes an angle w with θ^+. We can visualize the structure tensor using the ellipse $P(w)$. When both eigenvalues are small at a point x, we can see that there is relatively small variations in a neighborhood of x. If λ^+ is large and λ^- is small relatively, then there is a wide difference between the magnitude of

variation with regard to the direction from x. Two large eigenvalues indicate that there are high variations in all directions. An illustration of visualized structure tensors at different points is presented in Figure 3.1.

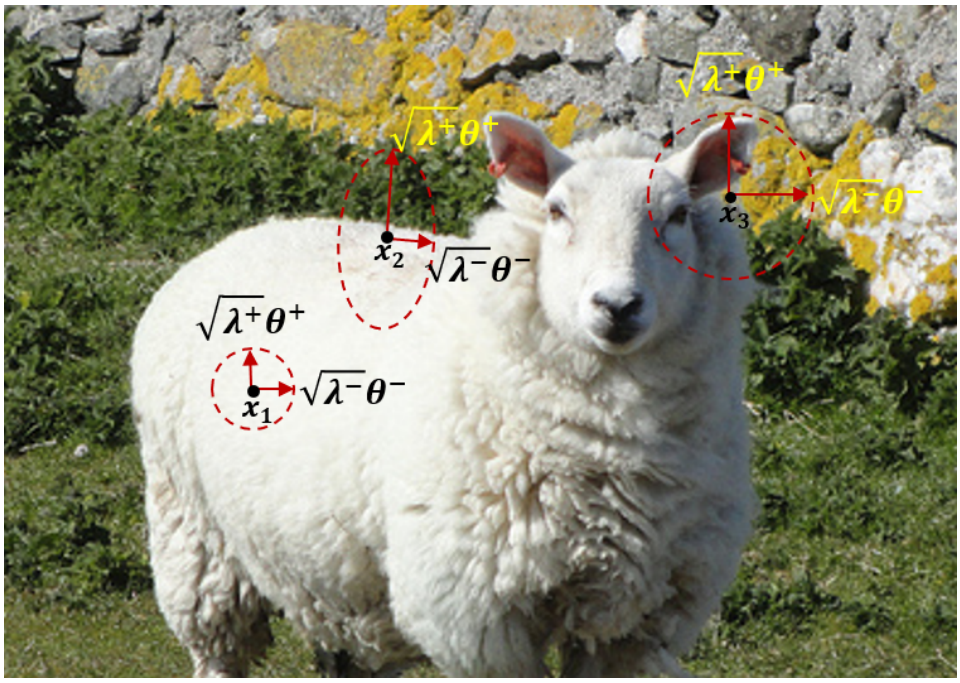


Figure 3.1: Visualized structure tensors at different points using $P(w)$. The shape of an ellipse varies according to the position of point and it indicates the magnitude of variation around point.

Structure tensor as a regularization

Now, we define a regularizer called STV using the eigenvalues obtained from the structure tensor $S_K\mathbf{u}$ [39]:

Definition 3.1.2. For an image $\mathbf{u} \in W^{1,2}(\Omega, \mathbb{R}^M)$ and $p \geq 1$, *the structure*

tensor total variation (STV) is defined as follows:

$$\mathrm{STV}_p(\mathbf{u}) = \int_\Omega \left\| (\sqrt{\lambda^+}, \sqrt{\lambda^-}) \right\|_p dx,$$

where λ^+, λ^- are eigenvalues of structure tensor $S_K\mathbf{u}$ at each point $x \in \Omega$.

STV acts as a regularization by constraining maximum and minimum intensity variation of a neighborhood of each point. It reflects the variation of a wider area around a point than TV does.

Remark 3.1.3. Suppose that we use the delta function $\delta(x)$ as a kernel $K(x)$ when constructing the structure tensor $S_K\mathbf{u}$. If an image u is a grayscale ($M = 1$), it becomes that the eigenvalues λ^+, λ^- of structure tensor are $\|\nabla u\|_2^2$, 0 respectively. Then, in this case, $\mathrm{STV}_p(u) = \mathrm{TV}(u)$ for $p \geq 1$. Similarly, when $M > 1$ and $p = 2$, we have $\mathrm{STV}_2(\mathbf{u}) = \mathrm{VTV}(\mathbf{u})$. It suggests that STV is an generalization of TV in that STV takes more information than TV by considering a neighborhood of a targeted point.

Remark 3.1.4. Throughout this section, we fix $p = 1$ and denote $\mathrm{STV}_1(\mathbf{u}) = \mathrm{STV}(\mathbf{u})$.

Using STV as a regularization, we can restore an noisy image which is corrupted by Gaussian noise. In Figure 3.2, we present the denoising results when we use TV and STV as a regularization, respectively, i.e., insert TV or STV to $\mathcal{R}(u)$ in (2.4).

In Figure 3.2, we present denoising results for noisy image corrupted by Gaussian noise using TV and STV as a regularization, respectively. By comparing the smoothness and remaining noisy features of the two results, we can identify that the image restored by STV shows better restoration quality than that by TV. It is also supported by quantitative results such as PSNR and SSIM which will be explained later. But we notice that there still exist staircase artifacts and stain-like features in both images.

(a) **(b)** **(c)**

Figure 3.2: Comparison of denoising results by TV and STV. (a) Noisy image corrupted by Gaussian noise $\sigma = 0.1$, (b) Restored image using TV as a regularization and (c) Restored image using STV as a regularization.

3.1.2 Proposed model

In the previous section, we verify that the solution of (2.4) using STV as a regularizer shows similar geometry to that of the ROF model - it is close to a piecewise constant function and shows staircase artifacts. Now, we propose a new regularizer which is a higher-order extension of STV and expect that use of the new regularizer would mitigate an occurrence of staircase artifacts. We utilize second-order derivative of an image to define the regularizer analogous to STV.

For an image $\mathbf{u} : \Omega \subset \mathbb{R}^2 \to \mathbb{R}^M,\ \mathbf{u} = (u_1, \cdots, u_M)$, we consider a Hessian operator $\mathbf{H}$ as follows:

$$\mathbf{Hu}(x) = [Hu_1(x), \cdots, Hu_M(x)]^T,$$

where $Hu_i(x) = \Big[\dfrac{\partial^2 u_i}{\partial x^2}, \dfrac{\partial^2 u_i}{\partial x \partial y}, \dfrac{\partial^2 u_i}{\partial y \partial x}, \dfrac{\partial^2 u_i}{\partial y^2}\Big]^T$ for $1 \le i \le M$.

Definition 3.1.5. We define *the second-order structure tensor* $S_K^{(2)}\mathbf{u}$ at point

x as

$$S_K^{(2)}\mathbf{u}(x) = K * (\mathbf{H}\mathbf{u}^T\mathbf{H}\mathbf{u})(x)$$

where K is a nonnegative, rotationally symmetric kernel.

Because the second-order structure tensor is a 4×4 symmetric positive semi-definite matrix, it has four nonnegative eigenvalues. Note that hessian vector of an image can be viewed as a gradient of gradient of an image. If we have a vector $\mathbf{n} = [\mathbf{n}_1^T, \mathbf{n}_2^T]^T \in \mathbb{R}^4$ which consists of unit direction vectors $\mathbf{n}_1, \mathbf{n}_2 \in \mathbb{R}^2$, the variation of gradient of each channel of image $\mathbf{u}$ with respect to $\mathbf{n}$ at point x can be expressed as follows:

$$\frac{\partial(\mathbf{J}\mathbf{u})}{\partial\mathbf{n}} := (\mathbf{H}\mathbf{u}(x))\mathbf{n}$$

where $\mathbf{J}$ is the Jacobian operator. If we are to include information of a neighborhood of a point, the magnitude of variation of gradient of image in local region is

$$\begin{aligned}\sqrt{K * \left\|\frac{\partial(\mathbf{J}\mathbf{u})}{\partial\mathbf{n}}\right\|_2^2} &= \sqrt{\mathbf{n}^T(S_K^{(2)}\mathbf{u})\mathbf{n}} \\ &= \sqrt{\lambda_1\langle\mathbf{n},\theta_1\rangle^2 + \lambda_2\langle\mathbf{n},\theta_2\rangle^2 + \lambda_3\langle\mathbf{n},\theta_3\rangle^2 + \lambda_4\langle\mathbf{n},\theta_4\rangle^2},\end{aligned}$$

where $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ and $\theta_1, \theta_2, \theta_3, \theta_4$ are eigenvalues and eigenvectors of $S_K^{(2)}\mathbf{u}$ respectively. It indicates that the eigenvalues of second order structure tensor are related to the magnitude of variation of gradient of an image in the local area.

Analogous to STV, we define the second order structure tensor TV (STV2) using the eigenvalues of the second order structure tensor.

Definition 3.1.6. For an image $\mathbf{u} \in W^{2,2}(\Omega, \mathbb{R}^M)$ and $q \geq 1$, *the second*

order structure tensor total variation (STV2) is defined as follows:

$$\mathrm{STV2}_q(\mathbf{u}) = \int_\Omega \left\| (\sqrt{\lambda_1}, \sqrt{\lambda_2}, \sqrt{\lambda_3}, \sqrt{\lambda_4}) \right\|_q dx,$$

where $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ are eigenvalues of structure tensor $S_K^{(2)}\mathbf{u}$ at each point $x \in \Omega$.

Remark 3.1.7. If we take the kernel $K(x)$ to be the delta function $\delta(x)$ for constructing the second order structure tensor $S_K^{(2)}(\mathbf{u})$, then its eigenvalues are $\|Hu\|_2^2, 0, 0, 0$ when u is a grayscale image. Then, $\mathrm{STV2}_q(u)$ corresponds to $\|\nabla^2 u\|_1$ for any $q \geq 1$.

Remark 3.1.8. As STV, we fix $q = 1$ for STV2 throughout this section and denote $\mathrm{STV2}_q(u) = \mathrm{STV2}(u)$.

Note that STV2 contains second-order derivative information of an image. When STV2 is used as a regularizer in a variational model for denoising, we expect that the solution avoids staircasing artifacts unlike STV which favors a piecewise-constant solution. Therefore using both STV and STV2 as regularizations simultaneously gives us benefits of each regularizer. We expect that while the solution preserves edges well, it does not produce staircasing artifacts. By combining both regularizations, we consider the hybrid STV minimization problem :

$$\arg\min_{\mathbf{u}} \tau_1 \mathrm{STV}(\mathbf{u}) + \tau_2 \mathrm{STV2}(\mathbf{u}) + \frac{1}{2}\|\mathbf{u} - \mathbf{f}\|_2^2, \;\; \tau_1, \tau_2 > 0, \tag{3.2}$$

where $\mathbf{f}$ is a given noisy image.

Now, we provide the convex properties of each functional, STV and STV2. Its proof is given in Appendix A.1.

Theorem 3.1.9. For $\mathbf{u} \in W^{2,2}(\Omega, \mathbb{R}^M)$, $\mathrm{STV}(\mathbf{u})$ and $\mathrm{STV2}(\mathbf{u})$ are convex.

3.1.3 Discretization of the model

In previous sections, we assume the model is on a continuous domain. In real photography, an image sensor in camera receives light so that each cell in the sensor determines intensity of an image. Therefore, to obtain practical results of our proposed model, it requires us to deal with the model on a discrete domain.

Let an image $\mathbf{u} = [u_1, \cdots, u_M]$ be defined on the discrete rectangular domain where each channel u_i, $(1 \leq i \leq M)$, is vectorized and has N pixels. For n-th pixel x_n in the domain, we compute $S_K\mathbf{u}(x_n)$, $S_K^{(2)}\mathbf{u}(x_n)$ and denote it by $[S_K\mathbf{u}]_n, [S_K^{(2)}\mathbf{u}]_n$ respectively. In the computation, we consider the kernel K to be also defined on the discrete domain and restrict it to have a compact support $\mathcal{P} = \{-L_K, \cdots, L_K\}^2$ where L_K is a positive integer.

Then, we define the discrete versions of STV and STV2 as follows:

$$\begin{aligned}
\mathrm{STV}(\mathbf{u}) &= \sum_{n=1}^{N} \left\| \left(\sqrt{\lambda_n^+}, \sqrt{\lambda_n^-} \right) \right\|_1, \\
\mathrm{STV2}(\mathbf{u}) &= \sum_{n=1}^{N} \left\| \left(\sqrt{\lambda_{1,n}}, \sqrt{\lambda_{2,n}}, \sqrt{\lambda_{3,n}}, \sqrt{\lambda_{4,n}} \right) \right\|_1,
\end{aligned}$$

where λ_n^+, λ_n^-, $(\lambda_n^+ \geq \lambda_n^-)$, are non-negative eigenvalues of $[S_K\mathbf{u}]_n$ and $\lambda_{1,n}, \lambda_{2,n}$, $\lambda_{3,n}, \lambda_{4,n}$ are non-negative eigenvalues of $[S_K^{(2)}\mathbf{u}]_n$ with decreasing order.

The optimization of discrete version of the energy functional (3.2) using the above regularizers is tough and burdensome. We seek to manageable representation of both regularizers. For simplicity, consider the case $M = 1$, i.e., a grayscale image. If we calculate $[S_K u]_n$, we have

$$\begin{aligned}
[S_K u]_n =& K * [\nabla u \nabla u^T](x_n) \\
=& \sum_{s \in \mathcal{P}} K(s) \cdot \begin{bmatrix} (\nabla_x u(x_n - s))^2, & \nabla_x u(x_n - s) \cdot \nabla_y u(x_n - s) \\ \nabla_x u(x_n - s) \cdot \nabla_y u(x_n - s), & (\nabla_y u(x_n - s))^2 \end{bmatrix},
\end{aligned}$$

where s ranges over compact support $\mathcal{P}$ of kernel K. It is noteworthy that the summation can be decomposed into multiplication of elaborately defined matrices. The patch-based Jacobian is introduced in [39] as a linear mapping $\mathbf{J}_K : \mathbb{R}^N \to \mathcal{X}$,

$$[\mathbf{J}_K u]_n = \begin{bmatrix} \sqrt{K(s_1)}\nabla_x u(x_n - s_1), & \sqrt{K(s_1)}\nabla_y u(x_n - s_1) \\ \vdots & \vdots \\ \sqrt{K(s_L)}\nabla_x u(x_n - s_L), & \sqrt{K(s_L)}\nabla_y u(x_n - s_L) \end{bmatrix},$$

where $\mathcal{X} = \mathbb{R}^{N\times(L\times 2)}$, $L = (2L_K + 1)^2$, $s_1, \cdots, s_L \in \mathcal{P}$ and $[\mathbf{J}_K u]_n$ is $L \times 2$ matrix at each pixel x_n. Then, we can observe by direct calculation that

$$[S_K u]_n = [\mathbf{J}_K u]_n^T [\mathbf{J}_K u]_n. \tag{3.3}$$

The above argument can be extended to the case $M > 1$. In this case, the patch-based Jacobian has the following form:

$$[\mathbf{J}_K \mathbf{u}]_n = \begin{bmatrix} \sqrt{K(s_1)}\nabla_x u_1(x_n - s_1), & \sqrt{K(s_1)}\nabla_y u_1(x_n - s_1) \\ \vdots & \vdots \\ \sqrt{K(s_L)}\nabla_x u_1(x_n - s_L), & \sqrt{K(s_L)}\nabla_y u_1(x_n - s_L) \\ \vdots & \vdots \\ \sqrt{K(s_1)}\nabla_x u_M(x_n - s_1), & \sqrt{K(s_1)}\nabla_y u_M(x_n - s_1) \\ \vdots & \vdots \\ \sqrt{K(s_L)}\nabla_x u_M(x_n - s_L), & \sqrt{K(s_L)}\nabla_y u_M(x_n - s_L) \end{bmatrix}, \tag{3.4}$$

where the patch-based Jacobians of each channel are concatenated vertically and we can identify that (3.3) still holds for multi-channel image $\mathbf{u}$. Analogously, we introduce the patch-based Hessian operator $\mathbf{PH}_K : \mathbb{R}^{NM} \to \mathcal{Y}$,

where $\mathcal{Y} = \mathbb{R}^{N\times(LM\times 4)}$.

$$[\mathbf{PH}_K\mathbf{u}]_n = \begin{bmatrix} \tilde{\nabla}^2 u_1(x_n - s_1) \\ \vdots \\ \tilde{\nabla}^2 u_1(x_n - s_L) \\ \vdots \\ \tilde{\nabla}^2 u_M(x_n - s_1) \\ \vdots \\ \tilde{\nabla}^2 u_M(x_n - s_L) \end{bmatrix}, \tag{3.5}$$

where $\tilde{\nabla}^2 u_i(x_n - s_j)$ equals

$$\sqrt{K(s_j)}\left[\Delta_{xx}(x_n - s_j),\ \ \Delta_{xy}(x_n - s_j),\ \ \Delta_{yx}(x_n - s_j),\ \ \Delta_{yy}(x_n - s_j)\right],$$

$L = (2L_K + 1)^2$, $s_1, \cdots, s_L \in \mathcal{P}$, $\Delta_{z_1 z_2}$ is the discrete second order partial derivative with respect to z_1, z_2 and $[\mathbf{PH}_K\mathbf{u}]_n$ is $LM \times 2$ matrix at each pixel x_n. Then, the second-order structure tensor is decomposed into the multiplication using the patch-based Hessian.

Proposition 3.1.10. The discrete STV and STV2 of $\mathbf{u}$ at pixel x_n are decomposed into multiplications utilizing patch-based Jacobian (3.4) and path-based Hessian (3.5) as follows:

$$[S_K\mathbf{u}]_n = [\mathbf{J}_K\mathbf{u}]_n^T[\mathbf{J}_K\mathbf{u}]_n,$$
$$[S_K^{(2)}\mathbf{u}]_n = [\mathbf{PH}_K\mathbf{u}]_n^T[\mathbf{PH}_K\mathbf{u}]_n.$$

The Proposition 3.1.10 enables that finding square roots of eigenvalues of STV and STV2 boils down to calculating singular values of patch-based Jacobian and patch-based Hessian. Moreover, because patch-based operators are linear, it is appropriate to apply mathematical theory for optimization process. Now, let us define the norm which represents the p-norm of singular

vector of a matrix:

Definition 3.1.11. For a matrix $X \in \mathbb{C}^{N_1 \times N_2}$, let $\sigma_1, \cdots, \sigma_{\min(N_1,N_2)}$ be the singular values of X. Then *the Schatten norm of order p of X* is defined as

$$\|X\|_{\mathcal{S}_p} = \left(\sum_{n=1}^{\min(N_1,N_2)} \sigma_n^p \right)^{\frac{1}{p}}.$$

Then we can represent the STV, STV2 using Schatten 1-norm:

$$\begin{aligned} \mathrm{STV}(\mathbf{u}) &= \sum_{n=1}^{N} \|[\mathbf{J}_K \mathbf{u}]_n\|_{\mathcal{S}_1} = \|\mathbf{J}_K \mathbf{u}\|_{1,1}, \\ \mathrm{STV2}(\mathbf{u}) &= \sum_{n=1}^{N} \|[\mathbf{P}\mathbf{H}_K \mathbf{u}]_n\|_{\mathcal{S}_1} = \|\mathbf{P}\mathbf{H}_K \mathbf{u}\|_{1,1}, \end{aligned}$$

where $\|\cdot\|_{1,1}$ denotes the mixed ℓ_1-S_1 norm.

Subsequently, let us equip inner product $\langle \cdot, \cdot \rangle_{\mathcal{X}}$ on space $\mathcal{X} = \mathbb{R}^{N \times (LM \times 2)}$ which are codomain of patch-based Jacobian operator: for $X_1, X_2 \in \mathcal{X}$,

$$\langle X_1, X_2 \rangle_{\mathcal{X}} = \sum_{i=1}^{N} \mathrm{trace}(X_{2,n}^T X_{1,n}),$$

where $X_{1,n}, X_{2,n} \in \mathbb{R}^{LM \times 2}$ and inner product $\langle \cdot, \cdot \rangle_{\mathcal{Y}}$ on space $\mathcal{Y} = \mathbb{R}^{N \times (LM \times 4)}$ which are codomain of patch-based Hessian operator can be also defined analogously.

Then, we are able to write the discrete versions of STV and STV2 using the dual norm of mixed ℓ_1-S_1 norm. See the following Lemma 3.1.12:

Lemma 3.1.12. ([40]) Let $p \geq 1$, and let q be the conjugate exponent of p, i.e., $\frac{1}{p} + \frac{1}{q} = 1$. Then, the mixed norm $\|\cdot\|_{\infty,q}$ is dual to the mixed norm $\|\cdot\|_{1,p}$.

By the above Lemma and the fact that dual of dual norm is original norm itself, the mixed ℓ_1-S_1 norm $\|\cdot\|_{1,1}$ on space $\mathcal{X}$ can be written as follows: for $X = [X_1^T, \cdots, X_N^T]^T \in \mathcal{X}$,

$$\|X\|_{1,1} = \max_{\Omega \in \mathcal{B}^{\mathcal{X}}_{\infty,\infty}} \langle \Omega, X \rangle_{\mathcal{X}},$$

where $\mathcal{B}^{\mathcal{X}}_{\infty,\infty}$ denotes the unit ball with respect to the ℓ_∞-S_∞ norm,

$$\mathcal{B}^{\mathcal{X}}_{\infty,\infty} = \{\Omega = [\Omega_1^T, \cdots, \Omega_N^T]^T \in \mathcal{X} : \|\Omega_n\|_{\mathcal{S}_\infty} \leq 1, \forall n = 1, \cdots, N\}.$$

The dual representation of the mixed ℓ_1-S_1 norm $\|\cdot\|_{1,1}$ on space $\mathcal{Y}$ can be also obtained similarly.

3.1.4 Numerical algorithm

Now, we reformulate the proposed model (3.2) in discrete domain utilizing the dual representation of the mixed ℓ_1-S_1 norm $\|\cdot\|_{1,1}$. So to speak, STV and STV2 are represented in a different way,

$$\mathrm{STV}(\mathbf{u}) = \max_{\Omega_1 \in \mathcal{B}^{\mathcal{X}}_{\infty,\infty}} \langle \Omega_1, \mathbf{J}_K \mathbf{u} \rangle_{\mathcal{X}}, \quad \mathrm{STV2}(\mathbf{u}) = \max_{\Omega_2 \in \mathcal{B}^{\mathcal{Y}}_{\infty,\infty}} \langle \Omega_2, \mathbf{P}\mathbf{H}_K \mathbf{u} \rangle_{\mathcal{Y}},$$

where $\mathcal{B}^{\mathcal{Y}}_{\infty,\infty}$ is the unit ball on $\mathcal{Y}$ with respect to the mixed norm $\|\cdot\|_{1,1}$ and, accordingly, we need to minimize the following energy functional:

$$\begin{aligned} &\min_{\mathbf{u}} \tau_1 \max_{\Omega_1 \in \mathcal{B}^{\mathcal{X}}_{\infty,\infty}} \langle \Omega_1, \mathbf{J}_K \mathbf{u} \rangle_{\mathcal{X}} + \tau_2 \max_{\Omega_2 \in \mathcal{B}^{\mathcal{Y}}_{\infty,\infty}} \langle \Omega_2, \mathbf{P}\mathbf{H}_K \mathbf{u} \rangle_{\mathcal{Y}} + \frac{1}{2}\|\mathbf{u} - \mathbf{f}\|_2^2 \\ =&\min_{\mathbf{u}} \max_{\substack{\Omega_1 \in \mathcal{B}^{\mathcal{X}}_{\infty,\infty} \\ \Omega_2 \in \mathcal{B}^{\mathcal{Y}}_{\infty,\infty}}} \tau_1 \langle \Omega_1, \mathbf{J}_K \mathbf{u} \rangle_{\mathcal{X}} + \tau_2 \langle \Omega_2, \mathbf{P}\mathbf{H}_K \mathbf{u} \rangle_{\mathcal{Y}} + \frac{1}{2}\|\mathbf{u} - \mathbf{f}\|_2^2. \end{aligned} \tag{3.6}$$

It is noteworthy that the objective function in (3.6) is convex with respect to $\mathbf{u}$ and concave with respect to Ω_1, Ω_2. It leads the strong max-min

property, i.e., the order of the maximum and the minimum has nothing to do with the solution. Therefore, we have that (3.6) is equal to

$$\max_{\substack{\Omega_1 \in \mathcal{B}^{\mathcal{X}}_{\infty,\infty} \\ \Omega_2 \in \mathcal{B}^{\mathcal{Y}}_{\infty,\infty}}} \min_{\mathbf{u}} \tau_1 \langle \Omega_1, \mathbf{J}_K \mathbf{u} \rangle_{\mathcal{X}} + \tau_2 \langle \Omega_2, \mathbf{P}\mathbf{H}_K \mathbf{u} \rangle_{\mathcal{Y}} + \frac{1}{2} \|\mathbf{u} - \mathbf{f}\|_2^2. \tag{3.7}$$

By differentiating with respect to $\mathbf{u}$, we earn that

$$\mathbf{u} = \mathbf{f} - \tau_1 \mathbf{J}_K^* \Omega_1 - \tau_2 \mathbf{P}\mathbf{H}_K^* \Omega_2, \tag{3.8}$$

where $\mathbf{J}_K^*$ and $\mathbf{P}\mathbf{H}_K^*$ denote the adjoint of patch-based Jacobian and patch-based Hessian, respectively. If we substitute (3.8) into (3.7), then we have

$$\begin{aligned} \max_{\substack{\Omega_1 \in \mathcal{B}^{\mathcal{X}}_{\infty,\infty} \\ \Omega_2 \in \mathcal{B}^{\mathcal{Y}}_{\infty,\infty}}} & \tau_1 \langle \mathbf{J}_K^* \Omega_1, \mathbf{f} \rangle_2 - \frac{\tau_1^2}{2} \|\mathbf{J}_K^* \Omega_1\|_2^2 + \tau_2 \langle \mathbf{P}\mathbf{H}_K^* \Omega_2, \mathbf{f} \rangle_2 - \frac{\tau_2^2}{2} \|\mathbf{P}\mathbf{H}_K^* \Omega_2\|_2^2 \\ & - \tau_1 \tau_2 \langle \mathbf{J}_K^* \Omega_1, \mathbf{P}\mathbf{H}_K^* \Omega_2 \rangle_2. \end{aligned} \tag{3.9}$$

We can solve the (3.9) using the projected gradient method. If we denote the objective function in (3.9) by $d(\Omega_1, \Omega_2)$, then its gradients with respect to variables Ω_1, Ω_2 are given as follows:

$$\begin{cases} \nabla_{\Omega_1} d = \tau_1 \mathbf{J}_K (\mathbf{f} - \tau_1 \mathbf{J}_K^* \Omega_1 - \tau_2 \mathbf{P}\mathbf{H}_K^* \Omega_2), \\ \nabla_{\Omega_2} d = \tau_2 \mathbf{P}\mathbf{H}_K (\mathbf{f} - \tau_1 \mathbf{J}_K^* \Omega_1 - \tau_2 \mathbf{P}\mathbf{H}_K^* \Omega_2). \end{cases} \tag{3.10}$$

After updating each variable Ω_1, Ω_2 using the gradients obtained in (3.10), we project these onto the constrained spaces $\mathcal{B}^{\mathcal{X}}_{\infty,\infty}, \mathcal{B}^{\mathcal{Y}}_{\infty,\infty}$. It can be accomplished by projecting each component of Ω_1, Ω_2 onto the unit S_∞-norm ball independently.

For $\Omega_1 \in \mathcal{X}$, each component $\Omega_{1,n} \in \mathbb{R}^{LM \times 2}$ of Ω_1 runs the SVD decomposition $\Omega_{1,n} = U_n \Sigma_n V_n^T$. Then we change diagonal components of Σ_n to 1 if they are bigger than 1 and keep untouched unless. For $\Omega_2 \in \mathcal{Y}$, a projection

process is same.

The entire algorithm for solving the proposed model is given in Algorithm 1. Here, we adopt the Nesterov's acceleration method [46] to gain a faster convergence rate than the standard gradient ascent method. Also, the step size of gradient ascent should be suitably chosen to guarantee the convergence. We can show that the objective function in (3.9) is Lipschitz continuous gradient with proper Lipschitz constants with respect to Ω_1, Ω_2, respectively. From the Lipschitz constants, we are able to determine the suitable step sizes for updating Ω_1 and Ω_2. See Appendix A.2 for further details.

Algorithm 1 Algorithm for solving the model (3.2)

Given noisy image $\mathbf{f}$
Parameters $\tau_1 > 0$ and $\tau_2 > 0$
Initialize $\Psi_1^1 = \Omega_1^0 = 0 \in \mathcal{X},\ \Psi_2^1 = \Omega_2^0 = 0 \in \mathcal{Y},\ t_1 = 1,\ n = 1$
repeat
  $\Omega_1^n \leftarrow \mathrm{Proj}_{\mathcal{B}^{\mathcal{X}}_{\infty,\infty}}\Big(\Psi_1^n + \frac{1}{8\tau_1}\mathbf{J}_k(\mathbf{f} - \tau_1\mathbf{J}_K^*\Psi_1^n - \tau_2\mathbf{P}\mathbf{H}_K^*\Psi_2^n)\Big)$
  $\Omega_2^n \leftarrow \mathrm{Proj}_{\mathcal{B}^{\mathcal{Y}}_{\infty,\infty}}\Big(\Psi_2^n + \frac{1}{64\tau_2}\mathbf{P}\mathbf{H}_k(\mathbf{f} - \tau_1\mathbf{J}_K^*\Psi_1^n - \tau_2\mathbf{P}\mathbf{H}_K^*\Psi_2^n)\Big)$
  $t_{n+1} \leftarrow \frac{1+\sqrt{1+4t_n^2}}{2}$
  $\Psi_i^{n+1} \leftarrow \Omega_i^n + \frac{t_n - 1}{t_{n+1}}(\Omega_i^n - \Omega_i^{n-1})$ for $i = 1, 2$
  $n \leftarrow n + 1$
until a stopping criterion is satisfied
Output $\mathbf{u} = \mathbf{f} - \tau_1\mathbf{J}_K^*\Omega_1^n - \tau_2\mathbf{P}\mathbf{H}_K^*\Omega_2^n$

3.1.5 Experimental results

We show the experimental results of image denoising problems by the proposed model and compare it with the other variational methods such as STV, TGV [8], TV-TV2 [49]. Throughout experiments, we rescale image intensities to have values in range $[0, 1]$ and use 8 test images as shown in

Figure 3.3: Test images. **a** Barbara face (250×250), **b** Child (255×255), **c** Walkbridge (256×256), **d** Woman (256×256), **e** Face (254×336), **f** Desert (321×481), **g** Girl (321×481), **h** Undersea (321×481).

Figure 3.3.

We generate a noisy image by adding Gaussian noise with a standard deviation $\sigma = 0.05, 0.1$ to a clean image $\mathbf{u}_0$:

$$\mathbf{f} = \mathbf{u}_0 + \mathbf{n}, \quad \mathbf{n} \sim N(0, \sigma^2),$$

where noise acts on each pixel with i.i.d. manner.

For a quantitative measure of the restored image, we adopt the peak signal-to-noise ratio (PSNR) and the structural similarity index (SSIM) [63] which are defined as follows:

$$\begin{aligned} \text{PSNR} &= 10 \log_{10} \left(\frac{NM}{\|u - u_0\|_2^2} \right), \\ \text{SSIM} &= \frac{(2\mu_u \mu_{u_0} + C_1)(2\sigma_{uu_0} + C_2)}{(\mu_u^2 + \mu_{u_0}^2 + C_1)(\sigma_u^2 + \sigma_{u_0}^2 + C_2)}, \end{aligned} \tag{3.11}$$

where $\mathbf{u}$ is the restored image, $\mathbf{u}_0$ is the original image with N pixels and M channels, $\mu_u, \mu_{u_0}, \sigma_u, \sigma_{u_0}$ are the means and standard deviations of u, u_0 respectively, σ_{uu_0} is the cross-covariance for u, u_0 and C_1, C_2 are positive constants.

For an image whose domain consists of $N_1 \times N_2$ pixels without considering the number of channels, we pad it symmetrically to make an $2N_1 \times 2N_2$ image and apply an denoising algorithm under the periodic boundary condition. After the process is terminated, we cut it to obtain an $N_1 \times N_2$ image. Basically, parameters are selected to give a best PSNR result for each model. But there is a case where PSNR value is highest but SSIM value is relativley low and restored image is not visually fine. In this case, parameters are slightly adjusted to obtain both relatively high PSNR and SSIM values. The non-negative, rotationally symmetric convolution kernel used in STV and the proposed model is the Gaussian kernel with support 3×3 pixels and standard deviation $\sigma = 0.5$. For TGV, we use the second order version, i.e. TGV^2. For TV-TV2 and TGV, color image is processed for each channel respectively.

In Table 3.1, 3.2, we report the PSNR, SSIM values of the four different methods for 8 test images. The proposed method attains highest PSNR and SSIM values for all test images. Let us see the qualitative aspects. Since the proposed method is a combined version of first and second order regularizer, it smoothes homogeneous regions with much less staircase artifacts than the first order model. In Figures 3.4, 3.5, 3.6, 3.7 and 3.8, we demonstrate the noisy and enlarged restored images by STV and our method. Let us compare the results of our method with those of STV. For the *Desert* image in Figure 3.4, our method removes the noise in the sky region better. For the *Walkbridge* image in Figure 3.5, pay attention to water and pebbles where our method represents the details better. For the *Woman* image in Figure 3.6, the texture of hair is recovered better and staircase artifacts in the cheek of woman are reduced in our method. Lastly, for the *Face* and

Child image in Figure 3.7 and 3.8, staircase artifacts around nose and cheek are mitigated in our method.

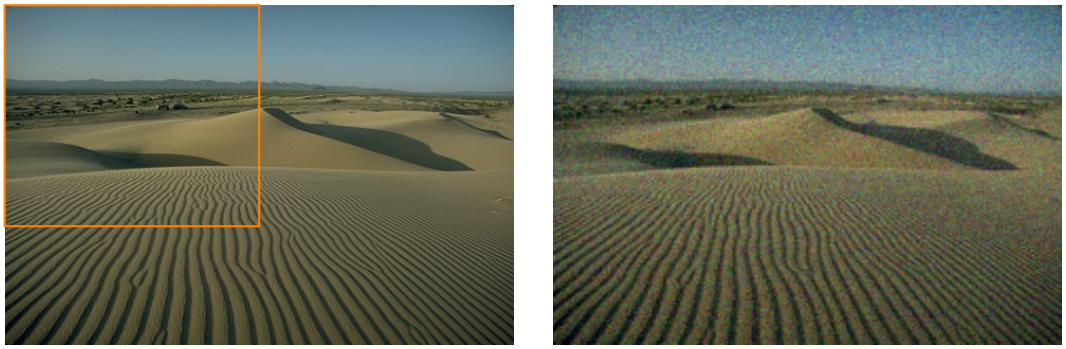
From the Figures, we can see that the results of our method show some blurry feature in return for removing staircase artifact.

	PSNR					SSIM				
	Noisy	TGV	TV-TV2	STV	Ours	Noisy	TGV	TV-TV2	STV	Ours
Barbara face	26.01	29.62	27.24	30.50	**30.82**	0.6786	0.8516	0.8097	0.8875	**0.8967**
Child	26.01	32.83	33.18	32.86	**33.31**	0.6555	0.9153	0.9259	0.9249	**0.9287**
Walkbridge	25.99	29.11	27.86	29.34	**29.59**	0.7792	0.8651	0.8305	0.8673	**0.8786**
Woman	25.99	33.41	32.99	33.67	**34.14**	0.5688	0.9192	0.9109	0.9133	**0.9233**
Face	26.01	35.73	35.74	35.16	**36.11**	0.5403	0.9418	0.9348	0.9337	**0.9476**
Desert	26.03	31.63	31.27	32.39	**32.53**	0.7823	0.9271	0.9210	0.9388	**0.9399**
Girl	26.03	32.54	31.65	33.38	**33.57**	0.7995	0.9509	0.9478	0.9625	**0.9634**
Undersea	26.03	30.42	29.64	31.15	**31.58**	0.8912	0.9582	0.9533	0.9643	**0.9676**

Table 3.1: PSNR and SSIM values for the noisy and restored images by different methods ($\sigma = 0.05$).

	PSNR					SSIM				
	Noisy	TGV	TV-TV2	STV	Ours	Noisy	TGV	TV-TV2	STV	Ours
Barbara face	19.99	26.11	25.52	26.60	**26.87**	0.4278	0.7335	0.7115	0.7586	**0.7746**
Child	19.99	29.85	29.70	29.22	**30.23**	0.4044	0.8659	0.8666	0.8622	**0.8676**
Walkbridge	19.97	25.68	25.77	25.98	**26.17**	0.5197	0.7280	0.7373	0.7452	**0.7572**
Woman	19.97	30.47	30.29	30.39	**30.84**	0.3063	0.8569	0.8559	0.8552	**0.8686**
Face	19.99	32.57	32.31	31.70	**32.72**	0.2654	0.9042	0.8862	0.8776	**0.9043**
Desert	20.01	28.74	29.02	29.37	**29.48**	0.4937	0.8625	0.8693	0.8789	**0.8820**
Girl	20.01	29.48	29.46	30.14	**30.35**	0.5869	0.9153	0.9169	0.9266	**0.9274**
Undersea	20.01	27.17	27.37	27.77	**28.02**	0.6988	0.9152	0.9207	0.9261	**0.9303**

Table 3.2: PSNR and SSIM values for the noisy and restored images by different methods ($\sigma = 0.1$).



Ground truth

Noisy image

STV

Ours

Figure 3.4: Enlarged denoising results on image *Desert* by STV and proposed method ($\sigma = 0.05$).

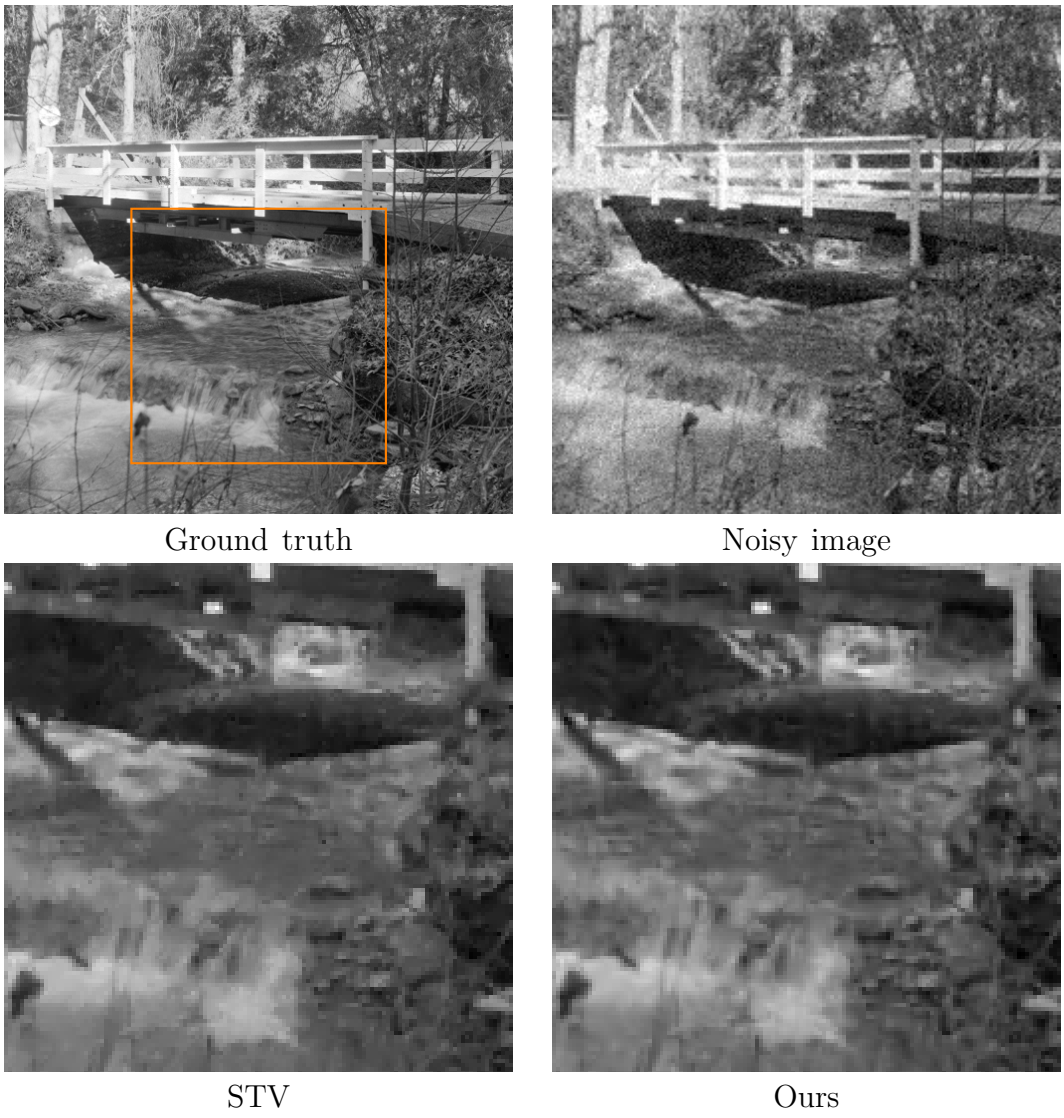


Figure 3.5: Enlarged denoising results on image *Walkbridge* by STV and proposed method ($\sigma = 0.05$).

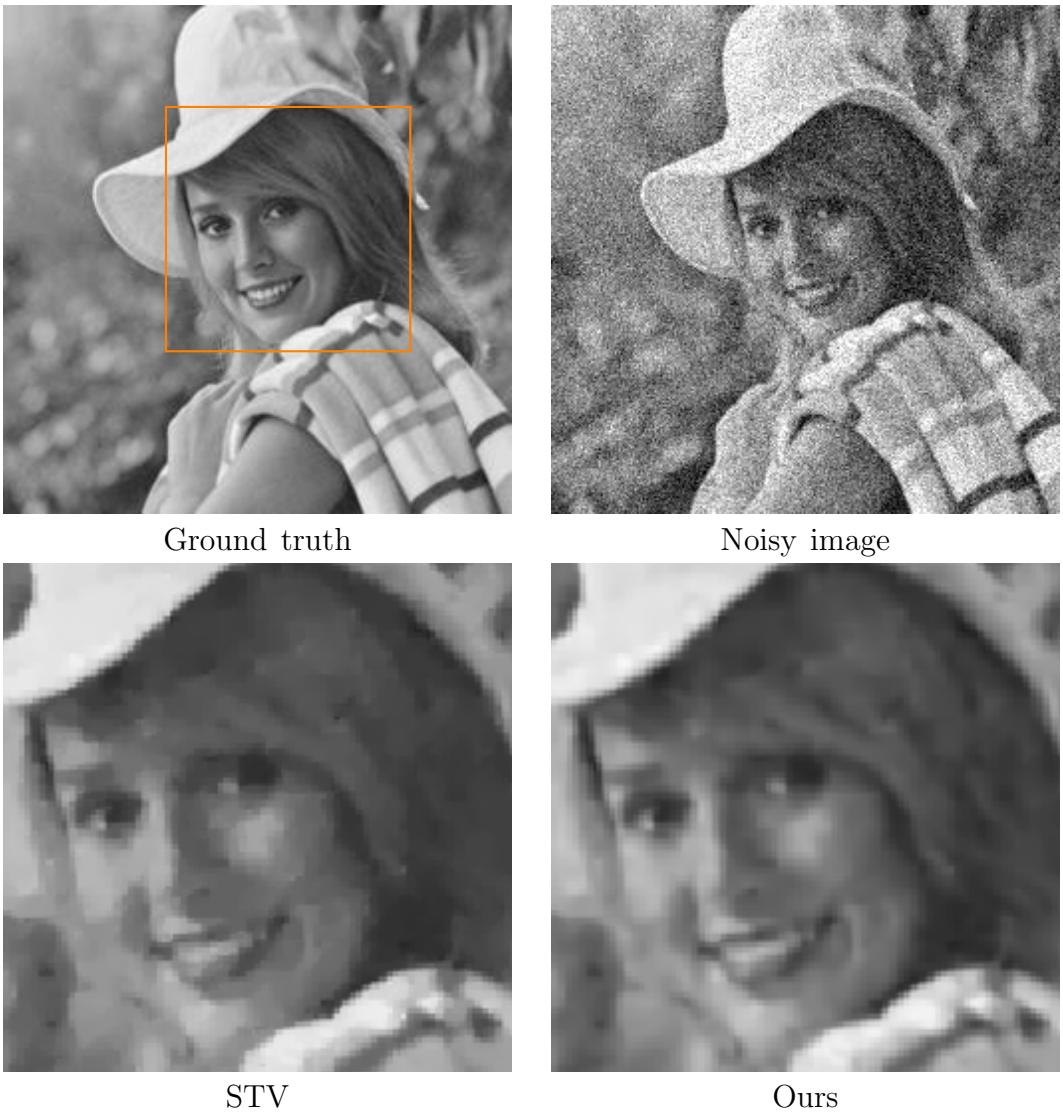


Figure 3.6: Enlarged denoising results on image *Woman* by STV and proposed method ($\sigma = 0.1$).

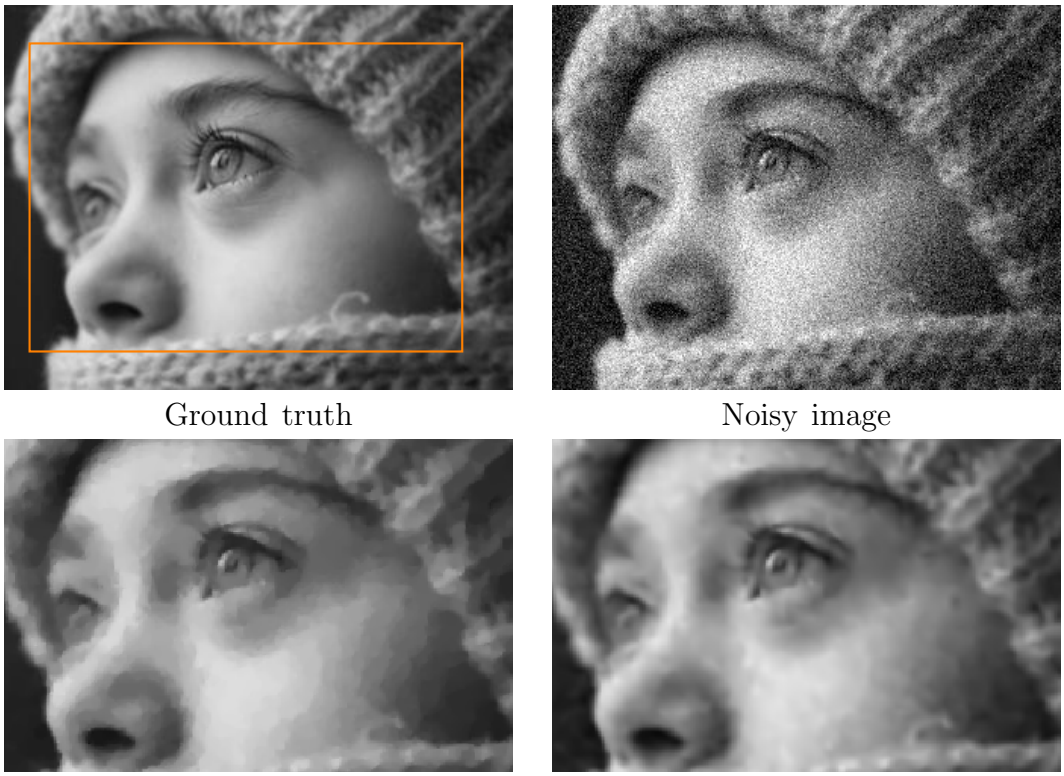


Figure 3.7: Enlarged denoising results on image *Face* by STV and proposed method ($\sigma = 0.1$).

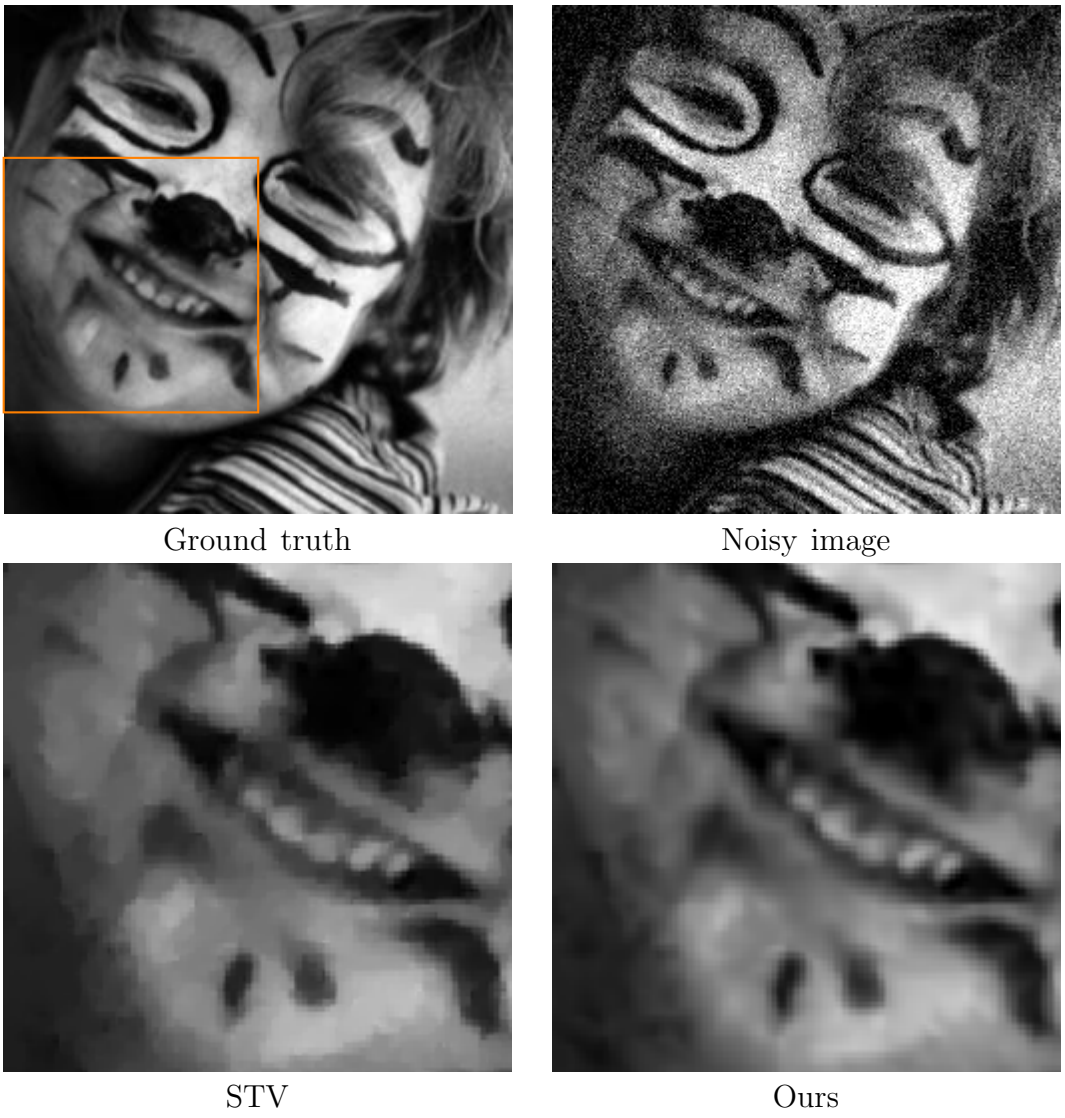


Figure 3.8: Enlarged denoising results on image *Child* by STV and proposed method ($\sigma = 0.1$).

3.2 Weighted nuclear norm minimization for Cauchy noise

Recently, weighted nuclear norm minimization (WNNM), which regularizes singular values of an input matrix with different strengths according to given weights, has demonstrated impressive results in low-level vision tasks such as additive Gaussian noise removal, deblurring and image inpainting [28, 29, 65]. In this section, we apply WNNM to remove additive Cauchy noise in images. A variational model is adopted based on maximum a posteriori (MAP) estimate, which contains a data fidelity term that is appropriate for noise following Cauchy distribution. Weighted nuclear norm is used as a regularizer in the proposed algorithm, and we utilized similar patches in the image by nonlocal similarity. We adopted the nonconvex alternating direction method of multiplier (ADMM) to solve the problem iteratively. Numerical experiments are presented to demonstrate the superior denoising performance of our algorithm compared with other existing methods in terms of quantitative measure and visual quality.

3.2.1 Variational models for Cauchy noise

As explained in the introduction, the impulsive-like noise appears in many scientific and engineering circumstances and it can be modeled by utilizing Cauchy distribution. We assume that the image y is degraded by the additive Cauchy noise n which follows Cauchy distribution $\mathcal{C}(\delta, \gamma)$ is given as follows:

$$y = u_0 + n,$$

where the original image u_0 is defined on a bounded domain $\Omega \subset \mathbb{R}^2$, and the probability density function (PDF) of Cauchy distribution $\mathcal{C}(\delta, \gamma)$ for each

$x \in \Omega$ is given by:

$$P(x; \delta, \gamma) = \frac{1}{\pi\gamma} \frac{\gamma^2}{\gamma^2 + (x - \delta)^2}, \tag{3.12}$$

where $\gamma > 0$, $\delta \in \mathbb{R}$ are called the scale and localization parameter, respectively.

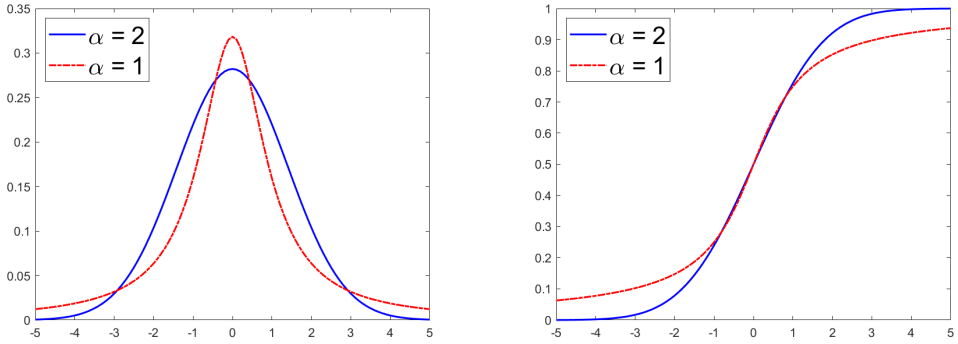
Cauchy distribution

Cauchy distribution is a special case of an analytically expressible α-stable distribution [47]. An α-stable distribution has four parameters: an index of stability $\alpha \in (0, 2]$, a skewness parameter $\beta \in [-1, 1]$, a scale parameter $\gamma \geq 0$, and a location parameter $\delta \in \mathbb{R}$. For a distribution to be symmetric, the skewness parameter β should be 0. If a random variable X follows the symmetric α-stable distribution, then its PDF $p(x)$ is given by:

$$p(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp\left(it\delta - \gamma^\alpha |t|^\alpha\right) e^{-itx} dt.$$

The index of stability α is a parameter that determines the thickness of the tail in the distribution, δ is the location parameter that indicates the position of the distribution peak, and γ is the scale parameter that represents the half-width at half-maximum. Specifically, $\alpha = 1$ and $\alpha = 2$ represent Cauchy distribution and normal distribution, respectively. Figure 3.9 shows the difference in shape between α-stable distributions while α varies. We can see that, as α decreases, the tail of the distribution becomes thicker, implying that noise following Cauchy distribution has more impulsive characteristic than that following the normal distribution. The PDF of Cauchy distribution is given in (3.12), and the cumulative distribution function (CDF) of Cauchy distribution [25] is given by

$$F(x; \delta, \gamma) = \frac{1}{\pi} \arctan\left(\frac{x - \delta}{\gamma}\right) + \frac{1}{2}.$$



(a) (b)

Figure 3.9: Graphs of (a) probability density functions and (b) cumulative distribution functions of α-stable distribution. The normal distribution corresponds to the case when $\alpha = 2$ and Cauchy distribution corresponds to the case when $\alpha = 1$.

Note that an interesting relationship exists between the normal distribution and Cauchy distribution:

Proposition 3.2.1. Let $N(\mu, \sigma^2)$ denote the normal distribution with mean μ, standard deviation σ, and X and Y be two independent random variables. Assume that $X \sim N(0, \sigma_x^2)$ and $Y \sim N(0, \sigma_y^2)$. Then, for the random variable $Z = X/Y$, $Z \sim C(0, \frac{\sigma_x^2}{\sigma_y^2})$.

Proof. The joint distribution of X and Y has the following PDF:

$$f_{X,Y}(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} e^{-\frac{x^2}{2\sigma_x^2} - \frac{y^2}{2\sigma_y^2}}.$$

From [20], the PDF of the distribution of $Z = X/Y$ is as follows:

$$\begin{aligned} p(z) &= \int_{-\infty}^{\infty} |y| f_{X,Y}(zy, y) dy = \frac{1}{\pi\sigma_x\sigma_y} \int_0^{\infty} y e^{-y^2\left(\frac{z^2}{2\sigma_x^2} + \frac{1}{2\sigma_y^2}\right)} dy \\ &= \frac{1}{\pi\sigma_x\sigma_y} \frac{1}{\frac{z^2}{\sigma_x^2} + \frac{1}{\sigma_y^2}} = \frac{1}{\pi} \frac{\sigma_x/\sigma_y}{z^2 + (\sigma_x/\sigma_y)^2}. \end{aligned}$$

Therefore, according to (3.12), $p(z)$ is the PDF of Cauchy distribution with location parameter 0 and scale parameter σ_x/σ_y. □

Numerically, Proposition 3.2.1 is useful. We can generate Cauchy noise with arbitrary scale parameter γ if we could generate Gaussian noise with the corresponding standard deviation.

Denoising model via MAP estimate

We use MAP estimate to derive a variational model for the additive Cauchy noise assuming some priors on the distribution of images as in Section 2.1.1. We define Ω as an image domain and assume that a clean image is contaminated by additive Cauchy noise, which follows $C(0, \gamma)$. Therefore, by the MAP, the estimated image u from noisy image y can be obtained as follows:

$$\begin{aligned} u^* &= \arg\max_u P(u|y) = \arg\max_u \frac{P(y|u)P(u)}{P(y)} \\ &= \arg\max_u P(y|u)P(u). \end{aligned} \tag{3.13}$$

We take the logarithm of (3.13). Then, we have

$$\begin{aligned} u^* &= \arg\min_u -\log P(y|u) - \log P(u) \\ &= \arg\min_u \int_\Omega -\log P(y(x)|u(x))dx - \log P(u). \end{aligned} \tag{3.14}$$

The first term in (3.14) indicates the occurrence probability of a noisy pixel value $y(x)$ given a clean pixel value $u(x)$. It is noteworthy that the noise distribution affecting each $x \in \Omega$ is independent and identically distributed. Hence, from the PDF of Cauchy distribution (3.12), we have

$$\log P(y(x)|u(x)) = -\log \pi + \log \gamma - \log(\gamma^2 + (u(x) - y(x))^2).$$

Furthermore, a priori knowledge on the distribution of clean image u is required. For example, the Gibbs prior is used in [45, 56] such that $P(u) = \frac{1}{Z} \exp\left(-\beta \int_\Omega |Du|\right)$ where $\beta > 0$, Z is a normalization factor. After omitting the constant terms, the resulting variational model is

$$\arg\min_u \int_\Omega \log\left(\gamma^2 + (u(x) - y(x))^2\right) dx - \log P(u). \tag{3.15}$$

It is noteworthy that the first term in (3.15) serves as data fidelity for denoising an image corrupted by Cauchy noise.

Previous models for denoising Cauchy noise

Several approaches have been proposed to address image corruption by Cauchy noise. Chang *et al.* [17] suggested a recursive restoration algorithm based on a Markov random field model. Achim and Kuruoğlu [2] designed a bivariate maximum a posteriori (MAP) estimator and restored a degraded image through complex wavelet transform. Laus *et al.* [37] introduced a generalized myriad filter and utilized it nonlocally for estimating the localization parameter of Cauchy distribution, which can be considered as the latent original pixel value of the image. Sciacchitano *et al.* [56] introduced a variational model via the MAP with a TV regularizer. It is written as follows:

$$\inf_{u \in \mathrm{BV}(\Omega)} \int_\Omega |Du| + \frac{\lambda}{2} \int_\Omega \log\left(\gamma^2 + (u - y)^2\right), \tag{3.16}$$

where Du denotes the distributional derivative of u. Because the data fidelity term in (3.16) is nonconvex, the authors subsequently proposed a convexified model by inserting a quadratic term additionally:

$$\inf_{u\in \mathrm{BV}(\Omega)} \int_\Omega |Du| + \frac{\lambda}{2}\left(\int_\Omega \log\left(\gamma^2 + (u-y)^2\right) + \rho\|u-\tilde{u}\|_2^2\right), \tag{3.17}$$

where $\rho > 0$ and $\tilde{u}$ is obtained by applying a median filter to y. If γ, ρ in (3.17) satisfy $8\rho\gamma^2 \geq 1$, the (3.17) becomes strictly convex. From the convexity of the model, we can guarantee the uniqueness of the solution irrespective of the initial condition, but the presence of the quadratic term induces a deviation from the model (3.16) derived from the MAP principle, which could bias the solution.

Mei *et al.* [45] adopted the nonconvex ADMM [62] which established the global convergence of an algorithm for nonconvex nonsmooth optimization with linear constraints and directly applied it to (3.16) which outperformed the results in [56].

The previously mentioned variational models are based on a TV regularization, and only considers the local properties of an image. However, recent studies demonstrated that nonlocal similarity-based approaches achieved better performance in restoring images corrupted by Gaussian noise [10, 21]. Primarily, WNNM which is stemmed from methods approximating a low-rank matrix, has demonstrated performance comparable to that of state-of-the-art methods for the denoising problem [28, 29, 60, 64].

Based on low-rank matrix approximation, we establish a variational model to restore images corrupted by Cauchy noise using the weighted nuclear norm. Our contribution is threefold. First, we connect the variational approach of Cauchy noise model to the WNNM model. Next, we apply the nonconvex ADMM to our proposed model for Cauchy noise and prove the convergence of the algorithm to a stationary point. Finally, we present novel methods for obtaining similar patches based on the median filter and defining weights for

the nuclear norm, which is suitable for Cauchy noise.

3.2.2 Low rank minimization by weighted nuclear norm

Low-rank matrix approximation aims to recover a low-rank matrix from limited or degraded information. It is required in many instances such as the Netflix problem as a representative example in the area of recommendation systems, triangulation from incomplete data and a low-dimensional embedding of data in Euclidean space [12, 53]. It also has shown excellent performance in image processing such as image restoration and alignment [51, 60]. In image denoising, nonlocal-based methods have demonstrated notable results. Not being confined to the local information of a given patch, they utilize similar patches across the entire image. If we flatten and stack those similar patches into a matrix, we may assume that the stacked matrix exhibit the low-rank property. For a local patch y_j of the image y, we can construct a matrix $Y \in \mathbb{R}^{n\times m}$ by stacking the similar flattened patches of y_j as columns. Then, for the noisy image corrupted by Gaussian noise, we consider the following low-rank minimization problem:

$$\arg\min_X \frac{1}{2}\|Y - X\|_F^2 + \|\Sigma\|_0,$$

where Σ is a diagonal matrix whose diagonal elements are singular values of X (i.e., $U\Sigma V^T$ is the singular value decomposition (SVD) of X) and $\|\cdot\|_0$ is ℓ_0 quasi-norm, which counts the number of nonzero values. Because a direct minimization of the ℓ_0 quasi-norm is a NP-hard problem, a relaxed model that replaces non-convex ℓ_0 quasi-norm with convex ℓ_1 norm is suggested [12, 53]:

$$\arg\min_X \frac{1}{2}\|Y - X\|_F^2 + \|X\|_*,$$

where $\|X\|_*$ is the sum of singular values of X and is called the nuclear norm of X, which satisfies $\|X\|_* = \|\Sigma\|_1$. However, because the ℓ_1 norm only approximates the ℓ_0 quasi-norm, it is insufficient for achieving the desired low-rank property.

Candés *et al.* [14] demonstrated that the weighted ℓ_1 norm could enhance the sparsity of the ℓ_1 norm. The weighted nuclear norm was first applied to an image denoising model as a regularizer by Gu *et al.* [29]. They applied the weighted nuclear norm to the Gaussian denoising problem. Specifically, for a noisy matrix $Y \in \mathbb{R}^{n\times m}$ whose columns are flattened patches similar to the reference patch, the denoised matrix $\hat{X} \in \mathbb{R}^{n\times m}$ is given as follows:

$$\hat{X} = \arg\min_X \frac{1}{2}\|Y - X\|_F^2 + \|X\|_{\mathbf{w},*}, \tag{3.18}$$

where $\|X\|_{\mathbf{w},*} = \sum_i w_i\sigma_i(X)$ is the weighted nuclear norm for $w_i \geq 0$, $i = 1, \cdots, n_0$, $n_0 = \min(m, n)$ and $\mathbf{w} = [w_1, w_2, \ldots, w_{n_0}]$. In general, the weighted nuclear norm is non-convex, thus rendering the problem (3.18) more challenging to solve owing to the presence of local minima. However, according to [66], we can obtain the optimal solution of (3.18).

Theorem 3.2.2. Let $Y = U\Sigma V^T$ be the SVD of $Y \in \mathbb{R}^{n\times m}$. If the weights satisfy $0 \leq w_1 \leq \cdots \leq w_{n_0}$, $n_0 = \min(n, m)$, the problem (3.18) has an optimal solution,

$$\hat{X} = U\text{soft}_W(\Sigma)V^T, \tag{3.19}$$

where W is the diagonal matrix such that $W_{ii} = w_i$ for $i = 1, \cdots, n_0$, and $\text{soft}_W(\Sigma)$ is the generalized soft-thresholding operator:

$$\text{soft}_W(\Sigma)_{ij} = \max(\Sigma_{ij} - W_{ij}, 0).$$

Proof. The solution to the problem (3.18) should satisfy the following:

$$\mathbf{0} \in Y - \hat{X} + \partial\|\hat{X}\|_{w,*}.$$

Here, $\partial\|\hat{X}\|_{w,*}$ denotes the subgradient of the weighted nuclear norm $\|\hat{X}\|_{w,*}$. Let the SVD of a matrix $X \in \mathbb{R}^{n\times m}$ be $U_X\Sigma_X V_X^T$ and $\text{rank}(X) = r$. Therefore, according to [66], the subgradient of the weighted nuclear norm has the following property.

$$\partial\|X\|_{w,*} = \Big\{U_X W_r V_X^T + Z : Z \in \mathbb{R}^{n\times m},\ U_X^T Z = 0,\ ZV_X = 0, \\ \sigma_{i-r}(Z) \le w_i \text{ for } i = r+1,\dots,n_0\Big\}, \tag{3.20}$$

where W_r is a diagonal matrix whose first r diagonal elements are $w_1, \cdots, w_r$ and all other elements are zero.

We can rewrite Σ as $\Sigma_0 + \Sigma_1$ where Σ_0 is the diagonal matrix whose i-th diagonal element is $\sigma_i(Y)$ if it is larger than w_i and 0 otherwise; conversely, Σ_1 is a diagonal matrix whose i-th diagonal element is $\sigma_i(Y)$ if it is smaller than or equal to w_i and 0 otherwise. Additionally, we define U_0, V_0 (resp. U_1, V_1) whose columns consist of the singular vectors of Y if they correspond to the singular values in Σ_0 (resp. Σ_1) and zero vectors otherwise. Subsequently, the SVD of $Y = U\Sigma V^T$ can be written as $U_0\Sigma_0 V_0^T + U_1\Sigma_1 V_1^T$. Let us define $\hat{X}$ as

$$\hat{X} = U_0 \text{soft}_{W_k}(\Sigma_0) V_0^T,$$

where W_k is the diagonal matrix with the main diagonal elements $[w_1, \cdots, w_k, 0, \cdots, 0]$ and k is the rank of Σ_0. Then, we have

$$\begin{aligned} Y - \hat{X} &= U_0 W_r V_0^T + U_1\Sigma_1 V_1^T \\ &= U_0 W_r V_0^T + Z, \end{aligned} \tag{3.21}$$

where Z is the substitution of $U_1\Sigma_1 V_1^T$. Here, it is sufficient to show that (3.21) satisfies (3.20). By the definitions of U_0, U_1, V_0, and V_1, it follows that $U_0^T Z = 0$ and $ZV_0 = 0$. Furthermore, the singular values of Z are the diagonal entries of Σ_1; therefore, $\sigma_i(Z) \le w_{i+r}$ for $i = 1, \cdots, n_0 - r$. □

Moreover, according to [28], the solution (3.19) becomes the global optimum solution.

3.2.3 Proposed method

Now, we introduce our variational model for denoising Cauchy noise, which adopts the nonlocal method that uses the weighted nuclear norm as a regularization. We combine the pertinent fidelity term for Cauchy noise with the weighted nuclear norm. A direct insertion of the weighted nuclear norm as a regularization of the model (3.15) is not reasonable because the natural image is not guaranteed to be a low-rank. Instead, following the process in [29], we adopt the nonlocal approach, where we group similar patches according to a specific metric, then vectorize and stack them to form a two-dimensional matrix. On the matrix that we construct, we apply the WNNM with the data fidelity term for Cauchy noise. In other words, for a noisy im-

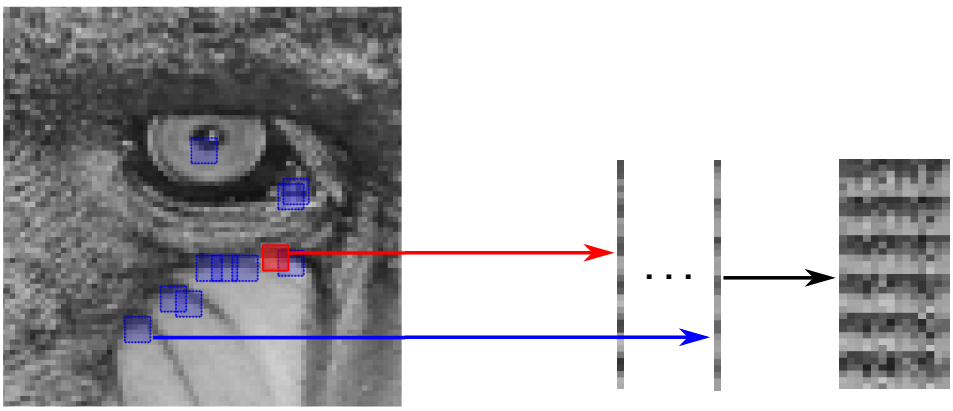


Figure 3.10: Illustration of the process of constructing matrix composed of similar patches.

age y, we consider a small patch y_j as a reference. Next, we obtain a group of patches similar to the reference, then flatten, and stack them to form Y_j.

On this Y_j, we solve the following problem:

$$\arg\min_{X_j} \ \|X_j\|_{\mathbf{w},*} + \sum_{x\in X_j, y\in Y_j} \log\left(\gamma^2 + (x-y)^2\right), \tag{3.22}$$

where $\|X\|_{\mathbf{w},*}$ is the weighted nuclear norm of X, and we adopted the discrete version of the fidelity term in (3.15).

3.2.4 ADMM algorithm

The minimization problem (3.22) still needs to be solved. In [45], the nonconvex ADMM algorithm was used for Cauchy noise denoising where the variational model consisted of a total variation regularizer and a fidelity term of the form as in (3.15). In this Section, we investigate the nonconvex ADMM algorithm and apply it to solve our problem. We can reformulate (3.22) as follows:

$$\arg\min_{X} \ \|X\|_{\mathbf{w},*} + \frac{\lambda}{2}\langle \log\left(\gamma^2 + (X-Y)^2\right), \mathbf{1}\rangle, \tag{3.23}$$

where discrete image $Y \in \mathbb{R}^{n\times m}$ is given and $\langle\cdot,\cdot\rangle$ is the inner product considering X, Y as vectors with mn elements. We substitute X in the fidelity term with the auxiliary variable V to such that the problem is of the constrained form.

$$\arg\min_{X,V} \ \|X\|_{\mathbf{w},*} + \frac{\lambda}{2}\langle \log\left(\gamma^2 + (V-Y)^2\right), \mathbf{1}\rangle, \quad X = V. \tag{3.24}$$

Next, we introduce the corresponding augmented Lagrangian of (3.24) with Lagrangian multiplier $W \in \mathbb{R}^{n\times m}$ and penalty parameter $\beta > 0$:

$$\mathcal{L}_\beta(X,V,W) = \|X\|_{\mathbf{w},*} + \frac{\lambda}{2}\langle \log\left(\gamma^2 + (V-Y)^2\right), \mathbf{1}\rangle + \langle W, X-V\rangle + \frac{\beta}{2}\|X-V\|^2. \tag{3.25}$$

Subsequently, the nonconvex ADMM asserts that we can alternatively minimize the Lagrangian $\mathcal{L}_\beta$ above with respect to X and V to solve the given constrained problem (3.24). The corresponding algorithm is given in Algorithm 2.

Algorithm 2 Nonconvex-ADMM for (3.22)

Given noisy Y
Parameter $\beta > 0$
Initialize X^0, V^0, W^0, $k = 0$
repeat
    $X^{k+1} = \arg\min_X \|X\|_{\mathbf{w},*} + \frac{\beta}{2}\left\|X - V^k + \frac{W^k}{\beta}\right\|_2^2$
    $V^{k+1} = \arg\min_V \frac{\lambda}{2}\left\langle \log\left(\gamma^2 + (V-Y)^2\right), \mathbf{1}\right\rangle + \frac{\beta}{2}\left\|X^{k+1} - V + \frac{W^k}{\beta}\right\|_2^2$
    $W^{k+1} = W^k + \beta(X^{k+1} - V^{k+1})$
    $k \leftarrow k + 1$
until X^{k+1} satisfies the stopping criterion
Output X^{k+1}

Following [45, 62], we can prove the convergence results of Algorithm 2 under suitable conditions.

Theorem 3.2.3. Let $\{(X^k, V^k, W^k)\}_{k\in\mathbb{N}}$ be the sequence generated by Algorithm 2. If $\beta > \frac{2\lambda}{\gamma^2}$, the sequence $\{(X^k, V^k, W^k)\}$ has at least one limit point and each limit point is a stationary point of $\mathcal{L}_\beta$.

A proof of Theorem 3.2.3 is given in Appendix B.1. The overall algorithm for denoising the given image corrupted by Cauchy noise is shown in Algorithm 3.

Algorithm 3 Algorithm for Cauchy noise denoising

Given noisy image y
Initialize $x^{(0)}$, $k = 0$
while $x^{(k)}$ do not satisfy the stopping criterion **do**
  for each patch $y_j^{(k)}$ in $x^{(k)}$ **do**
    Find similar patch group and stack it to form $Y_j^{(k)}$
    Estimate weight vector $\mathbf{w}$
    Apply nonconvex-ADMM to solve

$$\arg\min_{X_j^{(k)}} \|X_j\|_{\mathbf{w},*} + \sum_{x_k \in X_j^{(k)}, y_k \in Y_j^{(k)}} \log\left(\gamma^2 + (x_k - y_k)^2\right)$$

    return improved patch $y_j^{(k+1)}$
  end for
  Aggregate patches $y_j^{(k+1)}$ to form enhanced image $x^{(k+1)}$
  $k \leftarrow k + 1$
end while
Output Restored image $x^{(k)}$

3.2.5 Numerical method and experimental results

We present the results of several experiments to demonstrate the performance of the proposed algorithm for Cauchy noise removal. We used ten 256×256 gray-scale test images in $[0, 255]$ for our experiments, as shown in Figure 3.11. For the original image u_0, we generate the noisy image y corrupted by Cauchy noise n which follows Cauchy distribution $\mathcal{C}(0, \gamma)$:

$$y = u_0 + n = u_0 + \gamma \frac{\eta_1}{\eta_2},$$

where we used Proposition 3.2.1 and random variables η_1, η_2 following the normal distribution $N(0, 1)$.

We compare our algorithm with other methods for Cauchy noise removal including the convex total variation (CTV) method [56], nonconvex total variation (NCTV) method [45] and weighted generalized nonlocal myriad filters (wGNMF) [37]. For a quantitative measure of the restored image, we adopt the peak signal-to-noise ratio (PSNR) and the structural similarity index (SSIM) [63] which are defined as in (3.11).

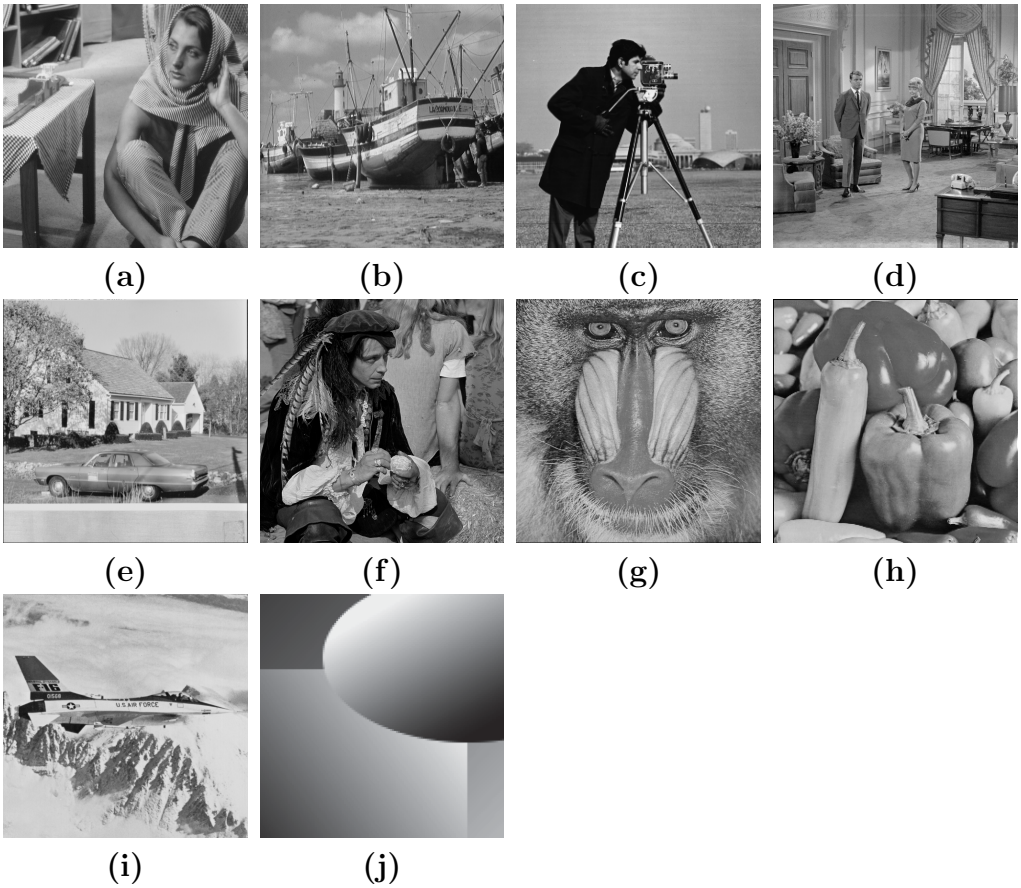


Figure 3.11: Test images with size of 256×256. **a** Barbara, **b** Boat, **c** Cameraman, **d** Couple, **e** House, **f** Man, **g** Mandrill, **h** Peppers, **i** Plane, **j** Synthetic

The parameters related to the proposed algorithm affect the experimental results. For the subproblem line 5 in Algorithm 2, we set the patch size to 6×6, stacked 70 similar patches for each reference patch moving with a stride of 4, and set β to be slightly larger than $\frac{2\lambda}{\gamma^2}$. For each stacked patch $\mathbf{Y}$, the weight vector $\mathbf{w} = [w_1, \cdots, w_n]$ is estimated as

$$w_i = \min\left(\frac{c\Theta^2\sqrt{n}}{\sigma_i(\mathbf{Y}) + \epsilon},\ \Theta^2\right),$$

where $\sigma_i(\mathbf{Y})$ is the i-th largest singular value of $\mathbf{Y}$, n is the number of stacked patches and c, Θ are constants. We observed in the experiment that as the denoising algorithm iterates, the small singular values of the newly stacked patch matrix $\mathbf{Y}$ approaches 0. By applying a soft-thresholding operator, the large weight corresponding to the small singular value of $\mathbf{Y}$ causes the singular value to be 0; thus detailed information with respect to the image is lost. To prevent the weight from being excessively large, we truncated the weight by a constant Θ^2. We set $\lambda = 1$, $c = 23$, $\Theta = 2.7$ for $\gamma = 5$ and $\Theta = 4$ for $\gamma = 10$ in our experiments.

Initialization is another issue because our model is non-convex. We obtained a starting image $x^{(0)}$ by eliminating noisy pixels that have extreme values. We applied a 3×3 median filter to a noisy image y to obtain a filtered image $\tilde{y}$ and truncated y by the minimum and maximum of all pixel values of $\tilde{y}$. In other words,

$$x^{(0)} = \max(m, \min(M, y)),$$

where m, M are minimum and maximum among all pixel values of $\tilde{y}$ respectively.

As we did not conduct an exhaustive search for other possible choices of parameters such as patch size and number of stacked patches, the performance of our algorithm can be further improved.

Block matching method

We used the Frobenius norm to compare the similarity of two patches. For patches p_a and p_b, we computed the distance between them as follows:

$$\mathrm{Dist}(p_a, p_b) = \|p_a - p_b\|_F^2.$$

However, because Cauchy noise is considerably impulsive compared with Gaussian noise [56], noisy pixels of impulsive nature can completely remove the original pixel value, thus complicating the similarity measurement of the patches. Instead of determining the distance directly, we applied a 3×3 median filter (`medfilt2` in MATLAB) to the noisy image and applied the Frobenius norm to obtain similar patches based on the filtered image.

Because the image improves as iteration step k increases, there is a k^*-th such that the recovered image X^{k^*} has better visual quality than $\texttt{medfilt2}(X^{k^*})$; which is obtained by applying the median filter to the recovered image. After step k^*, patch matching using the reconstructed image X^{k^*} produced better results than those using the image $\texttt{medfilt2}(X^{k^*})$. At this point, it is better to obtain similar patches based on the recovered image X^{k^*} instead of the filtered image. To obtain such k^*, we considered the total variation of the recovered image. The total variation for the image u, denoted by $\mathrm{TV}(u)$, is defined as follows [54]:

$$\mathrm{TV}(u) = \sum_{i,j} \sqrt{|u_{i+1,j} - u_{i,j}|^2 + |u_{i,j+1} - u_{i,j}|^2}.$$

We computed the ratio of the total variation of the recovered image X^k to that of the $\texttt{medfilt2}(X^k)$. At the early iteration step of the algorithm, the noisy pixel of impulsive nature still remained in the recovered image. The median filter erased most of these noisy pixels, thus reducing the sharpness of the image. Therefore, applying the median filter to the recovered image reduced the total variation of the image significantly. However, as the

iteration proceeded, the number of such impulsive pixels in the recovered image decreased and median filtering did not significantly reduce the total variation of the image. Hence, we set k^* as the first step k such that $\mathrm{TV}(X^k)/\mathrm{TV}(\texttt{medfilt2}(X^k)) \leq 3$ is satisfied.

Overall, at the early stage of the algorithm, we matched the similar patches based on $\texttt{medfilt2}(X^k)$. If the criterion $\mathrm{TV}(X^k)/\mathrm{TV}(\texttt{medfilt2}(X^k)) \leq 3$ is first satisfied at the iteration step k^*, then we obtain the patches based on X^k at the (k^*+1)-th step and thereafter.

Stopping criterion based on energy estimation

We define an energy function to be utilized for setting a stopping criterion of our algorithm. Because the weight of the weighted nuclear norm varies with respect to iteration, it is inappropriate to set the termination condition using the objective function (3.23). Instead, we use the nuclear norm for the energy function E:

$$E(X) = \|X\|_* + \frac{\lambda}{2} \sum_{x \in X, y \in Y} \log\left(\gamma^2 + (x-y)^2\right),$$

where X is a stacked matrix of flattened similar patches of the reference patch in the restored image and Y is the corresponding matrix whose elements are selected from the noisy image f. Theoretically, we have to compute $E(X)$ for every reference patch in the image. However, this requires high computational cost; therefore, we computed E only for the patch $\hat{p}$, which has the smallest standard deviation among the patches of filtered image $\texttt{medfilt2}(f)$. Subsequently, we terminated the algorithm at iteration k when the criterion

$$\frac{|E(X^{(k-1)}) - E(X^{(k)})|}{|E(X^{(k)})|} < \epsilon$$

was satisfied for threshold ϵ. In the experiments, we set $\epsilon = 1\mathrm{e}{-3}$.

Experimental results

In Table 3.2.5 and 3.2.5, we present the PSNR and SSIM of noisy images recovered by CTV [56], NCTV [45], nonlocal myriad filters [37] and our method. The best PSNR or SSIM values among all the denoising methods are marked in boldface. Based on the results, our proposed method performed better than the other methods in terms of the PSNR and SSIM values for most of the test images.

In Figures 3.12, 3.13, 3.14, 3.15, 3.16 and 3.17, we demonstrate the noisy and restored images by CTV, NCTV, wGNMF and the proposed method. As shown, CTV overly smoothes the image, thus resulting in detail loss and the appearance of staircase artifacts. NCTV preserves the details of the image better than CTV but some noisy features of Cauchy noise still remain, which are shown as salt-and-pepper-like pixels. The wGNMF causes the image to be slightly blurred. Our method exhibits sharp edges and preserves detailed features in the restored images while removing most of the impulsive-like noisy pixels. For example, for the *Cameraman* images in Figure 3.13, we can identify the mentioned characteristics in the background and the camera of the restored images. Additionally, we can verify that our method restored the details better than the compared methods, as shown in the checked pattern of the tablecloth in the image *Barbara* in Figure 3.12, and the letters on the body of the airplane in the image *Plane* in Figure 3.14. Our method well restores the smooth parts of the image *Synthetic* in Figure 3.15. For noisy images corrupted by a stronger noise level such as $\gamma = 10$, our method still demonstrated significantly better restoration performance compared with other methods, as shown in Figure 3.16, 3.17.

	PSNR					SSIM				
	Noisy	CTV	NCTV	Myriad	Ours	Noisy	CTV	NCTV	Myriad	Ours
Barbara	19.21	29.41	29.49	30.98	**33.17**	0.4265	0.8500	0.8533	0.8871	**0.9306**
Boat	19.19	29.45	29.92	29.53	**30.91**	0.4058	0.8519	0.8693	0.8523	**0.8905**
Cameraman	19.23	27.97	28.36	28.96	**31.00**	0.3563	0.8571	0.8482	0.8376	**0.9001**
Couple	19.19	28.54	29.55	29.34	**30.92**	0.4243	0.8190	0.8650	0.8557	**0.9004**
House	19.13	29.04	29.33	29.55	**30.49**	0.4333	0.8705	0.8775	0.8699	**0.9092**
Man	19.20	28.67	28.84	28.83	**29.85**	0.4744	0.8422	0.8604	0.8582	**0.8781**
Mandrill	19.21	26.46	27.00	27.92	**28.40**	0.5000	0.7499	0.8117	0.8332	**0.8525**
Peppers	19.24	30.92	31.00	30.67	**32.50**	0.3859	0.8929	0.8854	0.8712	**0.9166**
Plane	19.20	30.23	30.47	30.05	**31.37**	0.3727	0.9021	0.8896	0.8654	**0.9228**
Synthetic	19.21	39.36	41.51	35.68	**43.81**	0.1498	0.9609	0.9728	0.8394	**0.9833**

Table 3.3: PSNR and SSIM for the noisy and restored images by different methods ($\gamma = 5$).

	PSNR					SSIM				
	Noisy	CTV	NCTV	Myriad	Ours	Noisy	CTV	NCTV	Myriad	Ours
Barbara	16.31	27.57	27.81	27.70	**30.48**	0.2871	0.7963	0.8027	0.7665	**0.8758**
Boat	16.30	27.25	27.75	27.11	**28.86**	0.2738	0.7789	0.7944	0.7245	**0.8399**
Cameraman	16.33	26.40	26.86	26.86	**29.04**	0.2427	0.8039	0.7858	0.6910	**0.8627**
Couple	16.32	26.88	27.27	26.73	**28.64**	0.2834	0.7567	0.7814	0.7270	**0.8378**
House	16.26	26.78	27.19	26.83	**28.40**	0.3026	0.8022	0.8072	0.7424	**0.8631**
Man	16.31	26.52	26.96	26.50	**27.73**	0.3227	0.7627	0.7899	0.7515	**0.8049**
Mandrill	16.33	24.67	25.23	25.58	**26.27**	0.3416	0.6279	0.7164	0.7175	**0.7513**
Peppers	16.36	28.69	29.00	27.79	**30.49**	0.2629	0.8504	0.8511	0.7419	**0.8866**
Plane	16.31	27.86	28.21	27.67	**29.36**	0.2564	0.8528	0.8250	0.7235	**0.8924**
Synthetic	16.33	36.48	38.15	30.85	**41.39**	0.0765	0.9420	0.9523	0.6187	**0.9836**

Table 3.4: PSNR and SSIM for the noisy and restored images by different methods ($\gamma = 10$).

Ground truth

Noisy image

CTV [56]

NCTV [45]

Myriad filter [37]

Ours

Figure 3.12: Denoising results on image *Barbara* by different methods ($\gamma = 5$).

Ground truth

Noisy image

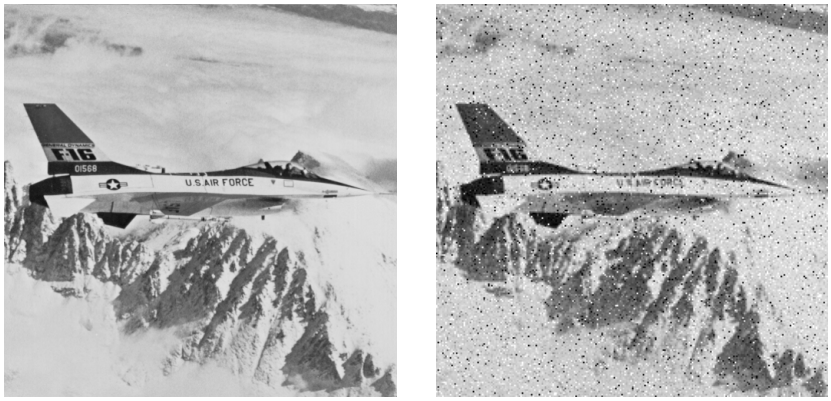
CTV [56]

NCTV [45]

Myriad filter [37]

Ours

Figure 3.13: Denoising results on image *Cameraman* by different methods ($\gamma = 5$).



Ground truth

Noisy image

CTV [56]

NCTV [45]

Myriad filter [37]

Ours

Figure 3.14: Denoising results on image *Plane* by different methods ($\gamma = 5$).

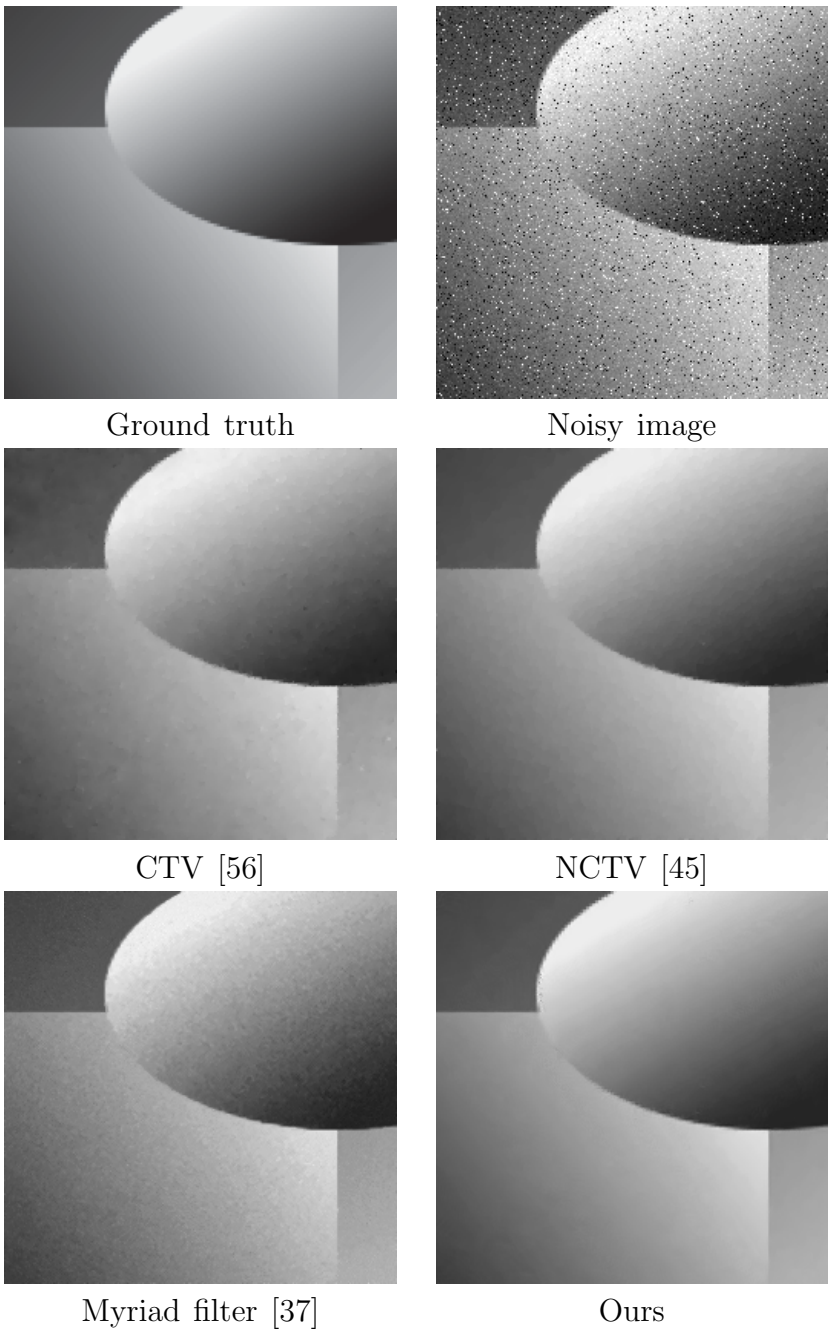


Figure 3.15: Denoising results on image *Synthetic* by different methods ($\gamma = 5$).

Ground truth

Noisy image

CTV [56]

NCTV [45]

Myriad filter [37]

Ours

Figure 3.16: Denoising results on image *Boat* by different methods ($\gamma = 10$).

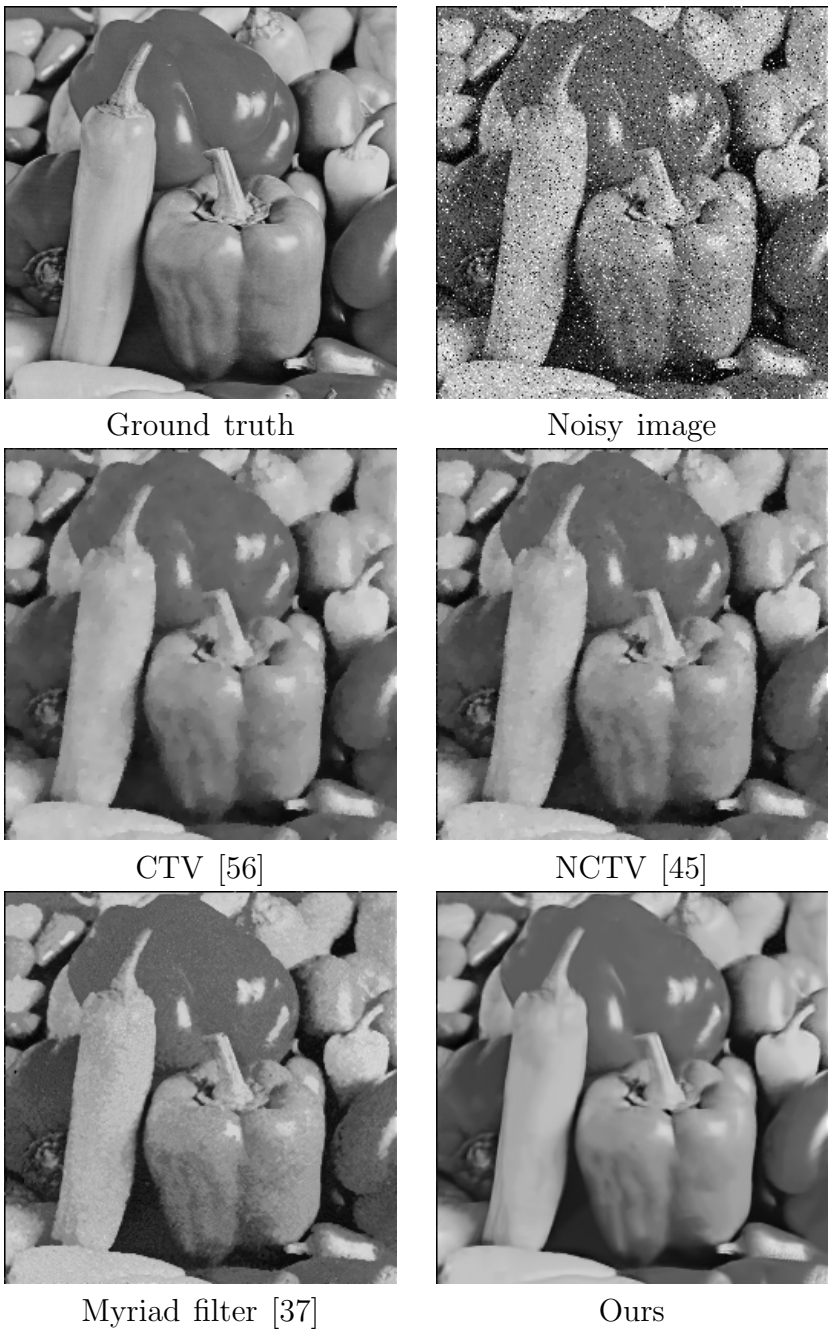


Figure 3.17: Denoising results on image *Peppers* by different methods ($\gamma = 10$).

Chapter 4

Image restoration in underwater

In this chapter, we propose a novel method for restoring an image taken in underwater. As we can see from the section 2.2, there exist various methods for restoring an image corrupted by haze in the atmosphere. In underwater, the attenuation of light depends on the wavelength and it makes a distinction between degraded image formation processes in atmosphere and underwater. Although there are several differences in aspects of image distortion process between air and underwater, we devise a restoration algorithm for underwater based on the methods used for dehazing an image taken in the air.

We adopt the color ellipsoid prior and modify it to utilize for image taken in the underwater. To prevent the artifact by the patch-based method, we extract the local region from the given image using the superpixel segmentation algorithm called SLIC. The color ellipsoid prior is applied only to green and blue channels and information of red channel is combined afterwards. We obtain background light by introducing an index which estimates magnitude and variation of intensity in the region of interest. Numerical experiments show that the proposed algorithm enhances the degraded image visually satisfactorily.

4.1 Scientific background

In underwater, image formation is governed by direct attenuation of light and background light due to scattering like hazy image formation in the atmosphere. However, in underwater, the degree of attenuation is not uniform with respect to wavelength.

The irradiance incident on the camera contains three components which are background light, direct attenuation and forward scattering [70]. Background light is due to the scattering of ambient light, direct attenuation is due to irradiance of non-scattered but reduced light from the object and forward scattering is due to scattered light from the object. Note that the effect of forward scattering is negligible [22] and therefore we can simplify the model.

It is known that the light attenuates exponentially as it travels through water and the transmission of object can be expressed as

$$t_\lambda(x) = \exp\left(-a_\lambda d(x)\right),$$

where a_λ is the attenuation coefficient for wavelength λ and $d(x)$ is a distance from the camera to the scene point x.

We assume that ambient light is approximately homogeneous. Then background light of the distance d is calculated by integrating the contribution of each infinitesimal volume component with respect to distance and angle relative to the line of camera's sight and the direction of incident light. It is in the form of

$$B_\lambda(d) = B_{\lambda,\infty}(1 - \exp\left(-a_\lambda d\right)),$$

where $B_{\lambda,\infty}$ is the background light at infinity.

If we gather the previously mentioned factors contributing to the degradation of image altogether, we obtain the following image formation model

in underwater:

$$I_c(x) = J_c(x)t_c(x) + B_{c,\infty}(1 - t_c(x)), \tag{4.1}$$

where for each color channel $c = R, G, B$ corresponding to suitable wavelength of light, I_c is observed degraded image, J_c is the scene radiance, $B_{c,\infty}$ is the background light at infinity, t_c is the transmission and a_c is the attenuation coefficient.

Since a color image contains red, green and blue channels, we select the wavelength corresponding to color of each channel. We choose 650nm, 525nm and 475nm for wavelengths representing red, green and blue, respectively.

Jerlov categorized waters into five oceanic types and five coastal types based on the transmittance data from near surface water clarity measurements [32]. The attenuation coefficients for selected wavelengths are obtained for all Jerlov water types [57]. Among them, we choose the coefficient of red light corresponding to the oceanic water type I. We can subsequently obtain the coefficients of other color channels by the following relations [70]:

$$\frac{a_c}{a_R} = \frac{B_{R,\infty}(-0.00113\lambda_c + 1.62517)}{B_{c,\infty}(-0.00113\lambda_R + 1.62517)}, \quad c \in \{G, B\}, \tag{4.2}$$

where λ_c represents the wavelength of corresponding color.

Figure 4.1 describes the process of image formation by effects of direct transmission and scattering in underwater.

4.2 Proposed method

In this section, we introduce our novel method for underwater image dehazing. Based on the color ellipsoid prior applied on green and blue channels of an image, we use superpixel segmentation to reduce the artifact arising from the use of patch and, furthermore, utilize value of red channel as an another indicator of depth map of scene.

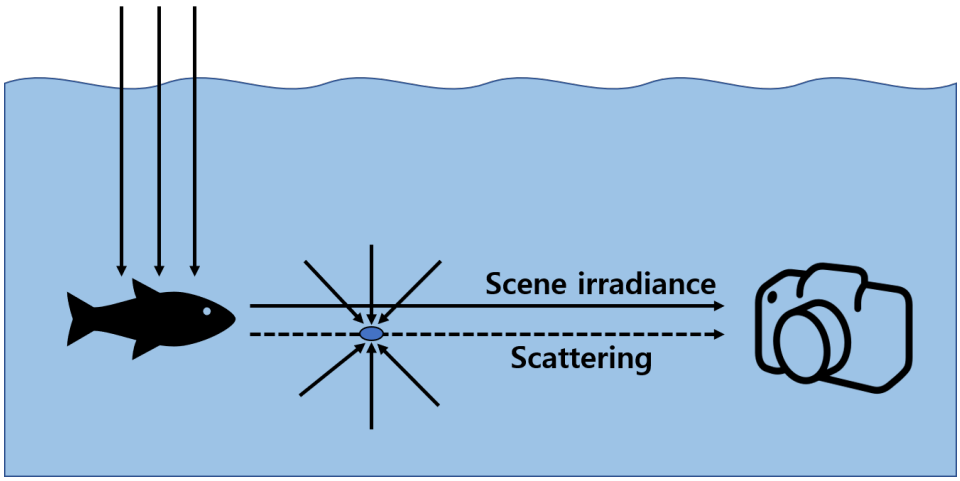
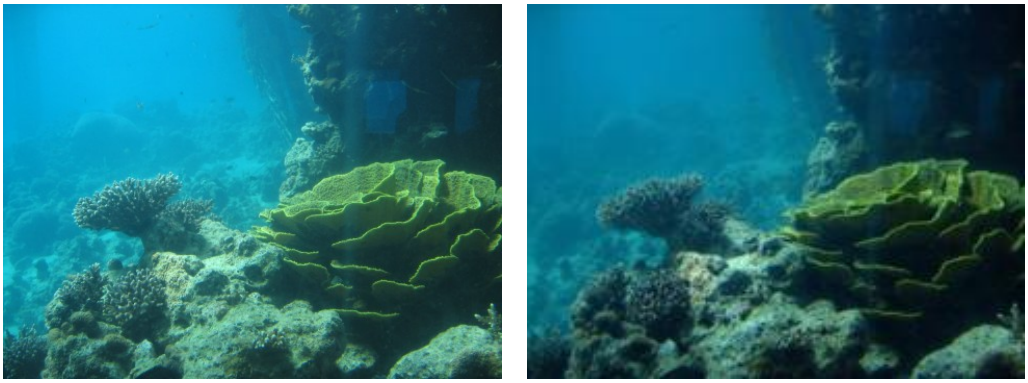


Figure 4.1: Illustration of light transmission and image formation processes in underwater.

4.2.1 Color ellipsoid prior on underwater

One could think of applying the dehazing methods which is useful in the atmosphere such as the dark channel prior or the color ellipsoid prior directly to the restoration problem in underwater. Note that for hazy image in the air, the result of applying the dark channel prior is that if haze is denser between the scene point and the observer, the intensity of the point becomes brighter. But, because red light attenuates rapidly in underwater, the dark channel value of underwater image shows different aspect from that of air such that the farther scene point from the observer becomes darker. Therefore, it is ineffective to directly use the dark channel prior or the color ellipsoid prior to restore an underwater image. Figure 4.2 shows the result of applying the dark channel prior to underwater image without any modifications.

Accordingly, following [22], we only consider green and blue channels of underwater image as an effective indicator of depth information when applying the dark channel based methods. We adopt the color ellipsoid prior only



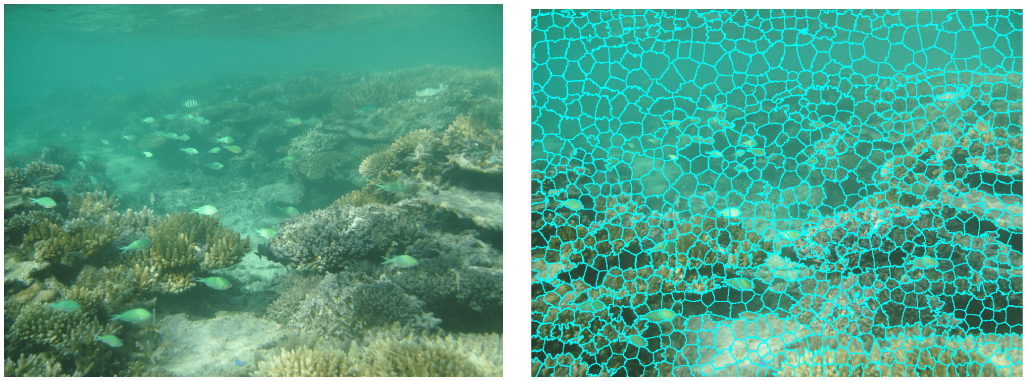
(a) (b)

Figure 4.2: Illustration of ineffectiveness of direct application of the dark channel prior method to the underwater image. (a) Given underwater image, (b) the result of dark channel prior method applied to (a).

on green and blue channels of image and extract depth of the scene.

Superpixel segmentation

In [11], when applying the color ellipsoid prior to the hazy image in the air, additional post-processing is necessary to reduce the so-called halo artifact. The artifact is primarily caused by simultaneous existence of more than two things (for example, branches of a tree and background) in a single patch. It makes difficult to approximate points scattered in RGB space by a single ellipsoid. To overcome mentioned weaknesses, we propose to use the color ellipsoid prior on a segmented region of image, not on a rectangular patch. We adopt one of the superpixel segmentaion methods, SLIC (Simple Linear Iterative Clustering), which is low computational and stable. A single superpixel usually contains similar pixels in which prevents artifact from the use of the color ellipsoid prior on the rectangular patch. Figure 4.3 shows the segmentation result of an underwater image by SLIC algorithm.



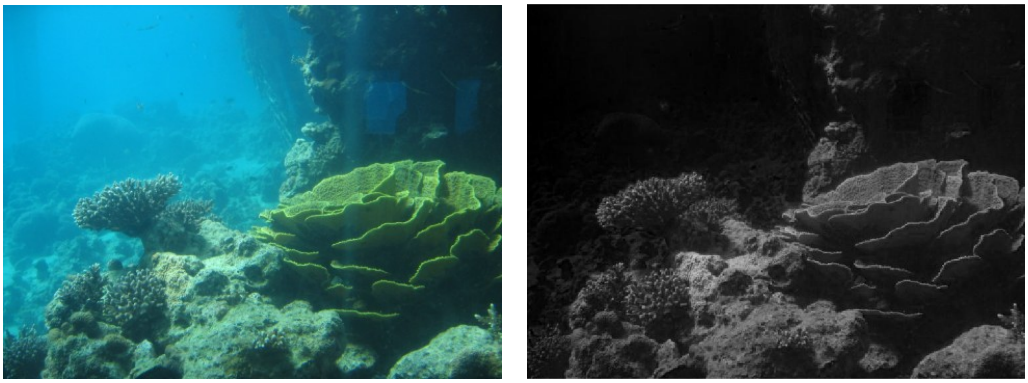
(a) (b)

Figure 4.3: Illustration of application of superpixel segmentation method (SLIC) to the underwater image. (a) Given underwater image, (b) the result of SLIC applied to the image (a).

Combining intensity map of red channel

Although it is not useful for application of dark channel based methods, the intensity map of red channel still contains valuable information. Because visibility of red light relies heavily on the distance from the observer, modest intensity of red channel of a scene point indicates that the point is not far from the observer. On the contrary, almost zero intensity of red channel implies that the point is distant from the observer. Therefore, we utilize the intensity map of red channel of an image by combining it with the normalized depth map obtained from the color ellipsoid prior on the green and blue channels. Fig 4.4 shows the intensity map of red channel of an underwater image. We can see the difference of intensities with respect to the distance from the observer.

For each superpixel Ω, the color ellipsoid prior is applied on (4.1) corresponding to green and blue channels. We assume that the transmission t_c is constant on each superpixel. If we divide both sides of (4.1) by $B_{c,\infty}$ and



(a) (b)

Figure 4.4: Illustration of intensity map of red channel of the underwater image. (a) Given underwater image, (b) Intensity map of red channel of the image (a).

take minimum with respect to pixels in the superpixel and green, blue channels, we have

$$\begin{aligned}\min_{c\in\{g,b\}}\left(\min_{y\in\Omega}\frac{I_c(y)}{B_{c,\infty}}\right) &= \min_{c\in\{g,b\}}\left(t_c\min_{y\in\Omega}\left(\frac{J_c(y)}{B_{c,\infty}}\right)+(1-t_c)\right)\\ &\approx \min_{c\in\{g,b\}}\left(t_c\min_{y\in\Omega}\left(\frac{J_c(y)}{B_{c,\infty}}\right)\right)+\min_{c\in\{g,b\}}(1-t_c)\\ &= \min_{c\in\{g,b\}}(1-t_c).\end{aligned}$$

Therefore,

$$\max_{c\in\{g,b\}} t_c = 1-\min_{c\in\{g,b\}}\left(\min_{y\in\Omega}\frac{I_c(y)}{B_{c,\infty}}\right).$$

By comparing the attenuation coefficients a_G and a_B given beforehand, we can determine whether $t_G > t_B$ or $t_G < t_B$. From the exponential relation between the depth and transmission, we can recover the depth map of an image.

Subsequently, we combine the intensity map I_R of red channel of an image with the acquired depth map. Because they have different scales, we normalize the depth map and denote it by d. Then, we get a refined depth map $\tilde{d}$ as follows:

$$\tilde{d}(x) = d_\infty \cdot \frac{2\mu_R(1 - (\mu_R + \sigma_R)) + d(x)}{2\mu_R + 1},$$

where μ_R, σ_R are average and standard deviation of $I_R(x)$ on a targeted superpixel Ω and d_∞ is a parameter determining a farthest distance from the observer to scene points of the image.

From the refined depth map $\tilde{d}$, we earn the transmission map $t_c = \exp(-a_c\tilde{d})$ so that the scene radiance is obtained from (4.1) as follows:

$$J_c = \frac{I_c - (1 - t_c)B_{c,\infty}}{t_c}. \tag{4.3}$$

4.2.2 Background light estimation

To recover a scene radiance J_c in (4.3), it remains to find a background light $B_{c,\infty}$ from the given intensity map of underwater image. For hazy image taken in the atmosphere, background light is estimated based on the brightest color value [18, 30, 50]. However, because red light is attenuated rapidly in underwater, red channel value of background light is not high. Rather, it is more likely to belong to the darkest part of red channel of image. Also, we predict that both green and blue channel values of background light are high. Furthermore, to estimate a background light precisely, it is better to find a homogeneous region which have no objects other than water. As a result, the background light would be acquired in a homogeneous region where red channel value is low and green, blue channel values are high.

Along with the above assumptions, we obtain background light using the method of [61] based on quad-tree decomposition. It repeats the process of

dividing the interested region into four equal parts and choosing one of them according to certain criterion. Eventually, it reaches a small, homogeneous region whose intensity is chosen as a background light.

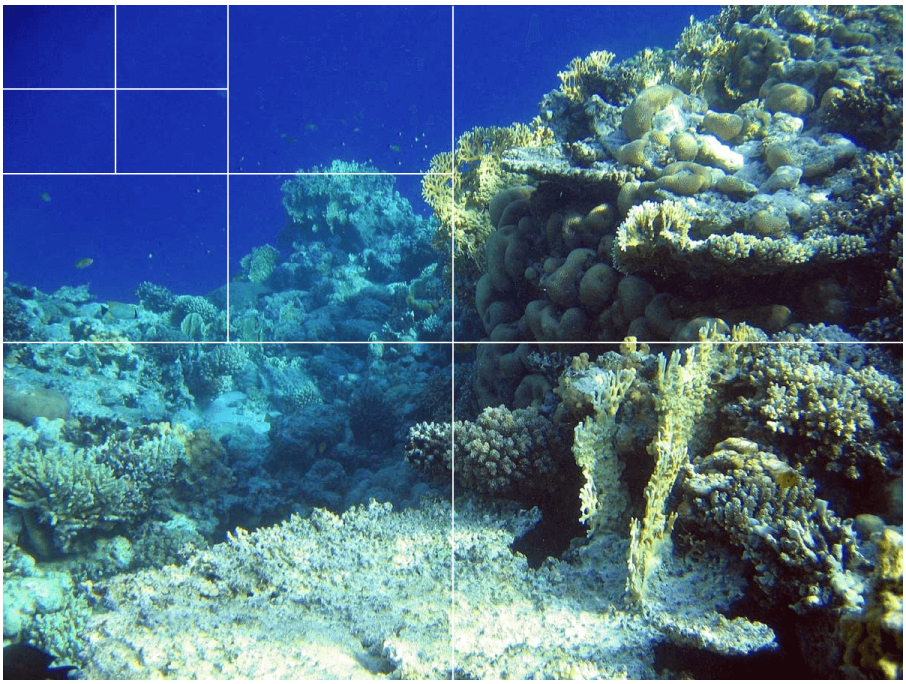


Figure 4.5: Illustration of process of determining background light. Iteratively divide the interested region into four and select one according to the ratio of image intensity and variation.

First, we modify the red channel I_R of an image to $I'_R = \max(I_R) - I_R$ so that a region with low red channel values has high values of I'_R. Then, we define the quantity Q on some region $\mathcal{D}$ which is a criterion for the quadtree decomposition:

$$Q = \frac{\sum_{x \in \mathcal{D}} (I'_R(x) + I_G(x) + I_B(x))}{\sum_{x \in \mathcal{D}} ((\nabla I'_R(x))^2 + (\nabla I_G(x))^2 + (\nabla I_B(x))^2)}. \tag{4.4}$$

Let us define I_0 as original image. We evenly divide I_0 into rectangles

which locate on upper-left, upper-right, lower-left and lower-right of I_0, respectively. We compute Q for each divided region and choose the one with highest Q values. For the chosen region, we repeat the above process. We terminate the process if the difference between maximum and minimum Q values among four divided regions exceeds a certain threshold.

Finally, we get a homogeneous patch and set the pixel as the seed point of background light which has the closest intensity to the average intensity of pixels of the obtained patch in the grayscale. Figure 4.5 depicts an overall process of finding the seed point of the background light.

The overall process of proposed method for restoring a given underwater image is shown in Algorithm 4.

Algorithm 4 Overall restoration process of the proposed method

Given underwater image I_c, $c = R, G, B$
Parameter $a_c > 0$
    Segment the image $I = (I_R, I_G, I_B)$ into superpixels by SLIC.
    For each superpixel, obtain the normalized depth map d by applying the color ellipsoid prior to I_G, I_B.
    Combine d with I_R to get a refined depth map $\tilde{d}$.
    $t_c = \exp(-a_c\tilde{d})$
    Find a background light $B_{c,\infty}$ based on the quad-tree decomposition.
    $J_c = (I_c - (1 - t_c)B_{c,\infty})/t_c$
Output J_c

4.3 Experimental results

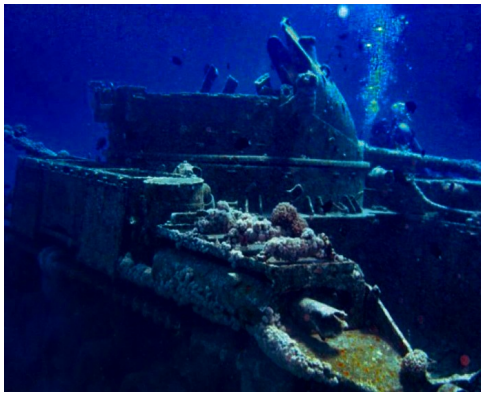
In this section, we present the results of numerical experiments to validate the superior performance of proposed method compared with other methods. Given images taken in underwater, we compare the results of our method with those of UDCP [22], GBRC [43], Nonlocal (NL) haze-line [6]

and IBLA [50]. The attenuation coefficients are critical factor affecting the results of experiments. We adopt the wavelengths of red, green and blue light for our method as 650nm, 525nm and 475nm respectively. Then, we bring attenuation coefficient of red light from [57] and obtain those of blue, green light by (4.2). Note that each restoration method selects their own attenuation coefficients and adjusts its algorithm to them. Therefore, we regard the attenuation coefficients as the parameters set to each method and follow them without any modifications when carrying out numerical experiments.

When we combine intensity map of red channel with depth map from the color ellipsoid prior, we stretch the intensity map of red channel so that its maximum and minimum takes 0 and 1 respectively and apply gamma correction with $\gamma = 0.5$.

In Figures 4.6, 4.7, 4.8 and 4.9, we demonstrate the results of numerical experiments by various restoration methods. We can see that our method performs well in comparison with other methods in terms of visual naturalness and restoration of original color of object. In Figure 4.6, our method restores color of image satisfactorily while in the results of other methods, red color is prevalent or the image becomes too dark. In Figure 4.7, the result of our method is visually natural but dehazing result of remote points is not good enough compared with that of NL haze-line. In Figure 4.8 and 4.9, the results of our method are visually natural in that the overall color of the scene is well restored conforming to color scheme of given image.

Underwater image



UDCP [22]

GBRC [43]

NL haze-line [6]

IBLA [50]

Ours

Figure 4.6: Dehazing results on underwater image by different methods.

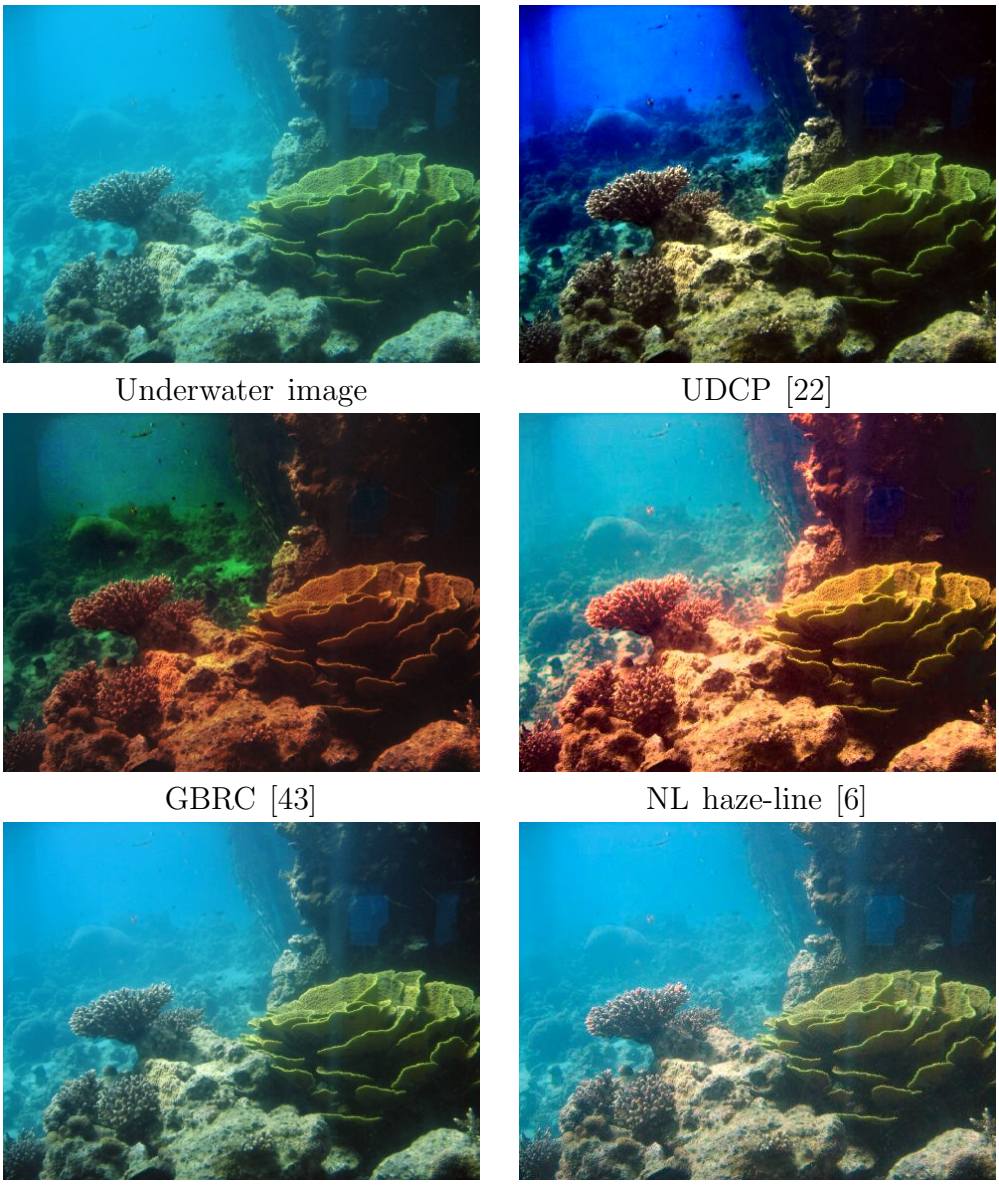


Figure 4.7: Dehazing results on underwater image by different methods.

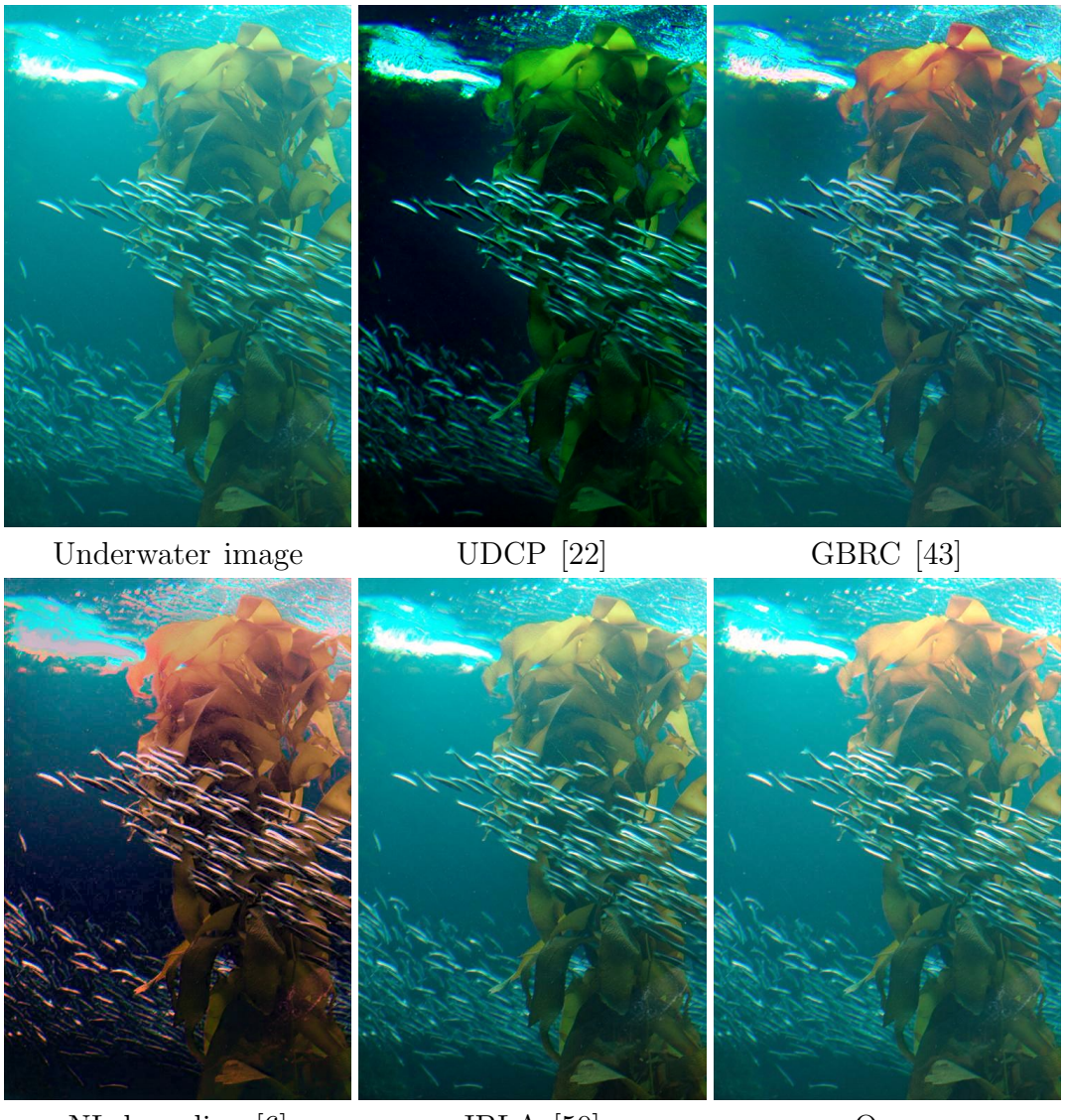


Figure 4.8: Dehazing results on underwater image by different methods.

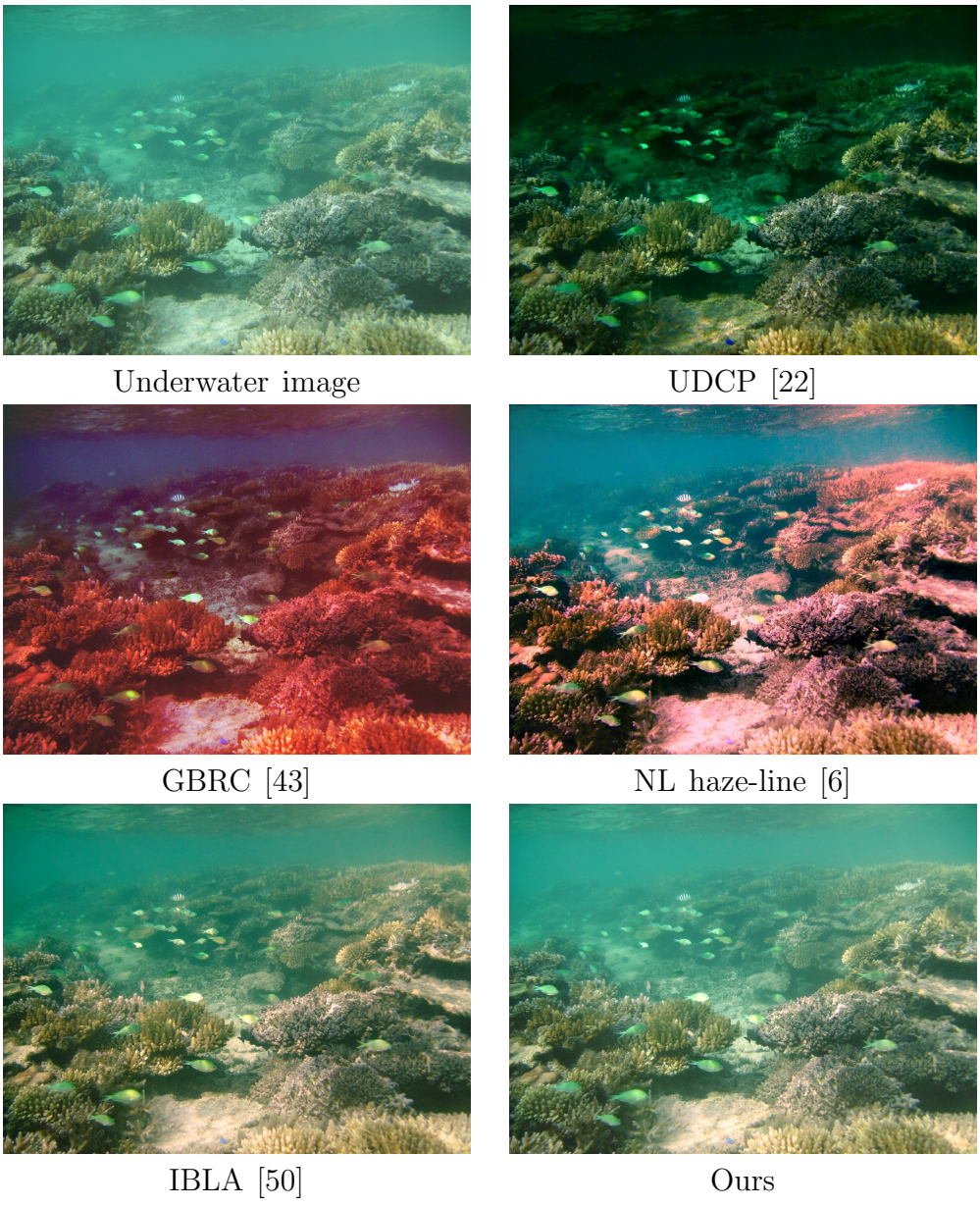


Figure 4.9: Dehazing results on underwater image by different methods.

Chapter 5

Conclusion

In this thesis, we propose regularization methods for restoring image corrupted by Gaussian or Cauchy noise and dehazing of underwater image. In image denoising for the additive Gaussian noise, we suggest the second-order extension of STV based on the fact that STV is a generalization of TV and adaptively takes the directions of highest and smallest intensity variation of neighborhood of a point, but still shows staircase artifact. We propose a hybrid method which combines STV and second-order STV. It improves restoration results by sufficiently smoothing homogeneous regions while preserving edge parts. A computational speed can be further improved if we use effective batched SVD algorithm which parallelizes computation of SVD requiring for every pixel of a given image.

Furthermore, for image corrupted by the additive Cauchy noise, we apply the weighted nuclear norm under nonlocal framework. A variational model is adopted based on MAP estimate, which contains a data fidelity term that is appropriate for noise following Cauchy distribution. Weighted nuclear norm is used as a regularizer in the proposed algorithm, and we utilized similar patches in the image on the basis of nonlocal similarity. We adopt the nonconvex ADMM to solve the problem iteratively and its convergence result is presented.

Subsequently, based on the color ellipsoid prior which is effective for restoring hazy image in the air, we introduce novel dehazing method which is adjusted to underwater situation. Because attenuation coefficient of light in water depends on wavelength of light and red light attenuates rapidly, we apply the color ellipsoid prior only for green and blue channels of a given image and utilize information of red channel for refining the obtained depth map afterwards. To avoid artifact arising from the use of patch based estimation of transmission map, we utilize SLIC, one of the superpixel segmentation algorithms, to proceed the restoration process for each superpixel.

Numerical experiments demonstrate that our proposed methods outperform other compared methods in terms of both quantitative measures such as PSNR, SSIM and qualitative aspect.

Appendix A

Proofs and supplementary material of Section 3.1

A.1 Proof of theorem 3.1.9

Proof. It is basically based on the proof in [39]. We define 2×2 matrix $G_K\mathbf{u}$ and 4×4 matrix $H_K\mathbf{u}$ as follows:

$$G_K\mathbf{u}(x) = \int_\Omega K(\mathbf{t} - x)\mathbf{Ju}(\mathbf{t})^T\mathbf{Ju}(\mathbf{t})d\mathbf{t}$$
$$H_K\mathbf{u}(x) = \int_\Omega K(\mathbf{t} - x)\mathbf{Hu}(\mathbf{t})^T\mathbf{Hu}(\mathbf{t})d\mathbf{t}$$

where $K(\mathbf{t})$ is a nonnegative, rotationally symmetric kernel and $\mathbf{Ju}$, $\mathbf{Hu}$ are Jacobian and Hessian operators described in Section 3.1. They corresponds to evaluating $S_K\mathbf{u}, S_K^{(2)}\mathbf{u}$ at the $x \in \Omega$, respectively and can be also repre-

sented in matrix form as follows:

$$G_K\mathbf{u}(x) = \begin{bmatrix} \int_{\mathbb{R}^2} K(\mathbf{t}-x)\sum_{i=1}^{M}\left(\frac{\partial u_i}{\partial x}(\mathbf{t})\right)^2 d\mathbf{t}, & \int_{\mathbb{R}^2} K(\mathbf{t}-x)\sum_{i=1}^{M}\left(\frac{\partial u_i}{\partial x}(\mathbf{t})\frac{\partial u_i}{\partial y}(\mathbf{t})\right) d\mathbf{t} \\ \int_{\mathbb{R}^2} K(\mathbf{t}-x)\sum_{i=1}^{M}\left(\frac{\partial u_i}{\partial y}(\mathbf{t})\frac{\partial u_i}{\partial x}(\mathbf{t})\right) d\mathbf{t}, & \int_{\mathbb{R}^2} K(\mathbf{t}-x)\sum_{i=1}^{M}\left(\frac{\partial u_i}{\partial y}(\mathbf{t})\right)^2 d\mathbf{t} \end{bmatrix},$$

$$H_K\mathbf{u}(x) = \begin{bmatrix} \int_{\mathbb{R}^2} K(\mathbf{t}-x)\sum_{i=1}^{M}\left(\frac{\partial^2 u_i}{\partial x^2}(\mathbf{t})\right)^2 d\mathbf{t}, & \cdots, & \int_{\mathbb{R}^2} K(\mathbf{t}-x)\sum_{i=1}^{M}\left(\frac{\partial^2 u_i}{\partial x^2}(\mathbf{t})\frac{\partial^2 u_i}{\partial y^2}(\mathbf{t})\right) d\mathbf{t} \\ \vdots & \ddots & \vdots \\ \int_{\mathbb{R}^2} K(\mathbf{t}-x)\sum_{i=1}^{M}\left(\frac{\partial^2 u_i}{\partial y^2}(\mathbf{t})\frac{\partial^2 u_i}{\partial x^2}(\mathbf{t})\right) d\mathbf{t}, & \cdots, & \int_{\mathbb{R}^2} K(\mathbf{t}-x)\sum_{i=1}^{M}\left(\frac{\partial^2 u_i}{\partial y^2}(\mathbf{t})\right)^2 d\mathbf{t} \end{bmatrix}.$$

Now, let us define the operator $A = A(\mathbf{u}, \mathbf{t}, x)$ such that for $\mathbf{v} = (v_1, v_2)^T \in \mathbb{R}^2$,

$$A : \mathbb{R}^2 \to \mathcal{H} = L^2(\mathbb{R}^2, \mathbb{R}^M), \quad A\mathbf{v} = v_1 f_1(\mathbf{t}) + v_2 f_2(\mathbf{t}),$$

where $f_i(\mathbf{t}) = k(\mathbf{t}-x)\Big(\frac{\partial u_1}{\partial x_i}(\mathbf{t}), \cdots, \frac{\partial u_M}{\partial x_i}(\mathbf{t})\Big)^T \in \mathcal{H}$, $(x_1 = x,\, x_2 = y)$, and $k(\mathbf{t}) = \sqrt{K(\mathbf{t})}$.

Similarly, we can define the operator $B = B(\mathbf{u}, \mathbf{t}, x)$ such that for $\mathbf{w} = (w_{11}, w_{12}, w_{21}, w_{22})^T \in \mathbb{R}^4$,

$$B : \mathbb{R}^4 \to \mathcal{I} = L^2(\mathbb{R}^4, \mathbb{R}^M), \quad B\mathbf{w} = \sum_{i,j=1}^{2} w_{ij} g_{ij}(\mathbf{t}),$$

where $g_{ij}(\mathbf{t}) = k(\mathbf{t}-x)\Big(\frac{\partial^2 u_1}{\partial x_i \partial x_j}(\mathbf{t}), \cdots, \frac{\partial^2 u_M}{\partial x_i \partial x_j}(\mathbf{t})\Big)^T \in \mathcal{I}$, $(x_1 = x, x_2 = y)$ and $k(\mathbf{t}) = \sqrt{K(\mathbf{t})}$.

It is noteworthy that $\mathcal{H}, \mathcal{I}$ are Hilbert spaces and they are equipped with

the inner products $\langle \cdot, \cdot \rangle_{\mathcal{H}}, \langle \cdot, \cdot \rangle_{\mathcal{I}}$ such that for $f, f' \in \mathcal{H}$,

$$\langle f, f' \rangle_{\mathcal{H}} = \sum_{i=1}^{M} \int_{\mathbb{R}^2} f_i(\mathbf{t}) f_i'(\mathbf{t}) d\mathbf{t},$$

and $\langle \cdot, \cdot \rangle_{\mathcal{I}}$ is defined similarly.

Then, there exist adjoint operators $A^* : \mathcal{H} \to \mathbb{R}^2$ and $B^* : \mathcal{I} \to \mathbb{R}^4$ which satisfies

$$\begin{aligned} \langle f, A\mathbf{v} \rangle_{\mathcal{H}} &= \langle A^* f, \mathbf{v} \rangle, \text{ for all } \mathbf{v} \in \mathbb{R}^2, f \in \mathcal{H}, \\ \langle g, B\mathbf{w} \rangle_{\mathcal{I}} &= \langle B^* g, \mathbf{w} \rangle \text{ for all } \mathbf{w} \in \mathbb{R}^4, g \in \mathcal{I}. \end{aligned}$$

The above adjoint relations bridge between structure tensor and operators A, B.

$$\begin{aligned} &\langle f, A\mathbf{v} \rangle_{\mathcal{H}} = \langle f, v_1 f_1 + v_2 f_2 \rangle_{\mathcal{H}} = \sum_{i=1}^{2} v_i \langle f, f_i \rangle_{\mathcal{H}} \\ &\Rightarrow A^* f = (\langle f, f_1 \rangle_{\mathcal{H}}, \langle f, f_2 \rangle_{\mathcal{H}})^T \\ &\Rightarrow A^* A\mathbf{v} = A^*(v_1 f_1 + v_2 f_2) = \begin{pmatrix} \langle f_1, f_1 \rangle_{\mathcal{H}} & \langle f_2, f_1 \rangle_{\mathcal{H}} \\ \langle f_1, f_2 \rangle_{\mathcal{H}} & \langle f_2, f_2 \rangle_{\mathcal{H}} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \\ &\Rightarrow A^* A = G_K \mathbf{u} \end{aligned}$$

By similar process, we also obtain that $B^* B = H_K \mathbf{u}$.

Note that both A and B are compact operators and it holds that element-wise square root of vector consisting of eigenvalues of $G_K \mathbf{u}$ (resp., $H_K \mathbf{u}$) is equal to vector which consists of effective singular values of A (resp., B) and whose length is equal to the rank of corresponding operator, which we will denote as $\boldsymbol{\sigma}(A)$ (resp., $\boldsymbol{\sigma}(B)$). Thus, we earn that

$$\mathrm{STV}(\mathbf{u}) = \int_{\Omega} \|\boldsymbol{\sigma}(A(x))\|_1 dx, \quad \mathrm{STV2}(\mathbf{u}) = \int_{\Omega} \|\boldsymbol{\sigma}(B(x))\|_1 dx.$$

The convexity of STV and STV2 would follows if we show the convexity of $\|\boldsymbol{\sigma}(A)\|_1$ and $\|\boldsymbol{\sigma}(B)\|_1$.

For compact operators $A_1, A_2 : \mathbb{R}^2 \to \mathcal{H}$, we have the following dual representation:

$$\|\boldsymbol{\sigma}(A_1)\|_1 = \sup_{\|\boldsymbol{\sigma}(A_2)\|_\infty=1} \operatorname{trace}(A_1 A_2^*).$$

Here, trace of an operator indicates that

$$\operatorname{trace}(A_1 A_2^*) = \sum_{m=1}^{2}\sum_{n=1}^{2} \sigma_m(A_1)\sigma_n(A_2)\langle u_m, \tilde{u}_n\rangle_{\mathcal{H}}\langle v_m, \tilde{v}_n\rangle_2,$$

if the singular decomposition of A_1, A_2 are represented as

$$A_1 = \sum_{m=1}^{2} \sigma_m(A_1)\, u_m \otimes v_m, \quad A_2 = \sum_{n=1}^{2} \sigma_n(A_2)\, \tilde{u}_n \otimes \tilde{v}_n$$

where $u_m, \tilde{u}_n \in \mathcal{H}$ (resp., $v_m, \tilde{v}_n \in \mathbb{R}^2$) are left (resp., right) singular vectors of A_1 and A_2, respectively.

Then for $0 \le t \le 1$, we have

$$\begin{aligned}
\|\boldsymbol{\sigma}(tA_1+(1-t)A_1')\|_1 &= \sup_{\|\boldsymbol{\sigma}(A_2)\|_\infty=1} \operatorname{trace}\Big((tA_1 + (1-t)A_1')A_2^*\Big) \\
&\le \sup_{\|\boldsymbol{\sigma}(A_2)\|_\infty=1} \operatorname{trace}(tA_1A_2^*) + \sup_{\|\boldsymbol{\sigma}(A_2)\|_\infty=1} \operatorname{trace}((1-t)A_1'A_2^*) \\
&= t \sup_{\|\boldsymbol{\sigma}(A_2)\|_\infty=1} \operatorname{trace}(A_1A_2^*) + (1-t) \sup_{\|\boldsymbol{\sigma}(A_2)\|_\infty=1} \operatorname{trace}(A_1'A_2^*) \\
&= t\|\boldsymbol{\sigma}(A_1)\|_1 + (1-t)\|\boldsymbol{\sigma}(A_1')\|_1
\end{aligned}$$

which shows that $\boldsymbol{\sigma}(A)$ is convex. By the same argument, we can show that $\boldsymbol{\sigma}(B)$ is also convex. Therefore, for $\mathbf{u} \in W^{2,2}(\mathbb{R}^2, \mathbb{R}^M)$, STV($\mathbf{u}$) and STV2($\mathbf{u}$) are convex. □

A.2 Finding the gradient step size for Algorithm 1

First, we investigate the adjoint of patch-based operators. From the definition of adjoint, the adjoint of patch-based Jacobian is given by [39]:

$$[\mathbf{J}_K\Omega_1]_k = \sum_{l=1}^{L}\Big[-\operatorname{div}\Big(T^*_{s_l,w}\circ\Omega_1^{((m-1)L+l,:)}\Big)\Big]_n,$$

where $k=(m-1)N+n$ with $1\le n\le N$ and $1\le m\le M$, div is the discrete divergence, $T_{s_l,w}$ is the weighted translation operator on the gradient with translation s_l and weight $w(s_l)=\sqrt{K(s_l)}$ and $\Omega_{1,n}^{(i,:)}$ is the i-th row of $\Omega_{1,n}$. Similarly, the adjoint of patch-based Hessian is given as follows:

$$\begin{aligned}[\mathbf{PH}_K\Omega_2]_k = \sum_{l=1}^{L}\Big(&\Big[\Delta^*_{xx}\circ T^*_{s_l,w}\Big]_n\Omega_{2,n}^{((m-1)L+l,1)} + \Big[\Delta^*_{xy}\circ T^*_{s_l,w}\Big]_n\Omega_{2,n}^{((m-1)L+l,2)}\\ &+\Big[\Delta^*_{yx}\circ T^*_{s_l,w}\Big]_n\Omega_{2,n}^{((m-1)L+l,3)} + \Big[\Delta^*_{yy}\circ T^*_{s_l,w}\Big]_n\Omega_{2,n}^{((m-1)L+l,4)}\Big),\end{aligned}$$

where $T_{s_l,w}$ is the weighted translation operator which acts on the hessian of image.

Now, we are going to show the following proposition:

Proposition A.2.1. Let $d(\Omega_1,\Omega_2)$ be the objective function in (3.9). Then, it is Lipschitz continuous gradient with respect to Ω_1 and Ω_2 with Lipschitz constants $8\tau_1^2$ and $64\tau_2^2$, respectively.

Proof. For $\Omega_1^{'},\Omega_1^{''}\in\mathcal{X}$ with $\Omega_2\in\mathcal{Y}$ fixed, we have

$$\begin{aligned}\|\nabla_{\Omega_1}d(\Omega_1^{'},\Omega_2)-\nabla_{\Omega_1}d(\Omega_1^{''},\Omega_2)\|_{\mathcal{X}} &= \|\tau_1^2\mathbf{J}_K\mathbf{J}_K^*(\Omega_1^{'}-\Omega_1^{''})\|_{\mathcal{X}}\\ &\le \tau_1^2\|\mathbf{J}_K\mathbf{J}_K^*\|\cdot\|\Omega_1^{'}-\Omega_1^{''}\|_{\mathcal{X}}\\ &= \tau_1^2\|\mathbf{J}_K\|^2\|\Omega_1^{'}-\Omega_1^{''}\|_{\mathcal{X}}\end{aligned}$$

Note that $\|\mathbf{J}_K\|^2=\|\mathbf{J}_K^*\mathbf{J}_K\|$ and $\mathbf{J}_K^*\mathbf{J}_K=-\operatorname{div}\circ\sum_{l=1}^{L}(T^*_{s_l,w}\circ T_{s_l,w})\circ\nabla$.

Therefore,

$$\|\mathbf{J}_K^* \mathbf{J}_K \mathbf{u}\|_2 \leq \|\nabla\|^2 \|T\| \|\mathbf{u}\|_2,$$

where $T = \sum_{l=1}^{L} (T_{s_l,w}^* \circ T_{s_l,w})$. Then, we can deduce that $\|\nabla\|^2 \leq 8$ and $\|T\| \leq 1$. It implies that $\|\mathbf{J}_K\|^2 \leq 8$.

Also, the Lipschitz constant for $\nabla_{\Omega_2} d$ follows from the fact that $\|\Delta_{xx}\|$, $\|\Delta_{yy}\|$ and $\|\Delta_{xy}\|$ are all smaller than 4. □

Appendix B

Proofs and supplementary material of Section 3.2

B.1 Proof of theorem 3.2.3

First, let us define the following:

$$\mathcal{F} : \mathbb{R}^{mn} \to \mathbb{R}, \quad \mathcal{F}(X) = \|X\|_{\mathbf{w},*}.$$
$$\mathcal{G} : \mathbb{R}^{mn} \to \mathbb{R}, \quad \mathcal{G}(X) = \frac{\lambda}{2}\langle \log\left(\gamma^2 + (X - Y)^2\right), \mathbf{1}\rangle.$$

We need the following Lemmas:

Lemma B.1.1. The iterates $\{(X^k, V^k, W^k)\}_{k\in\mathbb{N}}$ in Algorithm 2 satisfies:

1. $\mathcal{L}_\beta(X^k, V^k, W^k)$ is lower-bounded and non-increasing for all $k \in \mathbb{N}$.
2. $\{(X^k, V^k, W^k)\}$ is bounded.

Proof. By the definition of X^{k+1}, we can obtain

$$\mathcal{L}_\beta(X^k, V^k, W^k) - \mathcal{L}_\beta(X^{k+1}, V^k, W^k) \geq 0.$$

Furthermore, the following equalities hold

$$\begin{aligned}
&\mathcal{L}_\beta(X^{k+1}, V^k, W^k) - \mathcal{L}_\beta(X^{k+1}, V^{k+1}, W^k) \\
&= \mathcal{G}(V^k) - \mathcal{G}(V^{k+1}) - \langle W^k, V^k - V^{k+1}\rangle - \beta\langle X^{k+1} - V^{k+1}, V^k - V^{k+1}\rangle \\
&\quad + \frac{\beta}{2}\|V^k - V^{k+1}\|^2 \\
&= \mathcal{G}(V^k) - \mathcal{G}(V^{k+1}) - \langle W^{k+1}, V^k - V^{k+1}\rangle + \frac{\beta}{2}\|V^k - V^{k+1}\|^2,
\end{aligned}$$

$$\begin{aligned}
&\mathcal{L}_\beta(X^{k+1}, V^{k+1}, W^k) - \mathcal{L}_\beta(X^{k+1}, V^{k+1}, W^{k+1}) \\
&= \langle W^k - W^{k+1}, X^{k+1} - V^{k+1}\rangle = -\frac{1}{\beta}\|W^k - W^{k+1}\|^2,
\end{aligned}$$

Then, we have the following

$$\begin{aligned}
&\mathcal{L}_\beta(X^{k+1}, V^k, W^k) - \mathcal{L}_\beta(X^{k+1}, V^{k+1}, W^{k+1}) \\
&= \mathcal{G}(V^k) - \mathcal{G}(V^{k+1}) - \langle W^{k+1}, V^k - V^{k+1}\rangle - \frac{1}{\beta}\|W^k - W^{k+1}\|^2 \\
&\quad + \frac{\beta}{2}\|V^k - V^{k+1}\|^2.
\end{aligned} \tag{B.1}$$

Here, from the first-order optimality condition of line 6 in Algorithm 2,

$$\nabla\mathcal{G}(V^k) = W^k, \tag{B.2}$$

and the smoothness of $\mathcal{G}$ implies

$$\|W^{k+1} - W^k\| = \|\nabla\mathcal{G}(V^{k+1}) - \nabla\mathcal{G}(V^k)\| \le L_{\nabla\mathcal{G}}\|V^{k+1} - V^k\|, \tag{B.3}$$

where the Lipschitz constant $L_{\nabla\mathcal{G}} = \dfrac{\lambda}{\gamma^2}$ follows from $\nabla^2\mathcal{G} \le \dfrac{\lambda}{\gamma^2}$.

Then (B.1) becomes

$$\mathcal{G}(V^k) - \mathcal{G}(V^{k+1}) - \langle \nabla\mathcal{G}(V^{k+1}), V^k - V^{k+1}\rangle - \frac{1}{\beta}\|W^k - W^{k+1}\|^2 + \frac{\beta}{2}\|V^k - V^{k+1}\|^2$$
$$\geq \left(-\frac{L_{\nabla\mathcal{G}}}{2} - \frac{L_{\nabla\mathcal{G}}^2}{\beta} + \frac{\beta}{2}\right)\|V^k - V^{k+1}\|^2,$$

from (B.3) and the fact that $\nabla\mathcal{G}$ is Lipschitz continuous.

If we let $C_1 = -\frac{L_{\nabla\mathcal{G}}}{2} - \frac{L_{\nabla\mathcal{G}}^2}{\beta} + \frac{\beta}{2} > 0$ which is equivalent to $\beta > \frac{2\lambda}{\gamma^2}$, we have

$$\mathcal{L}_\beta(X^k, V^k, W^k) - \mathcal{L}_\beta(X^{k+1}, V^{k+1}, W^{k+1}) \geq C_1\|V^k - V^{k+1}\|^2, \tag{B.4}$$

indicating that $\mathcal{L}_\beta(X^k, V^k, W^k)$ is non-increasing for $k \in \mathbb{N}$.

Now considering that $\nabla\mathcal{G}$ is Lipschitz continuous, we have

$$\mathcal{F}(X^k) \geq 0, \quad \mathcal{G}(V^k) \geq \frac{\lambda}{2}\langle \log\gamma^2, \mathbf{1}\rangle,$$

and

$$\begin{aligned}
&\mathcal{G}(V^k) + \langle W^k, X^k - V^k\rangle + \frac{\beta}{2}\|X^k - V^k\|^2 \\
&= \mathcal{G}(V^k) + \langle \nabla\mathcal{G}(V^k), X^k - V^k\rangle + \frac{\beta}{2}\|X^k - V^k\|^2 \\
&\geq \mathcal{G}(X^k) - \frac{L_{\nabla\mathcal{G}}}{2}\|X^k - V^k\|^2 + \frac{\beta}{2}\|X^k - V^k\|^2 \\
&\geq \frac{\lambda}{2}\langle \log\gamma^2, \mathbf{1}\rangle.
\end{aligned}$$

Thus, $\mathcal{L}_\beta(X^k, V^k, W^k)$ is lower-bounded. Because $\mathcal{L}_\beta(X^k, V^k, W^k)$ is non-increasing and $\mathcal{F}$ is coercive, $\{(X^k, V^k)\}$ is bounded. Additionally, from (B.2), it is clear that $\{W^k\}$ is bounded. □

Lemma B.1.2. For all $k \geq 1$, there exists a constant $C_2 > 0$ and $p^{k+1} \in$

$\partial\mathcal{L}_\beta(X^{k+1}, V^{k+1}, W^{k+1})$ such that $\|p^{k+1}\| \leq C_2\|V^k - V^{k+1}\|$.

Proof. Here, we denote the partial of Lagrangian $\partial\mathcal{L}_\beta$ by

$$\partial\mathcal{L}_\beta(X^{k+1}, V^{k+1}, W^{k+1}) = (\partial_X\mathcal{L}_\beta, \nabla_V\mathcal{L}_\beta, \nabla_W\mathcal{L}_\beta)(X^{k+1}, V^{k+1}, W^{k+1}).$$

By direct computation and (B.3),

$$\|\nabla_V\mathcal{L}_\beta(X^{k+1}, V^{k+1}, W^{k+1})\| = \|W^k - W^{k+1}\| \leq L_{\mathcal{G}}\|V^k - V^{k+1}\|,$$
$$\|\nabla_W\mathcal{L}_\beta(X^{k+1}, V^{k+1}, W^{k+1})\| = \frac{1}{\beta}\|W^k - W^{k+1}\| \leq \frac{L_{\mathcal{G}}}{\beta}\|V^k - V^{k+1}\|.$$

Furthermore, observe that

$$\partial_X\mathcal{L}_\beta(X^{k+1}, V^{k+1}, W^{k+1}) = \partial_X\mathcal{F}(X^{k+1}) + W^{k+1} + \beta(X^{k+1} - V^{k+1})$$
$$= \partial_X\mathcal{F}(X^{k+1}) + W^k + \beta(X^{k+1} - V^k) + (W^{k+1} - W^k) + \beta(V^k - V^{k+1}).$$

From the optimality condition of line 5 in Algorithm 2, we have $0 \in \partial_X\mathcal{F}(X^{k+1}) + W^k + \beta(X^{k+1} - V^k)$. Therefore, if we define p^{k+1} as follows:

$$p^{k+1} := \left((W^{k+1} - W^k) + \beta(V^k - V^{k+1}),\ W^k - W^{k+1},\ \frac{1}{\beta}(W^k - W^{k+1})\right),$$

then $p^{k+1} \in \partial\mathcal{L}_\beta(X^{k+1}, V^{k+1}, W^{k+1})$ and

$$\|p^{k+1}\| \leq \left(L_{\mathcal{G}}(2 + \frac{1}{\beta}) + \beta\right)\|V^k - V^{k+1}\|,$$

where $C_2 = L_{\mathcal{G}}(2 + \frac{1}{\beta}) + \beta$. □

Now, we can prove the Theorem 3.2.3.

(Proof of Theorem 3.2.3). Because $\{(X^k, V^k, W^k)\}$ is bounded, there exists a subsequence $\{(X^{k_s}, V^{k_s}, W^{k_s})\}$ which converges to (X^*, V^*, W^*) as $s \to \infty$. Because $\mathcal{L}_\beta(X^k, V^k, W^k)$ is non-increasing and lower-bounded, it converges.

According to (B.4), $\|V^k - V^{k+1}\| \to 0$ as $k \to \infty$. Then, by Lemma B.1.2, there exists a sequence of subdifferentials $p^k \in \partial\mathcal{L}_\beta(X^k, V^k, W^k)$ that satisfies $\|p^k\| \to 0$ as $k \to \infty$. In particular, $\|p^{k_s}\| \to 0$ as $s \to \infty$. Because the roots of a polynomial depend continuously on its coefficients, $\mathcal{F}$ is continuous. This implies the continuity of $\mathcal{L}_\beta$ and $\lim_{s\to\infty} \mathcal{L}_\beta(X^{k_s}, V^{k_s}, W^{k_s}) = \mathcal{L}_\beta(X^*, V^*, W^*)$. Consequently, we have $0 \in \partial\mathcal{L}_\beta(X^*, V^*, W^*)$. □

Chapter B. Proofs and supplementary material of Section 3.2

Bibliography

[1] Radhakrishna Achanta, Appu Shaji, Kevin Smith, Aurelien Lucchi, Pascal Fua, and Sabine Süsstrunk. Slic superpixels compared to state-of-the-art superpixel methods. *IEEE transactions on pattern analysis and machine intelligence*, 34(11):2274–2282, 2012.

[2] Alin Achim and Ercan E Kuruoglu. Image denoising using bivariate α-stable distributions in the complex wavelet domain. *IEEE Signal Processing Letters*, 12(1):17–20, 2004.

[3] Derya Akkaynak and Tali Treibitz. A revised underwater image formation model. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 6723–6732, 2018.

[4] Derya Akkaynak and Tali Treibitz. Sea-thru: A method for removing water from underwater images. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1682–1691, 2019.

[5] Dana Berman, Shai Avidan, et al. Non-local image dehazing. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1674–1682, 2016.

[6] Dana Berman, Deborah Levy, Shai Avidan, and Tali Treibitz. Underwater single image color restoration using haze-lines and a new quantitative dataset. *arXiv preprint arXiv:1811.01343*, 2018.

[7] P. Blomgren and T. F. Chan. Color tv: total variation methods for restoration of vector-valued images. *IEEE Transactions on Image Processing*, 7(3):304–309, March 1998.

[8] Kristian Bredies, Karl Kunisch, and Thomas Pock. Total generalized variation. *SIAM Journal on Imaging Sciences*, 3(3):492–526, 2010.

[9] Kristian Bredies and Tuomo Valkonen. Inverse problems with second-order total generalized variation constraints.

[10] Antoni Buades, Bartomeu Coll, and J-M Morel. A non-local algorithm for image denoising. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, volume 2, pages 60–65. IEEE, 2005.

[11] Trung Minh Bui and Wonha Kim. Single image dehazing using color ellipsoid prior. *IEEE Transactions on Image Processing*, 27(2):999–1009, 2017.

[12] Emmanuel J Candès and Benjamin Recht. Exact matrix completion via convex optimization. *Foundations of Computational mathematics*, 9(6):717, 2009.

[13] Emmanuel J Candes, Justin K Romberg, and Terence Tao. Stable signal recovery from incomplete and inaccurate measurements. *Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences*, 59(8):1207–1223, 2006.

[14] Emmanuel J Candes, Michael B Wakin, and Stephen P Boyd. Enhancing sparsity by reweighted ℓ 1 minimization. *Journal of Fourier analysis and applications*, 14(5-6):877–905, 2008.

[15] Antonin Chambolle, Vicent Caselles, Daniel Cremers, Matteo Novaga, and Thomas Pock. An introduction to total variation for image analy-

sis. *Theoretical foundations and numerical methods for sparse recovery*, 9(263-340):227, 2010.

[16] Tony Chan, Antonio Marquina, and Pep Mulet. High-order total variation-based image restoration. *SIAM Journal on Scientific Computing*, 22(2):503–516, 2000.

[17] Yuh-Chin Chang, Srinivas R Kadaba, Peter C Doerschuk, and Saul B Gelfand. Image restoration using recursive markov random field models driven by cauchy distributed noise. *IEEE Signal Processing letters*, 8(3):65–66, 2001.

[18] John Y Chiang and Ying-Ching Chen. Underwater image enhancement by wavelength compensation and dehazing. *IEEE Transactions on Image Processing*, 21(4):1756–1769, 2011.

[19] Mandar A Chitre, John R Potter, and S-H Ong. Optimal and near-optimal signal detection in snapping shrimp dominated ambient noise. *IEEE Journal of oceanic engineering*, 31(2):497–503, 2006.

[20] JH Curtiss. On the distribution of the quotient of two chance variables. *The Annals of Mathematical Statistics*, 12(4):409–421, 1941.

[21] Kostadin Dabov, Alessandro Foi, Vladimir Katkovnik, and Karen Egiazarian. Image denoising by sparse 3-d transform-domain collaborative filtering. *IEEE Transactions on image processing*, 16(8):2080–2095, 2007.

[22] P. Drews Jr, E. do Nascimento, F. Moraes, S. Botelho, and M. Campos. Transmission estimation in underwater single images. In *2013 IEEE International Conference on Computer Vision Workshops*, pages 825–830, Dec 2013.

[23] Hamza El Ghannudi, Laurent Clavier, Nourddine Azzaoui, François Septier, and Paul-Alain Rolland. α-stable interference modeling and cauchy receiver for an ir-uwb ad hoc network. *IEEE Transactions on Communications*, 58(6):1748–1757, 2010.

[24] Raanan Fattal. Dehazing using color-lines. *ACM transactions on graphics (TOG)*, 34(1):13, 2014.

[25] Willliam Feller. *An introduction to probability theory and its applications*, volume 2. John Wiley & Sons, 2008.

[26] Adrian Galdran, Javier Vazquez-Corral, David Pardo, and Marcelo Bertalmio. Enhanced variational image dehazing. *SIAM Journal on Imaging Sciences*, 8(3):1519–1546, 2015.

[27] Guy Gilboa and Stanley Osher. Nonlocal operators with applications to image processing. *Multiscale Modeling & Simulation*, 7(3):1005–1028, 2008.

[28] Shuhang Gu, Qi Xie, Deyu Meng, Wangmeng Zuo, Xiangchu Feng, and Lei Zhang. Weighted nuclear norm minimization and its applications to low level vision. *International journal of computer vision*, 121(2):183–208, 2017.

[29] Shuhang Gu, Lei Zhang, Wangmeng Zuo, and Xiangchu Feng. Weighted nuclear norm minimization with application to image denoising. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2862–2869, 2014.

[30] Kaiming He, Jian Sun, and Xiaoou Tang. Single image haze removal using dark channel prior. *IEEE transactions on pattern analysis and machine intelligence*, 33(12):2341–2353, 2010.

[31] Moshe Idan and Jason L Speyer. Cauchy estimation for linear scalar systems. *IEEE transactions on automatic control*, 55(6):1329–1342, 2010.

[32] N.G. Jerlov. *Marine Optics*. Number V. 14 in Elsevier oceanography series. Elsevier Scientific Publishing Company, 1976.

[33] Saleem A Kassam. *Signal detection in non-Gaussian noise*. Springer Science & Business Media, 2012.

[34] Claude Knaus and Matthias Zwicker. Dual-domain image denoising. In *2013 IEEE International Conference on Image Processing*, pages 440–444. IEEE, 2013.

[35] Darwin T Kuan, Alexander A Sawchuk, Timothy C Strand, and Pierre Chavel. Adaptive noise smoothing filter for images with signal-dependent noise. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, (2):165–177, 1985.

[36] Ercan E Kuruoglu, William J Fitzgerald, and Peter JW Rayner. Near optimal detection of signals in impulsive noise modeled with a symmetric/spl alpha/-stable distribution. *IEEE Communications Letters*, 2(10):282–284, 1998.

[37] Friederike Laus, Fabien Pierre, and Gabriele Steidl. Nonlocal myriad filters for cauchy noise removal. *Journal of Mathematical Imaging and Vision*, 60(8):1324–1354, 2018.

[38] Jong-Sen Lee. Digital image enhancement and noise filtering by use of local statistics. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, (2):165–168, 1980.

[39] Stamatios Lefkimmiatis, Anastasios Roussos, Petros Maragos, and Michael Unser. Structure tensor total variation. *SIAM Journal on Imaging Sciences*, 8(2):1090–1122, 2015.

[40] Stamatios Lefkimmiatis, John Paul Ward, and Michael Unser. Hessian schatten-norm regularization for linear inverse problems. *IEEE transactions on image processing*, 22(5):1873–1888, 2013.

[41] Jaakko Lehtinen, Jacob Munkberg, Jon Hasselgren, Samuli Laine, Tero Karras, Miika Aittala, and Timo Aila. Noise2Noise: Learning image restoration without clean data. In *Proceedings of the 35th International Conference on Machine Learning*, volume 80 of *Proceedings of Machine Learning Research*, pages 2965–2974, Stockholmsmässan, Stockholm Sweden, 10–15 Jul 2018. PMLR.

[42] C. Li, J. Quo, Y. Pang, S. Chen, and J. Wang. Single underwater image restoration by blue-green channels dehazing and red channel correction. In *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 1731–1735, March 2016.

[43] Chongyi Li, Jichang Quo, Yanwei Pang, Shanji Chen, and Jian Wang. Single underwater image restoration by blue-green channels dehazing and red channel correction. In *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 1731–1735. IEEE, 2016.

[44] Ding Liu, Bihan Wen, Yuchen Fan, Chen Change Loy, and Thomas S Huang. Non-local recurrent network for image restoration. In *Advances in Neural Information Processing Systems*, pages 1680–1689, 2018.

[45] Jin-Jin Mei, Yiqiu Dong, Ting-Zhu Huang, and Wotao Yin. Cauchy noise removal by nonconvex admm with convergence guarantees. *Journal of Scientific Computing*, 74(2):743–766, 2018.

[46] Y. E. NESTEROV. A method for solving the convex programming problem with convergence rate $\mathcal{O}(1/k^2)$. *Dokl. Akad. Nauk SSSR*, 269:543–547, 1983.

[47] John P Nolan. Numerical calculation of stable densities and distribution functions. *Communications in statistics. Stochastic models*, 13(4):759–774, 1997.

[48] Jun Ohta. *Smart CMOS image sensors and applications*. CRC press, 2007.

[49] Konstantinos Papafitsoros and Carola-Bibiane Schönlieb. A combined first and second order variational approach for image reconstruction. *Journal of mathematical imaging and vision*, 48(2):308–338, 2014.

[50] Yan-Tsung Peng and Pamela C Cosman. Underwater image restoration based on image blurriness and light absorption. *IEEE transactions on image processing*, 26(4):1579–1594, 2017.

[51] Yigang Peng, Arvind Ganesh, John Wright, Wenli Xu, and Yi Ma. Rasl: Robust alignment by sparse and low-rank decomposition for linearly correlated images. *IEEE transactions on pattern analysis and machine intelligence*, 34(11):2233–2246, 2012.

[52] Javier Portilla, Vasily Strela, Martin J Wainwright, and Eero P Simoncelli. Image denoising using scale mixtures of gaussians in the wavelet domain. *IEEE Trans Image Processing*, 12(11), 2003.

[53] Benjamin Recht, Maryam Fazel, and Pablo A Parrilo. Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization. *SIAM review*, 52(3):471–501, 2010.

[54] Leonid I Rudin, Stanley Osher, and Emad Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D: nonlinear phenomena*, 60(1-4):259–268, 1992.

[55] Otmar Scherzer, Markus Grasmair, Harald Grossauer, Markus Haltmeier, and Frank Lenzen. *Variational methods in imaging*. Springer, 2009.

[56] Federica Sciacchitano, Yiqiu Dong, and Tieyong Zeng. Variational approach for restoring blurred images with cauchy noise. *SIAM Journal on Imaging Sciences*, 8(3):1894–1922, 2015.

[57] Michael G. Solonenko and Curtis D. Mobley. Inherent optical properties of jerlov water types. *Appl. Opt.*, 54(17):5392–5401, Jun 2015.

[58] GA Tsihrintzis, P Tsakalides, and CL Nikias. Signal detection in severely heavy-tailed radar clutter. In *Conference Record of The Twenty-Ninth Asilomar Conference on Signals, Systems and Computers*, volume 2, pages 865–869. IEEE.

[59] Siva Ram Krishna Vadali, Priyadip Ray, Subrahmanyam Mula, and Pramod K Varshney. Linear detection of a weak signal in additive cauchy noise. *IEEE Transactions on Communications*, 65(3):1061–1076, 2017.

[60] Shenlong Wang, Lei Zhang, and Yan Liang. Nonlocal spectral prior model for low-level vision. In *Asian Conference on Computer Vision*, pages 231–244. Springer, 2012.

[61] Wencheng Wang, Xiaohui Yuan, Xiaojin Wu, and Yunlong Liu. Dehazing for images with large sky region. *Neurocomputing*, 238:365–376, 2017.

[62] Yu Wang, Wotao Yin, and Jinshan Zeng. Global convergence of admm in nonconvex nonsmooth optimization. *Journal of Scientific Computing*, 78(1):29–63, 2019.

[63] Zhou Wang, Alan C Bovik, Hamid R Sheikh, Eero P Simoncelli, et al. Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing*, 13(4):600–612, 2004.

[64] Jun Xu, Lei Zhang, David Zhang, and Xiangchu Feng. Multi-channel weighted nuclear norm minimization for real color image denoising. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 1096–1104, 2017.

[65] Noam Yair and Tomer Michaeli. Multi-scale weighted nuclear norm image restoration. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3165–3174, 2018.

[66] Zhiyuan Zha, Xin Yuan, Bei Li, Xinggan Zhang, Xin Liu, Lan Tang, and Ying-Chang Liang. Analyzing the weighted nuclear norm minimization and nuclear norm minimization based on group sparse representation. *arXiv preprint arXiv:1702.04463*, 2017.

[67] Jianping Zhang and Ke Chen. A total fractional-order variation model for image restoration with nonhomogeneous boundary conditions and its numerical solution. *SIAM Journal on Imaging Sciences*, 8(4):2487–2518, 2015.

[68] Kai Zhang, Wangmeng Zuo, Yunjin Chen, Deyu Meng, and Lei Zhang. Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising. *IEEE Transactions on Image Processing*, 26(7):3142–3155, 2017.

[69] Yulun Zhang, Yapeng Tian, Yu Kong, Bineng Zhong, and Yun Fu. Residual dense network for image restoration. *arXiv preprint arXiv:1812.10477*, 2018.

[70] Xinwei Zhao, Tao Jin, and Song Qu. Deriving inherent optical properties from background color and underwater image enhancement. *Ocean Engineering*, 94:163–172, 2015.

[71] Manfred Zimmermann and Klaus Dostert. Analysis and modeling of impulsive noise in broad-band powerline communications. *IEEE transactions on Electromagnetic compatibility*, 44(1):249–258, 2002.

국문초록

본 논문에서 우리는 가우시안 또는 코시 분포를 따르는 잡음으로 오염된 영상과 물 속에서 얻은 영상을 복원하기 위한 정규화 방법에 대해 논의한다. 영상 잡음 문제에서 우리는 덧셈 가우시안 잡음의 해결을 위해 구조 텐서 총변이의 이차 확장을 도입하고 이것을 이용한 혼합 방법을 제안한다. 나아가 덧셈 코시 잡음 문제를 해결하기 위해 우리는 가중 핵 노름을 비국소적인 틀에서 적용하고 비볼록 교차 승수법을 통해서 반복적으로 문제를 푼다. 이어서 대기 중의 안개 낀 영상을 복원하는데 효과적인 색 타원면 가정에 기초하여, 우리는 물 속의 상황에 알맞은 영상 복원 방법을 제시한다. 물 속에서 빛의 감쇠 정도는 빛의 파장에 따라 달라지기 때문에, 우리는 색 타원면 가정을 영상의 녹색과 청색 채널에 적용하고 그로부터 얻은 깊이 지도를 적색 채널의 강도 지도와 혼합하여 개선된 깊이 지도를 얻는다. 수치적 실험을 통해서 우리가 제시한 방법들을 다른 방법과 비교하고 질적인 측면과 평가 지표에 따른 양적인 측면 모두에서 우수함을 확인한다.

주요어휘: 영상 잡음 제거, 수중 영상 개선, 가우시안 잡음, 코시 잡음, 구조 텐서, 가중 핵 노름, 색 타원면 가정
학번: 2013-22914

감사의 글

우선 저를 넓은 마음으로 지도해 주시고 언제나 지원을 아끼지 않으셨던 강명주 교수님께 감사드립니다. 부족한 부분들을 너그러이 이해해주시고 많은 기회를 제시해주시는 교수님의 포용력이 있었기에 교수님 밑에서 마음껏 연구와 프로젝트를 하면서 소중한 경험들을 쌓을 수 있었습니다. 바쁜 와중에도 흔쾌히 저의 논문 심사를 맡아주시고 논문에 대해 여러 가지 조언을 아끼지 않으신 김판기 교수님, 정미연 교수님, 이병준 교수님, 홍병우 교수님 감사합니다. 홍병우 교수님께는 함께 연구를 하면서 많은 것을 가르쳐 주신 것 또한 감사드립니다. 그리고 제가 대학원에 들어왔을 때, 연구자로서의 기본 자세를 가르쳐 주시고 몸소 열정적인 모습을 보여주신 하승열 교수님, 이미지 세미나를 통해서 영상처리를 알려주시고 연구를 이끌어주신 강명민 교수님 감사합니다.

같은 공간에서 오랫동안 함께한 433호 사람들 그리고 NCIA 연구실 선후배들 감사합니다. 장점이 제각각인 뛰어난 사람들이기 때문에 그들을 통해서 다양하고 깊이 있는 지식을 습득할 수 있었고, 토론을 하면서 함께 논문을 완성시켜 나가는 소중한 경험을 할 수 있었습니다. 또한 일 외적으로도 오랜 시간을 함께 하며 좋은 추억을 많이 쌓을 수 있어서 즐거웠습니다. 대학원 13후기 동기들에게도 고마움을 전하고 싶습니다. 모두 제겐 친구같이 편안한 사람들이고 같이 대학원을 다니면서 의지가 되었습니다. 규칙적인 생활 와중에 가끔씩 만나 스트레스를 풀 수 있었던 시간들은 신선했고, 즐거웠습니다. 그리고 오랜 시간 함께 해준 제 친구들에게 고맙습니다. 각자의 영역에서 열심히 살아가면서 발전하는 모습을 보고 저도 꿋꿋하게 연구를 해나갈 원동력을 얻었습니다. 힘든 순간들이 있을 때마다 힘이 되어주었고 만날 때면 많은 에너지를 얻었습니다.

마지막으로 가족들에게 감사를 전합니다. 부모님이 언제나 저를 굳건하게 믿어주시고 헌신적으로 지원해 주셨기에 포기하지 않고 힘든 과정을 극복할 수 있었고, 제 박사논문이 탄생할 수 있었습니다. 긴 시간동안 항상 옆에서 묵묵하게 응원해준 부모님과 동생에게 고맙고 사랑한다고 전하고 싶습니다.