

Single Retinal Image Restoration

Citation for published version (APA):

Zhang, S. (2023). *Single Retinal Image Restoration*. [Doctoral Thesis, Maastricht University]. Maastricht University. <https://doi.org/10.26481/dis.20231211sz>

Document status and date:

Published: 01/01/2023

DOI:

[10.26481/dis.20231211sz](https://doi.org/10.26481/dis.20231211sz)

Document Version:

Publisher's PDF, also known as Version of record

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

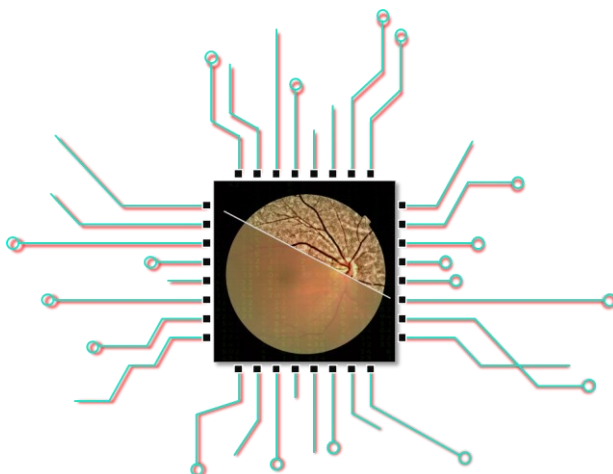
If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

Single Retinal Image Restoration

单幅眼底图像增强算法



Shuhe Zhang

张书赫

The research presented in this thesis was supported by the China Scholarship Council (No: CSC 201908340078).



Copyright ©, Shuhe Zhang, Maastricht University, Maastricht, 2023

All rights are reserved. No part of this book may be reproduced or transmitted in any form or by any means without prior permission in writing by the author.

Layout: Shuhe Zhang
Cover: Shuhe Zhang
Production: Ridderprint, www.ridderprint.nl
ISBN: 978-94-6483-428-4
DOI: <https://doi.org/10.26481/dis.20231211sz>

Single Retinal Image Restoration

DISSERTATION

To obtain the degree of Doctor at the Maastricht University,
on the authority of the Rector Magnificus,

Prof. dr. Pamela Habibović

in accordance with the decision of the Board of Deans,
to be defended in public
on Monday 11 December 2023, at 13:00 hours

by

SHUHE ZHANG

(张书赫)

Born August 13, 1994
in Wuhan China

SUPERVISORS

Prof. dr. Carroll A. B. Webers

Dr. Tos T. J. M. Berendschot

CO- SUPERVISOR

Dr. Remco Duits, Eindhoven University of Technology

ASSESSMENT COMMITTEE

Prof. dr. Ron Heeren (Chair)

Prof. dr. Bart ter Haar Romeny (Eindhoven University of Technology)

Prof. dr. Ton van Leeuwen (Amsterdam University Medical Centre)

Dr. Henry C.A. Woodruff

To my father, Mr. Hongtao Zhang

致我的父亲，张洪涛先生

Index

Chapter 1	General introduction	5
Chapter 2	A reflective background padding algorithm for fundus images preprocessing	41
Chapter 3	DPFR: A double-pass fundus reflection model for efficient single retinal image enhancement <i>Signal processing, (192) 2023, DOI:10.1016/j.sigpro.2021.108400</i>	57
Chapter 4	MUTE: A multilevel-stimulated denoising strategy for single cataractous retinal image dehazing <i>Medical Image Analysis, (88) 2023, DOI: 10.1016/j.media.2023.102848</i>	103
Chapter 5	InQue: An intensity-quenching scheme for large-scale fully unsupervised retinal image enhancement <i>Submitted to Pattern Recognition</i>	201
Chapter 6	BDIC: Boosting the performance of optical microscopy using imaging processing <i>SPIE photonics Europe proceeding 2022, DOI: 10.1117/12.2620781</i>	253
Chapter 7	LR2L: Luminosity rectified blind Richardson-Lucy deconvolution for single retinal image restoration <i>Computer Methods and Programs in Biomedicine (229), 2023, DOI: 10.1016/j.cmpb.2022.107297</i>	275
Chapter 8	Geodesic tracking of retinal vascular trees with optical and TV-flow enhancement in $SE(2)$ <i>International Conference on Scale Space and Variational Methods in Computer Vision (SSVM) 2023, DOI: 10.1007/978-3-031-31975-4_40</i>	339
Chapter 9	General discussions and concluding remarks	361

Table of contents

Appendix

- I** Impacts
- II** Acknowledgments
- III** Curriculum Vitae
- IV** List of publications

Chapter 1

General introduction:

A review of single retinal image restoration techniques & the outline of this thesis

Abstract

Background: Fundus cameras are widely used by ophthalmologists for monitoring and diagnosing retinal pathologies. Unfortunately, no optical system is perfect, and the visibility of retinal images can be greatly degraded due to the presence of problematic illumination, intraocular scattering, and blurriness caused by sudden movements. To improve image quality, different retinal image restoration/enhancement techniques have been developed, which play an important role in improving the performance of various clinical and computer-assistant applications. Due to this, the researchers are attracted to these techniques.

Scopes: This chapter presents a general introduction and review of retinal image restoration/enhancement techniques, and discusses their underlying mathematical models. Three main topics of retinal image restoration/enhancement techniques, i.e. illumination correction, dehazing, and deblurring are addressed. Finally, the aim and outline of this thesis are introduced.

1. The unique duality of eyes between their biological and optical roles

The visual system, as shown in Fig. 1-1 is a complex network of structures and processes that allows humans (and other animals) to perceive and interpret visual information from the surrounding environment. The visual system includes the ocular system, which detects light and sends visual information to the brain, as well as the neural pathways in the brain that process and interpret this information [1]. Within, the ocular system plays a special and unique role due to its duality feature between biological and optical roles.

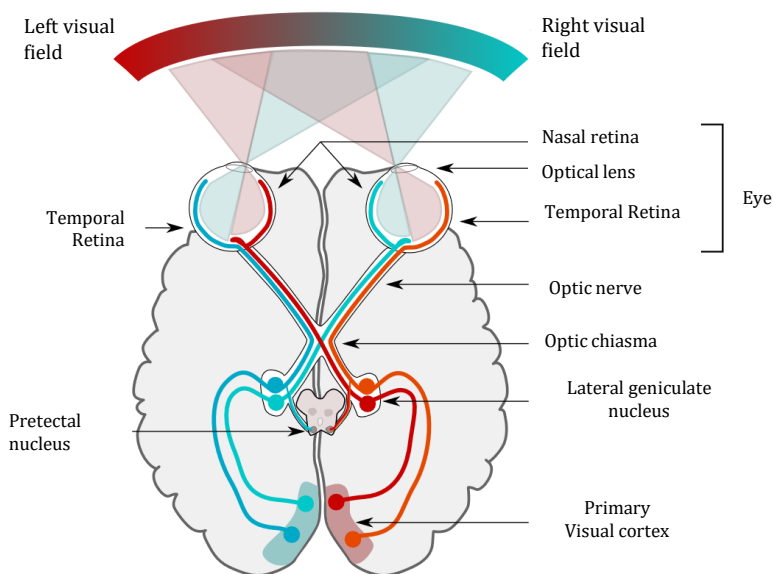


Fig. 1-1. The sketch of the human visual system. The visual system consists of the ocular part (eyes) and neural pathways that go through the thalamus into the visual cortex. The figure is adapted from Wiki.

1.1 From a biological perspective

From a biomedical perspective, the eyes are a complex organ that allows us to perceive the world around us through the detection of light. The eyes are made up of several

interconnected structures as shown in Fig. 1-2, including the cornea, lens, iris, retina, and optic nerve, which work together to focus and transmit visual information to the brain [2].

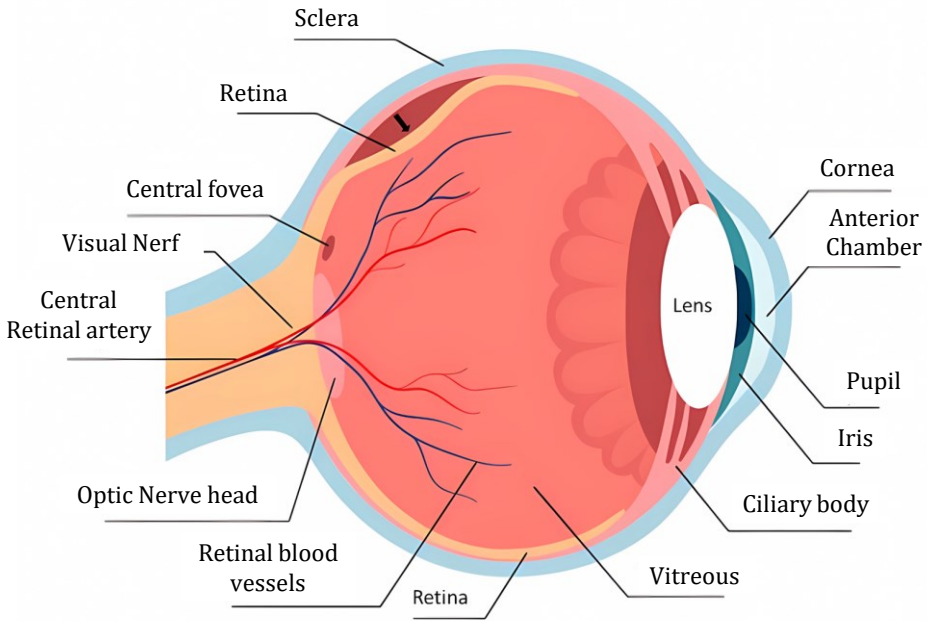


Fig. 1-2. The sketch sagittal plane of the anatomy of the eye. The figure is adapted from Wiki.

Besides vision, the eyes also play a crucial role in regulating the body's circadian rhythms [3], which help to control sleep-wake cycles, hormone secretion, and other physiological processes. The ocular system contains specialized cells called melanopsin-containing retinal ganglion cells, which are sensitive to light and play a critical role in regulating the body's circadian rhythm, or internal biological clock. These cells help to synchronize the body's physiological processes with the 24-hour cycle of the day and are important for maintaining optimal health and well-being.

1.2 From an optical perspective

From an optical perspective, the eyes function as an optical system that focuses light onto the retina [4]. The cornea, lens, and iris work together like a camera's lens system that captures and focuses light to the retina. The amount of incident light can be controlled by the iris. The retina is a layer of tissue located at the back of the eye that contains millions of light-sensitive cells called photoreceptors.

The retina plays a critical role in converting incoming light into electrical signals that can be transmitted to the brain for processing and interpretation, which works like the sensor plane of a camera.

2. Ophthalmoscope

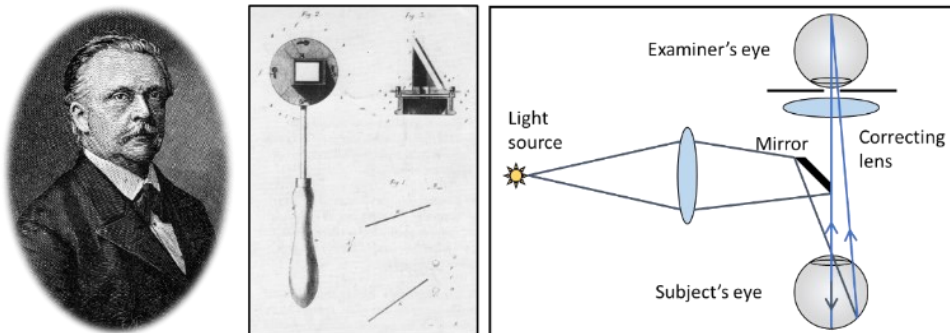


Fig. 1-3. Hermann von Helmholtz, and his handheld ophthalmoscope. Images from left to right are Helmholtz, his ophthalmoscope, and a sketch of its optical path.

The introduction of the ophthalmoscope by Helmholtz [5, 6] allowed for the capture of images of the retina, establishing ophthalmology as a separate sub-area of medicine. In his design as shown in Fig. 1-3, the ophthalmoscope, the subject's eye, and the examiner's eye together form two optical systems: (1) the illumination system, and (2) the detection system. In the illumination system, lights from the light source are reflected from the mirror and incident into the subject's eye. The reflected light from

the retina is then collected and observed by the examiner's eye. This two-system design becomes the fundament of all successor ophthalmoscopes and fundus cameras.

The ophthalmoscope and fundus camera provide a non-invasive method for retinal checking. Since then, retinal images have been widely used by ophthalmologists for early detection, diagnosis, and monitoring of ocular diseases and their progression. Morphologic changes due to eye diseases like diabetic retinopathy [7, 8], glaucoma [9, 10], and age-related macular degeneration [11-13] can be directly observed in these images. Moreover, neurological diseases such as stroke and cognitive dysfunction can be also diagnosed through retinal images [14, 15]. Obviously, the higher the image clarity, the more detailed information can be observed from the image, and the better their diagnostic capabilities.

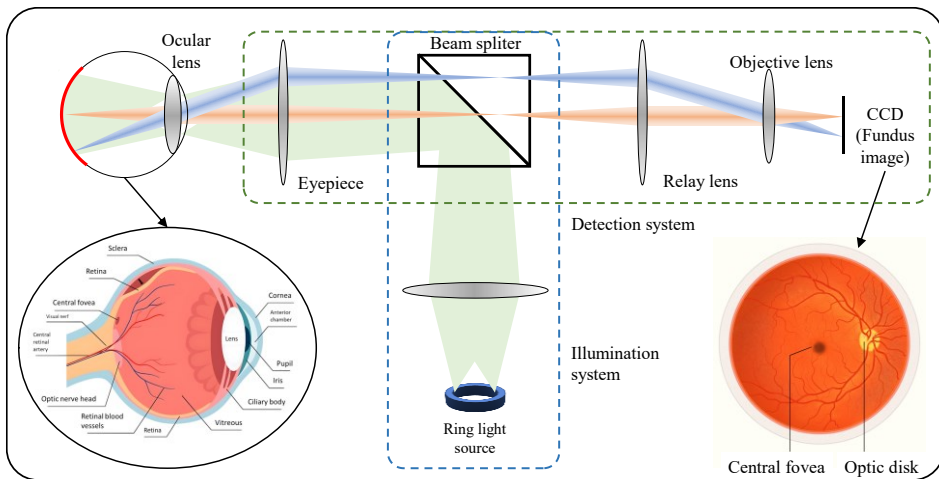


Fig. 1-4. Sketch of the optical design of a fundus camera.

Similar to an ophthalmoscope, a typical retinal imaging platform can be seen as two interconnected imaging systems, as depicted in Fig. 1-4. One is the ocular system, and the other is a reflective imaging system that normally illuminates the fundus through the pupil and collects the reflected light from the retina, forming the image on the camera sensor.

3. Introduction to retinal image enhancement

Retinal images can be severely degraded by opacities in the optical media of cataract eyes [16-18], retinal images for non-cataract subjects can as well be degraded by poor illumination conditions including uneven or insufficient illuminations shown in Fig. 1-5.

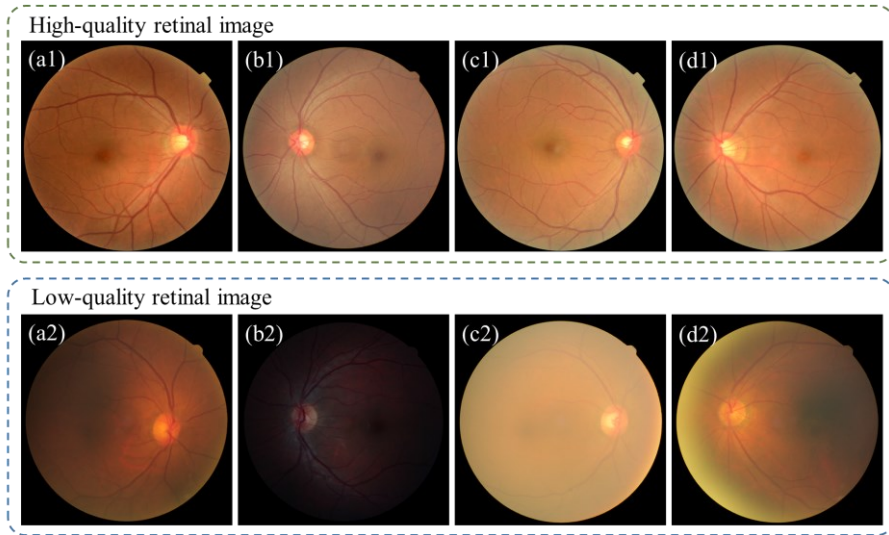


Fig. 1-5. Demonstration of retinal images of good and low quality. (a1) to (d1) are sample images of high quality. (a2) low-quality retinal image with haze and uneven illumination. (b2) insufficient illumination. (c2) haze effect. (d2) uneven illumination and blurriness

Diagnosing efficiency and precision are deeply related to the quality of retinal imaging. However, not every retinal image is perfect, and low-quality image occurrence is not a minor fact. Heaven et al. found 9.5% of all acquired images to be entirely unsatisfactory in a prospective study of 981 diabetic retinopathy patients [19]. Scanlon and Stephen found the ungradable image rate to be between 19.7% for nonmydriatic photography and 3.7% for mydriatic photography study of 3650 diabetic patients [20].

The quality of retinal imaging can be improved either by including high-end fundus cameras and using adaptive optics to tackle the optical aberrations [21], or by using

image enhancement processing to correct for illumination artifacts [22-24], to enhance contrast [25-27] and to use dehazing algorithms [28-30].

3.1 Categories of Retinal Image Enhancement Algorithms

In order to improve image clarity, a number of image enhancement/restoration methods have been proposed, which can be broadly classified into two categories: data domain methods and restored model methods [31] as shown in Fig. 1-6. Data domain methods can be further divided into two families based on their algorithms.

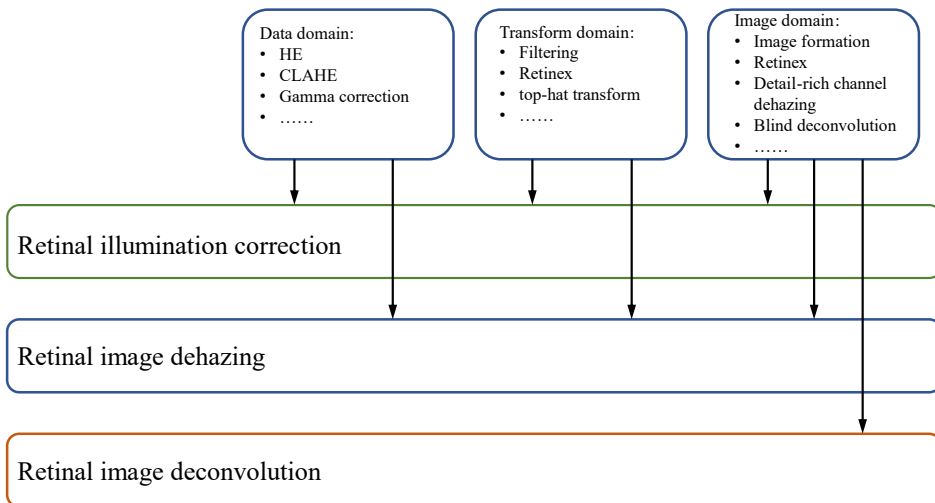


Fig. 1-6. Sketch of Retinal image restoration tasks and their solutions

The first family is known as transform-domain algorithms, which involve transforming a raw image into a new function of other parameters, such as the spatial frequency domain corresponding to the Fourier transform [32], or the structure feature domain corresponding to the top-hat transform [33, 34]. The transformed image is processed and then transformed back into a new image with enhanced contrast. This method allows for global or local modification of the weight for different structures within the image. However, due to its high computational cost, the second family of data domain methods, known as image-domain algorithms, is generally preferred.

The main concept behind the image-domain algorithm is gray-level adjustment. For quick and simple enhancement of retinal images, histogram equalization and its improved version, the contrast limited adaptive histogram equalization (CLAHE), are commonly used [25]. Other histogram modification methods such as q -quantile [35] and gray-scale global spatial entropy method also show promising results in improving the image's contrast [31, 36]. The global gray-level adjustment methods including the gamma map [37] and α -rooting [38] use a fixed function to convert the global gray-level distribution for adjusting the brightness of retina imaging.

Another group of image-domain algorithms uses filters to enhance contrast. These algorithms are similar to transform-domain algorithms but use a convolution kernel to separate the background and foreground information of an image [38, 39]. The foreground information usually corresponds to the detailed structure of an image. By modifying the weights between the background and the foreground, the contrast of the detailed structure can be enhanced. In general, data domain methods belong to pure signal (image) processing techniques that normally take a few considerations of the physical insight of the image formation and enhancement.

3.2 Inspired by Computational Imaging

In order to obtain self-consistent methods for retinal imaging enhancement, restored model methods have been developed. These restored model methods share a similar idea of computational imaging, i.e. a physical model is built to describe the optical process of forming an image under the impact of degeneration agents. By directly or indirectly measuring the optical properties of these degeneration agents, one can compensate for the degeneration agents by digitally mimicking the propagation of the optical wave and modifying the wavefront of light [40-43]. Imaging through scattering media, for example, is a well-known application of computational imaging [42, 43].

Different from computational imaging, the restored models for imaging enhancement do not measure the optical properties of the degeneration agents but try

to find the solutions corresponding to their statistical properties in optical or visual aspects. The solutions can be regarded as roughly estimated versions of those degeneration agents and can be also compensated by applying them to the image formation model, resulting in enhanced images.

Restoration model methods are widely used for imaging de-hazing [44, 45], underwater image enhancement, and night image enhancement [46], while only a few studies have reported their use in retinal imaging enhancement. To our knowledge, the first publication about the application of restored model methods in retinal image enhancement can be traced back to 1989 [16], where the model for imaging the retina in photographs taken through intraocular scatter is considered similar to the model used to represent imaging of the earth from a satellite in the presence of light cloud cover. The scattering effect is removed (or suppressed) by using the Retinex theory.

Based on the image formation model, Xiong *et al.* [31] proposed to use intensity correction and histogram adjustment to preprocess the image, then a transmission map is generated according to the intensity of the preprocessed image in each color channel. The haze effect can be therefore suppressed through dehazing. Although the performance of their approach is good, it relies on statistical and empirical properties of the retina imaging database to determine the algorithm parameters, which makes it hard to apply to different databases. A subsequent study [23] showed the results of using the illumination-reflectance model of image formation to correct the illumination of retinal images. In this research, the color-inversed dark-channel prior, also known as bright-channel prior [47], is employed which shows efficient illumination correction. Mitra *et al.* [48] proposed a model that includes a thin layer of cataracts. Gaudio *et al.* [49] demonstrated a pixel color amplification method for retina imaging enhancement which shows good performance in enhancing the detailed structure of retina images.

In this introductory chapter, we first revisit the mathematical model used for retinal image restoration, their physical/mathematical insight, and how they are related to each other in **Section 4**. We further show how these image formation models are

applied to retinal image restoration in illumination correction in **Section 5**, dehazing in **Section 6**, and deblurring in **Section 7**. A brief introduction to deep-learning-based methods is discussed in **Section 8**. Potential clinical applications that can benefit from retinal image restoration/enhancement are present in **Section 9**. Finally, the aim and outline of this thesis are presented in **Section 10**.

4. Mathematica model for retinal image restoration

4.1 Pixel value stretch model

Enhancing the quality of the retinal image can be achieved by simple manipulation of pixel values. For example, one can enlarge the pixel value if the original value is too small to be noticed or decrease the value of pixels if they are too bright.

Accordingly, the gamma correction, given in Eq. (1-1), provides a simple and straightforward way for pixel adjusting.

$$s_{new}(\mathbf{x}) = s_{old}^{\gamma}(\mathbf{x}), \quad (1-1)$$

When $\gamma < 1$, the nonlinear transform enlarges the small value of pixels, while if $\gamma > 1$, the small value is further suppressed.

$$s_{new}(\mathbf{x}) = HE[s_{new}(\mathbf{x})], \quad (1-2)$$

Another method for pixel adjustment is histogram equalization (HE). When the image is represented by a narrow range of intensity values, the HE is able to make the intensity better distributed among the full dynamic range as shown in Fig. 1-7 (b) and Fig. 1-7 (f). To avoid the over and under-exposure effect of HE, the CLAHE [50] is proposed to adaptively achieve the histogram equalization according to the local contrast in the image's sub-block.

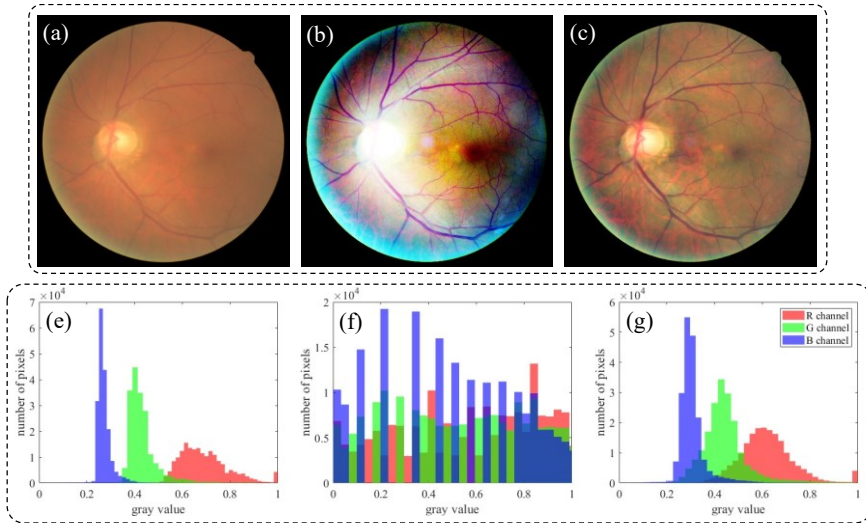


Fig. 1-7. HE and CLAHE enhancement on retinal image. (a) Raw retinal image. (b) Enhanced image using HE. (c) Enhanced image using CLAHE. (e) Histogram for the color channels of the image in (a). Histogram of image in (b). (g) Histogram for image in (c).

The HE and its improved version, the CLAHE, is widely used as a pre-processing method for retinal image enhancement, and the research has shown that the image formation model (IFM) -based methods gain better image restoration results than HE methods.

4.2 Image formation model

The widely used image formation model for retinal image enhancement is the illumination model which is given by

$$s(\mathbf{x}) = l(\mathbf{x}) \cdot r(\mathbf{x}), \quad (1-3)$$

here s is the captured image by the camera, l is the illumination pattern from the light source which is assumed to be spatially slow varying, and r is the retinal reflectance as shown in Fig. 1-8. This model is used for retinal image illumination correction and can be combined with the dehazing model, which we will show later.

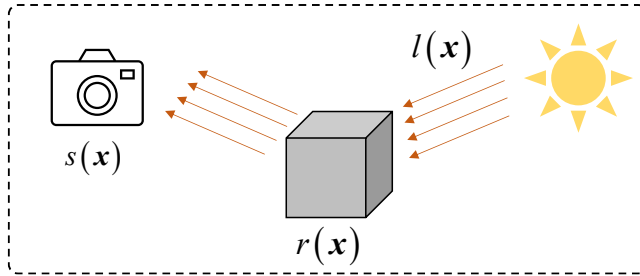


Fig. 1-8. Illumination model.

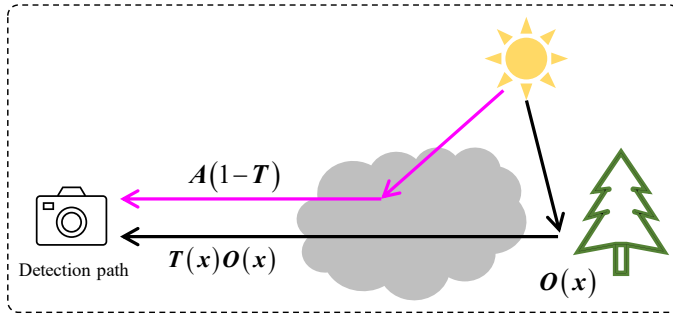


Fig. 1-9. Natural scene haze formation model.

To tackle the haze effect caused by intraocular scattering on retinal imaging, the haze formation model is adopted. The early-stage of the retinal image formation model was directly adopted from Koschmieder & McCartney’s model [51, 52] of hazy nature scenes, as shown in Fig. 1-9, which is given by

$$s(x) = t(x) \cdot o(x) + a[1 - t(x)], \quad (1-4)$$

where o is the haze-free image, t is the transmission matrix of the haze medium describing the portion of the light that is not scattered and reaches the camera. a is the global atmospheric light, and s is the observed image. Here is a large amount of natural scene dehazing research based on Eq (1-4) [53], however, it was developed for natural scenes and may not be the optimal choice for fundus imaging since it ignores the double-pass property of fundus photography.

Peli *et al.* [16] developed an optical model for imaging the retina through cataracts which is

$$s(\mathbf{x}) = \alpha \cdot l \cdot o(\mathbf{x}) + l \cdot [1 - t(\mathbf{x})] \quad (1-5)$$

where l is the flash illumination of the fundus camera and α is the attenuation of retinal illumination due to the cataract. Both l and α are considered as constant as shown in Fig. 1-10.

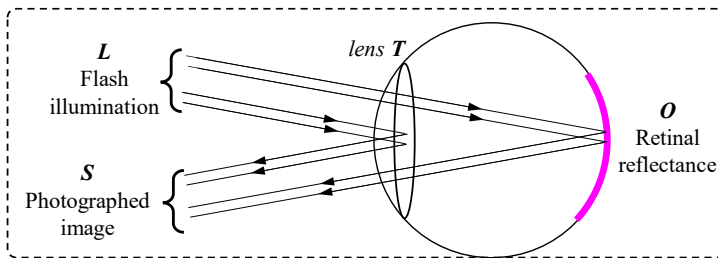


Fig. 1-10. Peli's retinal image formation model.

Different from Eq. (1-4), Eq. (1-5) reveals that the illumination pattern also impacts the quality of retinal imaging. However, as L is constant, Eq. (1-5) loses the ability to correct the uneven (spatially varying) illumination of retinal imaging. In addition, the existing parameter α shows the basic idea of the double-pass property where the illumination light interacts twice with the cataract layer. (When the light goes inside the eye, and when it is reflected out from the fundus).

In [54], an addition-formed retinal image formation model was proposed, which is

$$s(\mathbf{x}) = \text{clip}\left(\alpha \cdot \{[l \otimes G_L(r_L, \sigma_L)](\mathbf{x}) + o(\mathbf{x})\} + \beta; s\right). \quad (1-6)$$

G_L is the Gaussian kernel determined by parameter r_L and σ_L . l is the illumination pattern. clip is a function that truncates the pixel value of the image in a certain dynamic range. β is the additive noise signals.

Although this model didn't consider the optical processing when the image is produced, it shares a similar idea with the image structure model which will be

introduced later in Section 4.3. Based on Eq. (1-6), a Neural network was developed to achieve retinal image restoration. The network was trained based on the pair of high-quality retinal images and degraded images using Eq. (1-6).

4.3 Image structure model

Besides the image formation model, there are also image structure models used for retinal image enhancement [27, 38, 39, 55], and the image structure models can be summarized as

$$s(\mathbf{x}) = s_{background}(\mathbf{x}) + s_{structures}(\mathbf{x}), \quad (1-7)$$

where $s_{background}$ is the background information of the observed image which corresponds to the low-frequency components, while $s_{structures}$ denotes the detailed information implying the detailed structures and textures of the image as shown in Fig. 1-11.

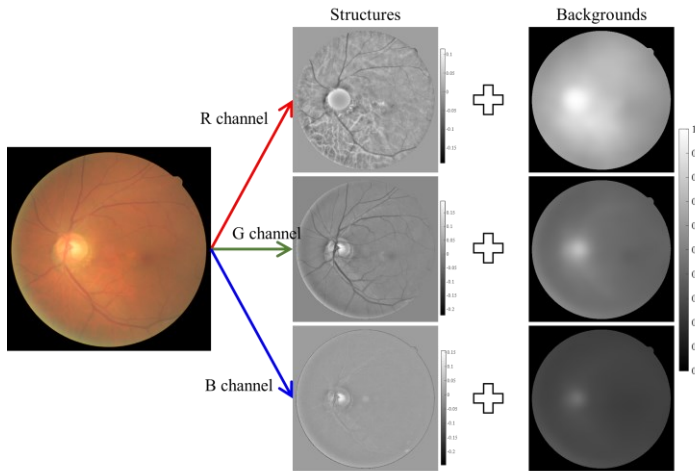


Fig. 1-11. Image structure model. The raw image is decomposed into structured (high-frequency) components and background components.

By giving a large weight to $s_{structures}$ and suppressing the $s_{background}$ one can obtain a contrast-enhanced image. $s_{background}$ can be obtained by low pass filtering of s [38,

55] and total variation regularization [27]. While $s_{structures}$ can be obtained by high pass filtering of s or subtracting $s_{background}$ from s . Note that Eq. (1-7) is not based on the optical process of how the image is formed and the physical insight is different from Eqs. (1-3) to (1-6).

4.4 Retinex theory

It is worth noting that the illumination model Eq (1-3) and image structure model Eq. (1-8) can be unified by the Retinex theory, which was developed first to explain the Land effect from a visual perspective [56]. It was later developed for uneven illumination correction in computer vision. A detailed implementation of Retinex can be categorized into several families including Variational Retinex [57-59], PDE Retinex [60], Threshold Retinex [61], center/surrounded Retinex (known also as filtering-based Retinex) [62], while the filtering-based Retinex gains a lot of research interesting due to its computational efficiency and implementation simplicity. Taking the logarithm to both sides of Eq. (1-3), we obtain

$$\log s(\mathbf{x}) = \log l(\mathbf{x}) + \log r(\mathbf{x}), \quad (1-8)$$

which is identical to the Eq. (1-8). Since $l(\mathbf{x})$ is assumed to be spatially slow-varying. A good estimation of $l(\mathbf{x})$ can be given by low-pass filtering $s(\mathbf{x})$ such that

$$l(\mathbf{x}) \approx \text{LowPassFiltering}(s) = F \otimes s, \quad (1-9)$$

F is the low-passing filter which is known also as the surround function. \otimes denotes the 2D convolution. A Gaussian kernel is a good candidate for such a low-pass filter. According to Eq. (1-8) and Eq. (1-9), $r(\mathbf{x})$ can be given by

$$r = \exp[\log s - \log(F \otimes s)]. \quad (1-10)$$

In the practical implementation, a pixel value normalization should be applied to Eq. (1-10), to avoid the intensity-distorted pixels.

5. Retinal image intensity correction

Retinal image intensity correction is a very basic task for retinal image restoration. Statistical analysis shows that many retinal images are suffering from problematic illumination of certain degrees. Problematic illumination conditions can be due to the imperfect photography skill of ophthalmologists, imperfect head/eye position of subject participants, and problematic illumination path of the fundus camera.

Since human visual assessment of image quality is highly related to the image's brightness, a simple intensity correction on a retinal image can produce great and significant improvement in the image's quality for visual assessment. In this section, we briefly introduce two methods: the gamma correction, and the retinex method for retinal image intensity corrections.

5.1 Intensity correction using Gamma correction

The intensity correction can be simply achieved by Gamma correction if $\gamma < 1$. As shown in Fig. 1-12 (a), when $\gamma = 1/2.2$, a small value, say, 0.218, becomes 0.5 after the Gamma correction. Using this, we can transform the input RGB retinal image shown in Fig. 1-12 (b) to HSV-color space, and then perform gamma correction to its V channel (Value). After that, the image is transformed back to RGB-color space resulting in illumination-corrected images as shown in Fig. 1-12 (e).

By adjusting the value of γ , one can achieve different strengths of illumination correction, while the image contrast is not yet improved significantly.

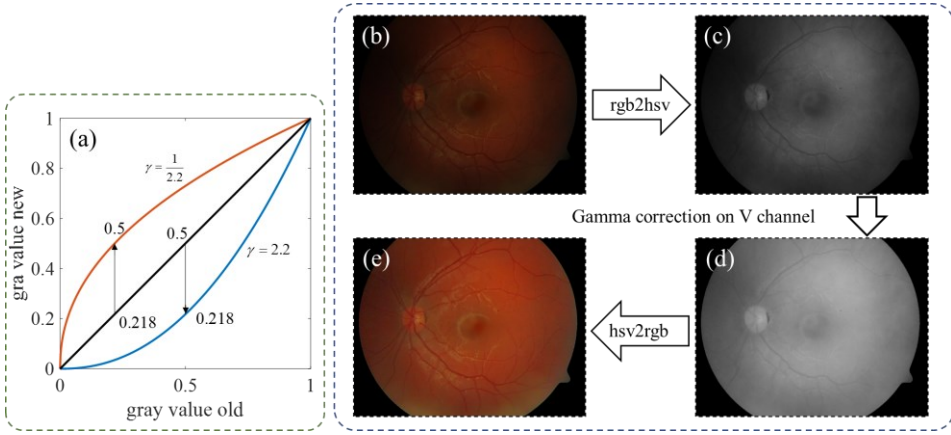


Fig. 1-12. Retinal image intensity correction using Gamma correction.(a) Gamma correction. (b) raw image. (c) V channel of the image. (d) V channel after gamma correction. (e) Enhanced image.

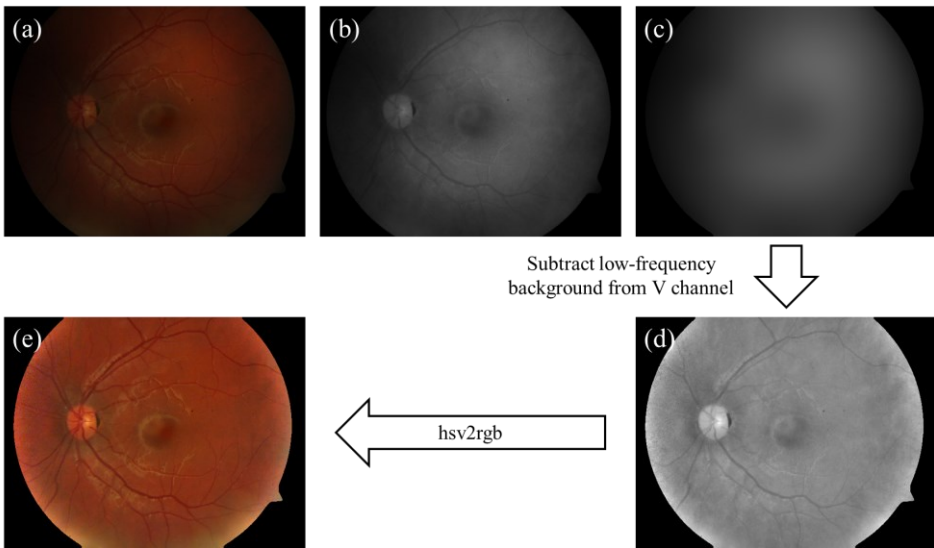


Fig. 1-13. Retinal image intensity correction using Gaussian filtering Retinex method. (a) raw image. (b) V channel. (c) background component of the V channel. (d) retinex correction. (e) Illumination corrected image.

5.2 Intensity correction using Retinex

As mentioned in Section 4.4, the intensity correction can be also achieved using Retinex. The low-frequency component of the V channel can be a good estimation of the illumination pattern as shown in Fig. 1-13 (c). Subtracting the illumination pattern from the original V channel (Fig. 1-13 (d)) and applying the intensity normalization, the output V channel is shown in Fig. 1-13 (d) and in Fig. 1-13 (e) for the RGB image where the uneven pattern is corrected.

Besides the basic Gaussian-filtering Retinex method shown above, researchers have developed a more complex framework, such as variational Retinex of different regularizations and non-local Retinex to achieve illumination correction.

It is also worth noting that the Retinex theory linked image illumination correction and image dehazing through some single algebras. This property will be further discussed in Section 6.2 and Section 6.3.

6. Retinal image dehazing

6.1 Dehazing using dark-channel prior

The dark channel prior (DCP) [63] has been widely used for natural scene dehazing including underwater image enhancement and haze removal even for thick fog situations. Inspired by the visual similarity between the natural fog-induced haze and cataract-induced haze, the dark-channel prior is extended to retinal image restoration. The dark channel of an RGB image is given as

$$\mathbf{J}_{dark}(x) = \min_{c \in \{R, G, B\}} \left\{ \text{Minimum-Filter} \left[\mathbf{J}_c(\rho), w \right] \right\}, \quad (1-11)$$

which first filters the three color channels of the image using a local minimum filter with a size of w pixels, and then calculates the minimum value within the three color channels.

The principle of DCP tells that in any haze-free image (in RGB color space) as shown in Fig. 1-14 (a), at least one pixel has zero intensity in at least one channel as shown in

Fig. 1-14 (c). As such the transmission map of a hazy image (Fig. 1-14 (b)) can be estimated using the dark channel of the image as shown in Fig. 1-14 (d).

$$\text{Dark}\left(\frac{s}{a}\right) = \text{Dark}\left(\frac{t \cdot o}{a}\right) + \text{Dark}(1-t), \text{ where } \text{Dark}\left(\frac{t \cdot o}{a}\right) \approx 0, \quad (1-12)$$

According to the image formation model in Eq. (1-4), the dehazed image can be then calculated.

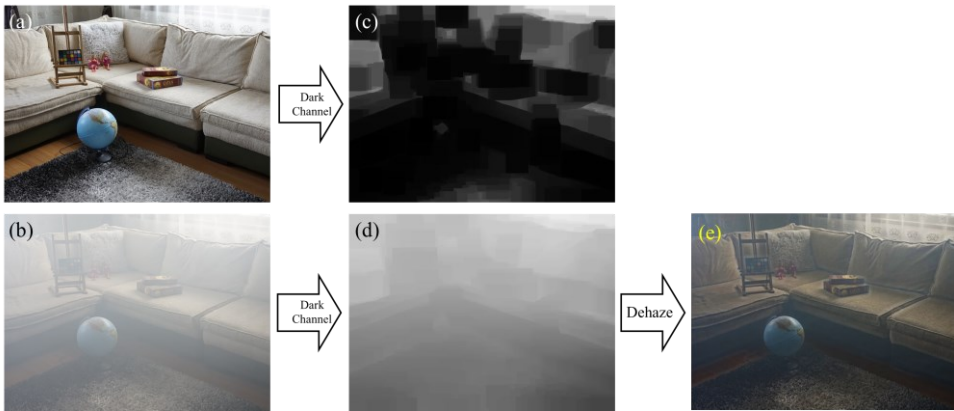


Fig. 1-14. Dehazing using dark-channel prior. (a) Hazy-free image. (b) Hazy image. (c) dark-channel of (a). (d) dark channel of (b). (e) digital dehazing results of (b)

The results of DCP dehazing are promising. However, it is found that the performance of DCP on retinal image dehazing is limited, especially for thick cataracts due to different color statistical features between natural scene images and retinal images. Although the DPC fails to estimate the transmission map of the retinal image in RGB color space, it works in the intensity domain of the image since DCP is valid for gray-scaled image dehazing and the result is also promising [63].

Accordingly, one is able to convert the retinal image from RGB color space to, for example, the CIE-LAB color space, and then perform dehazing to the L-channel (intensity channel). After that, the dehazed retinal image is obtained. Fig. 1-15 shows

the dehazing results on the cataractous retinal image after the illumination correction was applied. The haze effect is significantly suppressed as shown in Fig. 1-15.

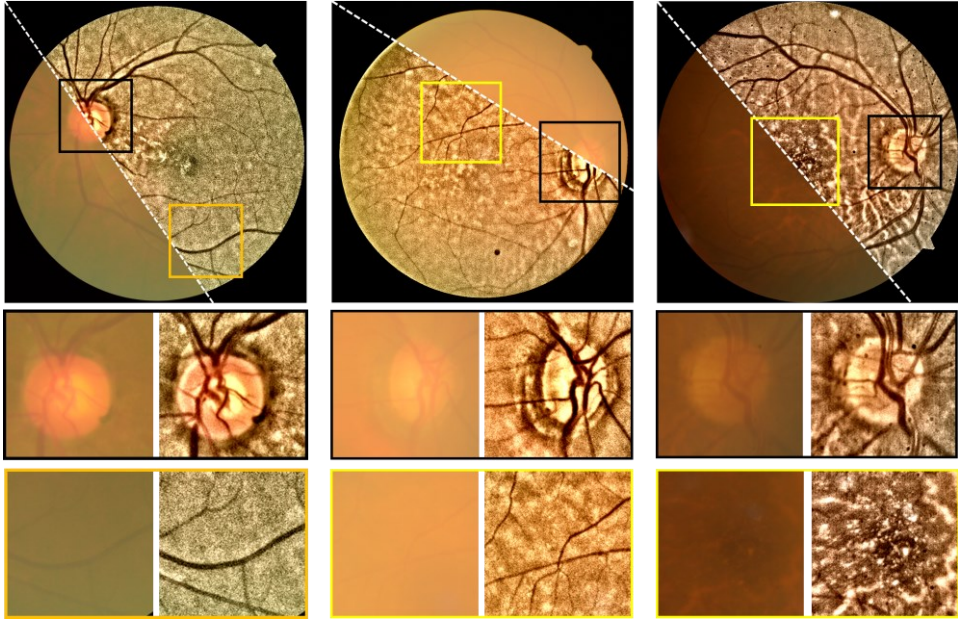


Fig. 1-15. Restoration of cataractous retinal images. First row: raw and enhanced images. Second row: partially zoomed-in images in the optical disk areas. Third row: zoomed-in images in the yellow boxes.

6.2 The duality between intensity correction and dehazing

The nature scene image dehazing seems to be unrelated to intensity correction since they are dealing with different problems. Later, as pointed out in Ref. [64] they are connected by a simple algebra modification of the haze formation model in Eq. (1-4) by assuming the input image is globally white-balanced that is

$$s = t \cdot o + (1 - t), \tag{1-13}$$

$$\Leftrightarrow 1 - s = 1 - t \cdot o - 1 + t \tag{1-14}$$

$$\Leftrightarrow (1 - s) = t(1 - o)$$

If we consider $1 - s = s_{new}$ and $1 - o = r$, Eq. (1-14) is identical to Eq. (1-3). This implies an interesting phenomenon that the color-inversed hazy image looks like an image suffering from insufficient illumination as shown in Fig. 1-16. Using Eq. (1-13) and Eq. (1-14), the dark-channel prior can be used in the color inverted image to enable illumination correction [65], and the color-inversed version of dark-channel prior is also known as bright-channel prior [47].

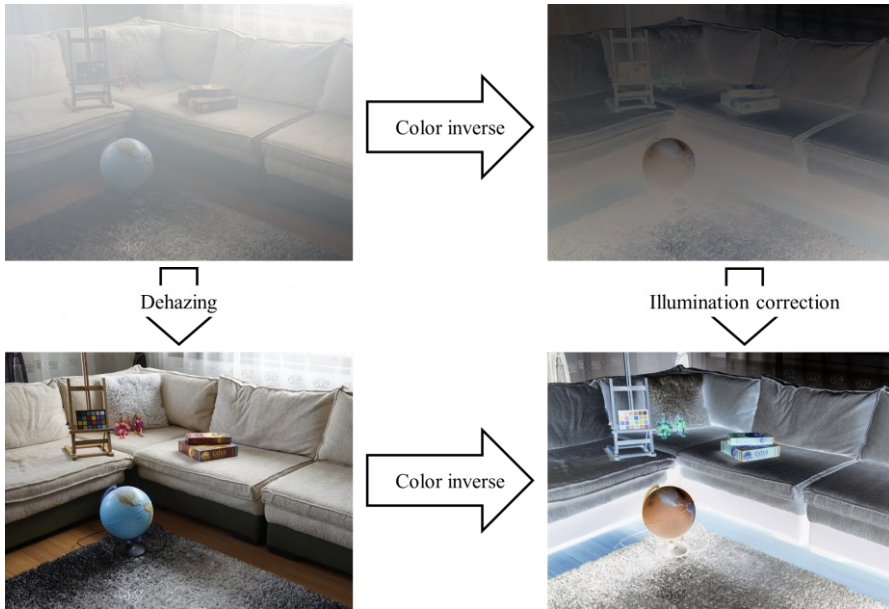


Fig. 1-16. Dehazing task can be converted into an intensity correction task in color-inversed domain.

Further, according to Retinex by assuming t is spatially slow varying, according to Eq. (1-14) and Eq. (1-13), we have

$$o = 1 - \text{Retinex}(1 - s) = \text{Dehazing}(s). \quad (1-15)$$

This formula is also proofed in Ref [64], and shows that the image dehazing task can be finished under the Retinex theory. As such, the Retinex theory is the bridge to image dehazing and image illumination correction [66].

6.3 Dehazing using Retinex

Tab. 1-1. List of publications on non-deep-learning based methods of single retinal image enhancement.

Literature	Image formation model	Key idea	Functions		
			Illumination correction	Contrast enhancement	Dehazing
[16]	Eq. (1-5)	Filtering Retinex	N/A	Yes	Yes
[67]		Image Filtering	Yes	Yes	N/A
[25]		CLAHE	N/A	Yes	N/A
[39]	Eq. (1-7)	Filtering Retinex	Yes	Yes	N/A
[68]	Eq. (1-3), Eq. (1-7)	Filtering Retinex	Yes	Yes	N/A
[23]	Eq. (1-3), Eq. (1-4)	DCP, Retinex	Yes	N/A	N/A
[31]	Eq. (1-4)	HE, Filtering	Yes	Yes	Yes
[69]	Eq. (1-7)	CLAHE	Yes	Yes	N/A
[70]		HE	Yes	Yes	N/A
[48]	Eq. (1-7)	Filtering Retinex, HE	Yes	Yes	Yes
[37]	Eq. (1-3)	Gamma correction, CLAHE	Yes	Yes	N/A
[35]	Eq. (1-3)	Gamma correction, HE	Yes	Yes	N/A
[71]	Eq. (1-3)	HE	Yes	Yes	N/A
[55]	Eq. (1-3), Eq. (1-7)	Filtering, Retinex	Yes	Yes	Yes
[38]	Eq. (1-3), Eq. (1-7)	Filtering, Retinex	Yes	Yes	Yes
[49]	Eq. (1-4)	DCP, extension of DCP	Yes	Yes	Yes
[26]	Eq. (1-4)	Extension of DCP	Yes	Yes	Yes
[27]	Eq. (1-3), Eq. (1-7)	Filtering, Retinex	Yes	Yes	N/A
[72]		Gamma correction, HE	Yes	Yes	N/A
[73]	Eq. (1-3), Eq. (1-7)	Filtering, Retinex	Yes	Yes	N/A

The application of Retinex theory in retinal image dehazing shares a similar idea of image structure model and filtering-based Retinex, where the haze layer is regarded as the slow-varying background component of the retinal image, and the dehazed image can be obtained by subtracting the background component from the hazy one. Tab. 1-

1 lists some of the start-of-the-art publications on non-deep-learning methods of single retinal image enhancement.

7. Retinal image deblurring

Apart from problematic illumination and haze impact, retinal images can be also degraded by optical aberrations such as defocus, and motion blur caused by unintended movement of a subject's head. Such degenerations are modeled by a convolution process between the ideal retinal image and a blur kernel, and cannot be restored using the contrast enhancement methods mentioned above. To tackle the blurriness degenerations, the blind deconvolution method is needed as the kernels are usually unknown.

Image blind deconvolution has been greatly developed and is mainly used for natural scene image deburring [74]. Much prior knowledge, including but not limited to the heavy-tail prior [75, 76], the dark-channel prior [77], and the local maximum gradient prior [78], has been explored to facilitate the single image blind deconvolution tasks. Nevertheless, blind deconvolution for retinal images is still problematic and challenging since there are a large number of retinal images suffering from poor illumination conditions which hide the structure (edge) information that is essential for proper deconvolutions. However, according to our investigation, only a few studies have reported on single retinal image blind deconvolution [79-82], which rather aimed to correct the blurriness caused by aberrations and motions during the retinal imaging.

Andrés et al. proposed a two-step retinal image blind deconvolution method [80], in which the first step is estimating and compensating for the uneven illumination using a fourth-order polynomial. The second step is blind deconvolution with TV-regularization corresponding to the Heavy-tail-prior to natural scene deburring. However, this method requires at least two paired retinal images of one identical subject. Francisco et al. limit the shape of the convolution kernel to a Gaussian shape and perform a line search to determine the size of the Gaussian kernel corresponding to the peak image quality score [81]. This method doesn't correct the illumination

pattern of the retinal image, in addition, not all retinal images are degraded by a simple Gaussian kernel. Therefore, there is a need for methods that can solve two ill-posed problems in one scene, and achieve blind illumination correction and deconvolution simultaneously, which can potentially benefit clinical applications on retinal images.

8. Deep learning methods

With the development of computational power, deep-learning-based retinal image enhancements attracted a lot of research interest. Due to the lack of paired real retinal images for good and degenerated quality, most learning-based retinal image restoration methods published recently can be categorized as extensions of GAN (Generative Adversarial Network). These methods convert the retinal image restoration task into a style-transform task that transforms the image style from a bad-quality retinal image to a good-quality one. To mitigate the risk of GANs introducing unexpected artifacts, many works focus on preserving information fidelity.

Since there are no paired real retinal images, researchers use synthetic/simulated degenerated retinal images to train the networks. For instance, based on the image formation model proposed by Peli et al. [16], Luo et al. [28] trained an unpaired GAN to achieve cataract retinal image dehazing for mild cataract cases. Li et al. [30] proposed an annotation-free GAN for cataractous retinal image restoration. Based on the natural scene haze formation model, Yang et al. [83] trained a modified cycle-GAN for artifact reduction and structure retention in retinal image enhancement. Shen et al. [54] proposed a new mathematical model to formulate the image-degrading process of fundus imaging and train a network for retinal image restoration. Other researchers have modified the structures of the network or loss function to improve the performance of the networks [84, 85].

While these learning-based methods produce impressive restoration results in both quality and naturalness preservation, they have limitations. Overfitting on synthetic data and lack of generalization are potential issues as we will show in Chapter 4 and Chapter 5. Additionally, the performance of trained networks is limited by the input

image resolution (typically 512×512), which is too small for clinical applications where image resolution, in general, is larger than 2000×1000. Furthermore, these methods lack interpretability and may introduce unexpected artifacts or elimination of important retinal structures, which can be detrimental to clinical applications. Thus, there is still a long way to go in both technical and ethical aspects of learning-based retinal image enhancement methods [86]

9. Potential applications of enhanced retinal images

9.1 Retinal image blood vessel segmentations

Retinal image blood vessel segmentation provides the shape distributions of blood vessels, which is important for clinical diagnosis as the morphological changes of blood vessels are biomarkers for diseases such as lacunar stroke [87], cognitive dysfunction [88], cardiovascular risk [89], diabetes [90] and glaucoma [91].

Blood vessel segmentation can be achieved by either human specialists or computer software. The former provides accurate results but is time-consuming. The latter option provides fast segmentation results but is less accurate than human specialists. Moreover, due to poor image contrast of the cataractous retinal image, hand-based segmentation is even more time-consuming, and automatic segmentation for hazy retinal images can be error-prone.

With enhanced retinal image, blood vessel segmentation can be better performed due to the increment of image visual quality.

9.2 Retinal image registration

Image registration is an important application as it provides in the fields of computer vision, pattern recognition, and medical image analysis [92-94]. It aligns two or more retinal images together in the same spatial axis to provide an overall comprehensive understanding [92]. A promising retinal image registration relies on precise feature detecting and matching for images to be registered. Registration of

cataractous retinal images can be failed due to poor feature-pairing results as the features are hidden by the haze effect of cataracts.

With the enhancement of image contrast, the registration algorithm can better find the paired feature for accurate registration.

9.3 Retinopathy diagnosis

Finally, enhancement results in the areas that indicate retinopathy like hard exudates and hemorrhages can also benefit the related diagnosis due to the improvement of the image's quality. For example, as shown in Fig. 1-17, the retinopathy such as exudates can be clearly observed in the enhanced image in Fig. 1-17 (g2).

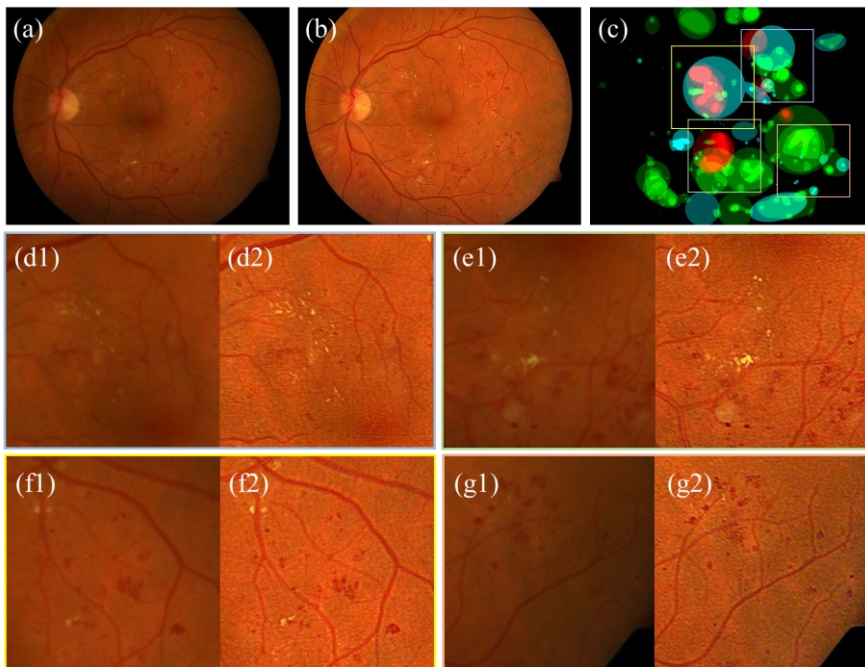


Fig. 1-17. Enhancement of retinopathy areas using LRRL. (a) Raw image. (b) Restored image. (c) Labels of retinopathy areas. Red: Hard exudates; Green: Hemorrhages; Cyan: red small dots. (d1) to (g2) are zoomed-in pictures for regions in blue, green, yellow, and red boxes.

10. Aim and outlines of this thesis

The research described in this thesis aims to develop comprehensible, non-deep learning-based algorithms for single retinal image restoration/enhancement, and investigate their applications in ophthalmology. Such algorithms and their inside principles (each step and the intermedia products) can be fully understood by human specialists, especially ophthalmologists. This is vital for the model to produce reliable and interpretable enhanced images for further clinical applications.

Research questions addressed in this thesis were:

- (1) What is the appropriate model, in both mathematical and optical aspects, that can best describe the image formation process in a fundus image?
- (2) How image illumination correction and image dehazing are combined for single cataract retinal image enhancement/restoration?
- (3) To what degrees can a cataract retinal image be enhanced using a dehazing algorithm?
- (4) Can the proposed model be further optimized? How?
- (5) How illumination correction and blind-deconvolution are combined for retinal image deblur?
- (6) How can the proposed models benefit the community of ophthalmology?

The retinal image consists of two areas: the black background, and the central region of interest (ROI). In order to perform convolutional methods to achieve retinal image enhancement, a background filling should be performed to avoid the boundary effect of convolution operation happening to the edge of the black background and ROI. Such background-filling pre-processing was introduced in **Chapter 2** where the background is filled in a reflective manner.

The *first* question is then addressed in **Chapter 3** where the double-pass fundus reflection (DPFR) model was proposed, inspired by the optical properties of the fundus camera. With the DPFR model, the *second question* is addressed as well.

To further improve the performance of the DPFR model, **Chapter 4** developed a multi-level denoising strategy to significantly increase the image contrast of cataractous retinal images, which partially solves the *third question* and the *fourth question*. Within, the potential clinical applications of the proposed enhancement method are validated based on both simulation and experimental data.

To fully answer the *third question*, we developed a new model in **Chapter 5** to achieve color preservation retinal image enhancement. **Chapter 6** together with **Chapter 7** address the *Fifth question*.

The *sixth* question is addressed in the combination of **Chapter 4**, **Chapter 5**, **Chapter 6**, and **Chapter 7**. Particularly in **Chapter 8** the improvement for retinal image blood vessel tracking using the enhanced image is detailed.

References

1. A. London, I. Benhar, and M. Schwartz, "The retina as a window to the brain—from eye research to CNS disorders," *Nature Reviews Neurology* **9**, 44-53 (2013).
2. R. S. Snell, and M. A. Lemp, *Clinical anatomy of the eye* (John Wiley & Sons, 2013).
3. S. W. Lockley, J. Arendt, and D. J. Skene, "Visual impairment and circadian rhythm disorders," *Dialogues in Clinical Neuroscience* **9**, 301-314 (2007).
4. P. Artal, A. Benito, and J. Taberner, "The human eye is an example of robust optical design," *Journal of vision* **6**, 1-1 (2006).
5. T. T. J. M. Berendschot, P. J. DeLint, and D. v. Norren, "Fundus reflectance—historical and present ideas," *Progress in retinal and eye research* **22**, 171-200 (2003).
6. C. R. Keeler, "The ophthalmoscope in the lifetime of Hermann von Helmholtz," *Archives of Ophthalmology* **120**, 194-201 (2002).
7. D. M. van Reyk, M. C. Gillies, and M. J. Davies, "The retina: oxidative stress and diabetes," *Redox report* **8**, 187-192 (2003).
8. J. Zhang, B. Dashtbozorg, F. Huang, T. T. J. M. Berendschot, and B. M. ter Haar Romeny, "Analysis of Retinal Vascular Biomarkers for Early Detection of Diabetes," in *VipIMAGE 2017*, J. M. R. S. Tavares, and R. M. Natal Jorge, eds. (Springer International Publishing, 2018), pp. 811-817.
9. D. M. S. Barros, J. C. C. Moura, C. R. Freire, A. C. Taleb, R. A. M. Valentim, and P. S. G. Morais, "Machine learning applied to retinal image processing for glaucoma detection: review and perspective," *BioMedical Engineering OnLine* **19**, 20 (2020).
10. R. J. Chalakkal, W. H. Abdulla, and S. C. Hong, "3 - Fundus retinal image analyses for screening and diagnosing diabetic retinopathy, macular edema, and glaucoma disorders," in *Diabetes and Fundus OCT*, A. S. El-Baz, and J. S. Suri, eds. (Elsevier, 2020), pp. 59-111.
11. M. Miura, M. Yamanari, T. Iwasaki, A. E. Elsner, S. Makita, T. Yatagai, and Y. Yasuno, "Imaging Polarimetry in Age-Related Macular Degeneration," *Investigative Ophthalmology Visual Science* **49**, 2661-2667 (2008).
12. M. Trieschmann, F. J. van Kuijk, R. Alexander, P. Hermans, P. Luthert, A. C. Bird, and D. Pauleikhoff, "Macular pigment in the human retina: histological evaluation of localization and distribution," *Eye* **22**, 132-137 (2008).
13. T. Theelen, T. T. J. M. Berendschot, C. B. Hoyng, C. J. Boon, and B. J. Klevering, "Near-infrared reflectance imaging of neovascular age-related macular degeneration," *Graefes Arch Clin Exp.Ophthalmol* **247**, 1625-1633 (2009).
14. T. E. Yap, S. I. Balendra, M. T. Almonte, and M. F. Cordeiro, "Retinal correlates of neurological disorders," *Therapeutic advances in chronic disease* **10**, 2040622319882205 (2019).
15. S. Zafar, J. McCormick, L. Giancardo, S. Saidha, A. Abraham, and R. Channa, "Retinal Imaging for Neurological Diseases: "A Window into the Brain", " **59**, 137-154 (2019).
16. E. Peli, and T. Peli, "Restoration of retinal images obtained through cataracts," *IEEE Transactions on Medical Imaging* **8**, 401-406 (1989).
17. C. W. Yang, X. U. Liang, S. Wang, and H. Yang, "The evaluation of screening for cataract needed surgery with digital nonmydriatic fundus camera," *Ophthalmol. China* **19**, 46-49 (2010).
18. L. Xiong, H. Li, and L. Xu, "An Approach to Evaluate Blurriness in Retinal Images with Vitreous Opacity for Cataract Diagnosis," *Journal of Healthcare Engineering* **2017** 1-16 (2017).
19. C. J. Heaven, J. Cansfield, and K. M. Shaw, "The quality of photographs produced by the non-mydratic fundus camera in a screening programme for diabetic retinopathy: A 1 year prospective study," *Eye* **7**, 787-790 (1993).
20. P. H. Scanlon, C. Foy, R. Malhotra, and S. J. J. D. C. Aldington, "The influence of age, duration of diabetes, cataract, and pupil size on image quality in digital photographic retinal screening," **28**, 2448-2453 (2005).
21. S. A. Burns, A. E. Elsner, K. A. Sapoznik, R. L. Warner, and T. J. Gast, "Adaptive optics imaging of the human retina," *Prog Retin Eye Res* **68**, 1-30 (2019).

22. M. Foracchia, E. Grisan, and A. Ruggeri, "Luminosity and contrast normalization in retinal images," *Medical Image Analysis* **9**, 179-190 (2005).
23. B. Savelli, A. Bria, A. Galdran, C. Marrocco, M. Molinara, A. e. Campilho, and F. Tortorella, "Illumination correction by dehazing for retinal vessel segmentation," *IEEE 30th International Symposium on Computer-Based Medical Systems* (2017).
24. Y. Ma, J. Liu, Y. Liu, H. Fu, Y. Hu, J. Cheng, H. Qi, Y. Wu, J. Zhang, and Y. Zhao, "Structure and illumination constrained GAN for medical image enhancement," *IEEE Transactions on Medical Imaging* **40**, 3955-3967 (2021).
25. A. W. Setiawan, T. R. Mengko, O. S. Santoso, and A. B. Suksmono, "Color Retinal Image Enhancement using CLAHE," *International Conference on ICT for Smart Society* (2013).
26. L. Cao, and H. Li, "Detail-richest-channel based enhancement for retinal image and beyond," *Biomedical Signal Processing and Control* **69**, 102933 (2021).
27. J. Wang, Y.-J. Li, and K.-F. Yang, "Retinal fundus image enhancement with image decomposition and visual adaptation," *Computers in Biology and Medicine* **128**, 104116 (2021).
28. Y. Luo, K. Chen, L. Liu, J. Liu, J. Mao, G. Ke, and M. Sun, "Dehaze of Cataractous Retinal Images using an Unpaired Generative Adversarial Network," *IEEE Journal of Biomedical and Health Informatics* (2020).
29. S. Zhang, C. A. B. Webers, and T. T. J. M. Berendschot, "A double-pass fundus reflection model for efficient single retinal image enhancement," *Signal Processing* **192**, 108400 (2022).
30. H. Li, H. Liu, Y. Hu, H. Fu, Y. Zhao, H. Miao, and J. Liu, "An annotation-free restoration network for cataractous fundus images," *IEEE Transactions on Medical Imaging* (2022).
31. L. Xiong, H. Li, and L. Xu, "An enhancement method for color retinal images based on image formation model," *Computer Methods and Programs in Biomedicine* **143**, 137-150 (2017).
32. J. Mukherjee, and S. K. Mitra, "Enhancement of Color Images by Scaling the DCT Coefficients," *IEEE Transactions on Image Processing* **17**, 1783-1794 (2008).
33. X. Bai, F. Zhou, and B. Xue, "Image enhancement using multi scale image features extracted by top-hat transform," *Optics & Laser Technology* **44**, 328-336 (2012).
34. M. Liao, Y.-q. Zhao, X.-h. Wang, and P.-s. Dai, "Retinal vessel enhancement based on multi-scale top-hat transformation and histogram fitting stretching," *Optics & Laser Technology* **58**, 56-62 (2014).
35. B. Gupta, and M. Tiwari, "Color retinal image enhancement using luminosity and quantile based contrast enhancement," *Multidimensional Systems and Signal Processing* **30**, 1829-1837 (2019).
36. T. Celik, "Spatial Entropy-Based Global and Local Image Contrast Enhancement," *IEEE Transactions on Image Processing* **23**, 5298 - 5308 (2014).
37. M. Zhou, K. Jin, S. Wang, J. Ye, and D. Qian, "Color Retinal Image Enhancement Based on Luminosity and Contrast Adjustment," *IEEE Transactions on Biomedical Engineering* (2018).
38. L. Cao, H. Li, and Y. Zhang, "Retinal image enhancement using low-pass filtering and α -rooting," *Signal Processing* **170**, 107445 (2020).
39. P. Dai, H. Sheng, J. Zhang, L. Li, J. Wu, and M. Fan, "Retinal Fundus Image Enhancement Using the Normalized Convolution and Noise Removing," *International Journal of Biomedical Imaging* **2016** 1-12 (2016).
40. N. D. Shemonski, Fredrick A. South, Y.-Z. Liu, S. G. Adie, P. S. Carney, and S. A. Boppart, "Computational high-resolution optical imaging of the living human retina," *Nat. Photonics* **9**, 1-5 (2015).
41. J. Chung, G. W. Martinez, K. C. Lencioni, S. R. Sadda, and C. C. Yang, "Computational aberration compensation by codedaperture- based correction of aberration obtained from optical Fourier coding and blur estimation," *Optica* **6**, 647-661 (2019).
42. A. Arias, and P. Artal, "Wavefront-shaping-based correction of optically simulated cataracts," *Optica* **7**, 22-27 (2020).
43. R. Dutta, S. Manzanera, A. Gambín-Regadera, E. Irlas, E. Tajahuerce, J. Lancis, and P. Artal, "Single-pixel imaging of the retina through scattering media," *Biomedical Optics Express* **10**, 4159-4167 (2019).
44. D. Singh, and V. Kumar, "A Comprehensive Review of Computational Dehazing Techniques," *Archives of Computational Methods in Engineering* (2018).

45. K. P. Senthilkumar, and P. Sivakumar, "A Review on Haze Removal Techniques," *Computer Aided Intervention and Diagnostics in Clinical and Medical Images*, 113-123 (2019).
46. S. Banerjee, and S. S. Chaudhuri, "Nighttime Image-Dehazing: A Review and Quantitative Benchmarking," *Archives of Computational Methods in Engineering* (2020).
47. Y. Wang, S. Zhuo, D. Tao, J. Bu, and N. Li, "Automatic local exposure correction using bright channel prior for under-exposed images," *Signal Processing* **93**, 3227-3238 (2013).
48. A. Mitra, S. Roy, S. Roy, and S. K. Setua, "Enhancement and restoration of non-uniform illuminated Fundus Image of Retina obtained through thin layer of cataract," *Computer Methods and Programs in Biomedicine* **156**, 169-178 (2018).
49. A. Gaudio, A. Smailagic, and A. e. Campilho, "Enhancement of Retinal Fundus Images via Pixel Color Amplification," *International Conference on Image Analysis and Recognition* (2020).
50. S. M. Pizer, E. P. Amburn, J. D. Austin, R. Cromartie, A. Geselowitz, T. Greer, B. ter Haar Romeny, J. B. Zimmerman, and K. Zuiderveld, "Adaptive histogram equalization and its variations," *Computer vision, graphics, and image processing* **39**, 355-368 (1987).
51. H. Koschmieder, "Luftlicht und sichtweite," *Naturwissenschaften*, **26**, 521-528 (1938).
52. E. J. McCartney, "Optics of the Atmosphere: Scattering by molecules and particles," New York, 408 (1976).
53. D. Mujbaile, and D. Rojatkhar, "Model based Dehazing Algorithms for Hazy Image Restoration – A Review," (2020).
54. Z. Shen, H. Fu, J. Shen, and L. Shao, "Modeling and Enhancing Low-Quality Retinal Fundus Images," *IEEE Transactions on Medical Imaging* **40**, 996-1006 (2021).
55. L. Cao, and H. Li, "Enhancement of blurry retinal image based on non-uniform contrast stretching and intensity transfer," *Medical & Biological Engineering & Computing* **59** (2020).
56. E. H. Land, and J. J. McCann, "Lightness and retinex theory," *Josa* **61**, 1-11 (1971).
57. R. Kimmel, M. Elad, D. Shaked, R. Keshet, and I. Sobel, "A variational framework for retinex," *International Journal of computer vision* **52**, 7-23 (2003).
58. W. Ma, and S. Osher, "A TV Bregman iterative model of Retinex theory," *Inverse Problems & Imaging* **6**, 697 (2012).
59. S. Park, S. Yu, B. Moon, S. Ko, and J. Paik, "Low-light image enhancement using variational optimization-based retinex model," *IEEE Transactions on Consumer Electronics* **63**, 178-184 (2017).
60. J. M. Morel, A. B. Petro, and C. Sbert, "A PDE formalization of Retinex theory," *IEEE Transactions on Image Processing* **19**, 2825-2837 (2010).
61. A. Blake, "Boundary conditions for lightness computation in Mondrian world," *Computer vision, graphics, and image processing* **32**, 314-327 (1985).
62. D. J. Jobson, Z. Rahman, and G. A. Woodell, "Properties and performance of a center/surround retinex," *IEEE Transactions on Image Processing* **6**, 451-462 (1997).
63. K. He, J. Sun, and X. Tang, "Single Image Haze Removal Using Dark Channel Prior," (2009).
64. A. Galdran, A. Alvarez-Gila, A. Bria, J. Vazquez-Corral, and M. Bertalmio, "On the Duality Between Retinex and Image Dehazing," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 8212-8221 (2018).
65. Z. Shi, M. M. Zhu, B. Guo, M. Zhao, and C. Zhang, "Nighttime low illumination image enhancement with single image using bright/dark channel prior," *EURASIP Journal on Image and Video Processing* **2018**, 1-15 (2018).
66. J. Wang, K. Lu, J. Xue, N. He, and L. Shao, "Single Image Dehazing Based on the Physical Model and MSRCR Algorithm," *IEEE Transactions on Circuits and Systems for Video Technology* **28**, 2190-2199 (2017).
67. M. Foracchia, E. Grisan, and A. Ruggeri, "Luminosity and contrast normalization in retinal images," *Med Image Anal* **9**, 179-190 (2005).
68. A. M. R. R. Bandara, and P. W. G. R. M. P. B. Giragama, "A Retinal Image Enhancement Technique for Blood Vessel Segmentation Algorithm," *2017 IEEE International Conference on Industrial and Information Systems*, 1-5 (2017).

69. Sonali, S. Sahu, A. K. Singh, S. P. Ghrera, and M. Elhoseny, "An approach for de-noising and contrast enhancement of retinal fundus image using CLAHE," *Optics & Laser Technology* **110**, 87-98 (2019).
70. N. Singh, L. Kaur, and K. Singh, "Histogram equalization techniques for enhancement of low radiance retinal images for early detection of diabetic retinopathy," *Engineering Science and Technology, an International Journal* **22**, 736-745 (2019).
71. I. Qureshi, J. Ma, and K. Shaheed, "A Hybrid Proposed Fundus Image Enhancement Framework for Diabetic Retinopathy," *Algorithms* **12**, 14 (2019).
72. R. Kumar, and A. Kumar Bhandari, "Luminosity and contrast enhancement of retinal vessel images using weighted average histogram," *Biomedical Signal Processing and Control* **71**, 103089 (2022).
73. R. Han, C. Tang, M. Xu, B. Liang, T. Wu, and Z. Lei, "Enhancement method with naturalness preservation and artifact suppression based on an improved Retinex variational model for color retinal images," *Journal of the Optical Society of America A* **40**, 155-164 (2023).
74. R. Fergus, B. Singh, A. Hertzmann, S. T. Roweis, and W. T. Freeman, "Removing camera shake from a single photograph," in *Acm Siggraph 2006 Papers*(2006), pp. 787-794.
75. J. Kotera, F. Šroubek, and P. Milanfar, "Blind Deconvolution Using Alternating Maximum a Posteriori Estimation with Heavy-Tailed Priors," *International Conference on Computer Analysis of Images and Patterns* 59-66 (2013).
76. A. Levin, Y. Weiss, F. Durand, and W. T. Freeman, "Understanding and evaluating blind deconvolution algorithms," in *2009 IEEE conference on computer vision and pattern recognition*(IEEE2009), pp. 1964-1971.
77. J. Pan, D. Sun, H. Pfister, and M.-H. Yang, "Blind Image Deblurring Using Dark Channel Prior," *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2016).
78. L. Chen, F. Fang, T. Wang, and G. Zhang, "Blind Image Deblurring with Local Maximum Gradient Prior," *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (2019).
79. U. Qidwai, and U. Qidwai, "Blind Deconvolution for Retinal Image Enhancement," *IEEE EMBS Conference on Biomedical Engineering & Sciences* (2010).
80. A. G. Marrugo, M. S. Millan, M. Sorel, and F. Šroubek, "Retinal image restoration by means of blind deconvolution," *J. of Biomedical Optics* **16**, 116016 (2011).
81. F. J. Ávila, J. Ares, M. C. Marcellán, M. V. Collados, and L. Remón, "Iterative-Trained Semi-Blind Deconvolution Algorithm to Compensate Straylight in Retinal Images," *J. Imaging* **7**, 1-15 (2021).
82. S. Zhang, C. A. B. Webers, and T. T. J. M. Berendschot, "Luminosity rectified blind Richardson-Lucy deconvolution for single retinal image restoration," *Computer Methods and Programs in Biomedicine* **229**, 107297 (2023).
83. B. Yang, H. Zhao, L. Cao, H. Liu, N. Wang, and H. Li, "Retinal image enhancement with artifact reduction and structure retention," *Pattern Recognition* **133**, 108968 (2023).
84. C. Wan, X. Zhou, Q. You, J. Sun, J. Shen, S. Zhu, Q. Jiang, and W. Yang, "Retinal Image Enhancement Using Cycle-Constraint Adversarial Network," *Front. Med.* **8**, 793726 (2021).
85. S. Chen, Z. Qian, and Z. Hua, "A Novel Un-Supervised GAN for Fundus Image Enhancement with Classification Prior Loss," *Electronics* **11** (2022).
86. K. TSIMA, "The reproducibility issues that haunt health-care AI," *Nature* **613** (2023).
87. F. N. Doubal, T. J. MacGillivray, N. Patton, B. Dhillon, M. S. Dennis, and J. M. Wardlaw, "Fractal analysis of retinal vessels suggests that a distinct vasculopathy causes lacunar stroke," *Neurology* **74**, 1102-1107 (2010).
88. C. Y. Cheung, S. Ong, M. K. Ikram, Y. T. Ong, C. P. Chen, N. Venketasubramanian, and T. Y. Wong, "Retinal vascular fractal dimension is associated with cognitive dysfunction," *J Stroke Cerebrovasc Dis* **23**, 43-50 (2014).
89. P. Zhu, F. Huang, F. Lin, Q. Li, Y. Yuan, Z. Gao, and F. Chen, "The relationship of retinal vessel diameters and fractal dimensions with blood pressure and cardiovascular risk factors," *PLoS one* **9**, e106551 (2014).
90. F. Huang, B. Dashtbozorg, J. Zhang, E. Bekkers, S. Abbasi-Sureshjani, T. T. Berendschot, and B. M. Ter Haar Romeny, "Reliability of Using Retinal Vascular Fractal Dimension as a Biomarker in the Diabetic Retinopathy Detection," *J. Ophthalmol.* **2016**, 6259047 (2016).

91. M. Ciancaglini, G. Guerra, L. Agnifili, R. Mastropasqua, V. Fasanella, M. Cinelli, C. Costagliola, and L. Ambrosone, "Fractal dimension as a new tool to analyze optic nerve head vasculature in primary open angle glaucoma," *In Vivo* **29**, 273-279 (2015).
92. T. Chanwimaluang, G. Fan, and S. R. Fransen, "Hybrid retinal image registration," *IEEE transactions on information technology in biomedicine* **10**, 129-142 (2006).
93. T. Kauppi, V. Kalesnykiene, J.-K. Kamarainen, L. Lensu, I. Sorri, A. Raninen, R. Voutilainen, H. Uusitalo, H. Kälviäinen, and J. Pietilä, "DIARETDB1 diabetic retinopathy database and evaluation protocol," presented at the Medical Image Understanding and Analysis2007.
94. G. Wang, Z. Wang, Y. Chen, and W. Zhao, "Robust point matching method for multimodal retinal image registration," *Biomedical Signal Processing and Control* **19**, 68-76 (2015).

Chapter 2

A reflective background padding algorithm for fundus images preprocessing

Shuhe Zhang, Shuo Zhang, and Tos T. J. M. Berendschot

Abstract

2

Padding the black background for fundus images using similar colors allows the convolution kernel to move smoothly above the image and benefits all convolution-based image processing methods. However, according to our knowledge, there is no report that is dedicated to reporting relevant techniques on fundus images. In order to fill this gap, in this study, we propose a simple and straightforward algorithm to achieve fundus image background padding in a reflective manner. The color of a given pixel in the black background area is replaced by the color of another pixel that is symmetric with respect to the cross point between the edge of the fundus area and the straight line between those two pixels. Experimental results show that our algorithm is suitable for any kind of fundus image and supports all convolution treatments.

Keywords

Retinal image; Image processing; Background padding; Convolution

1. Introduction

Many fundus imaging processing algorithms rely on convolution methods, such as enhancing retinal images by the Retinex theory [5-11] or blood vessels segmentation using Match-filters [1, 12]. During those processes, the convolution kernel moves above the retinal image and extracts or removes the image's features.

Ideally, the convolution kernel is required to move smoothly (no rapid changes of pixel value in raw images). Most mathematical software packages including TensorFlow, MATLAB, and Mathematica, have already embedded the boundary padding processes to deal with the edge effect during the convolution [13, 14]. However, due to the universal feature of fundus images where the retinal area is within a disk region separated from the black background, the traditional padding method cannot handle the black ground since they are designed for squared areas. As a result, the convolution kernel suffers from sudden changes in pixel value at the boundary of the retinal area and black background, yielding problematic pixels at the boundary depending on the size of the convolution kernel.

It is known that such problematic pixels can lead to failure in retinal image enhancement or misleading feature extracting [10, 11, 14], since the problematic pixels will be enhanced and accumulated in each processing step. To avoid the boundary problem at the edge of the retinal area, a background padding procedure that fills the black background with similar colors is needed.

The content-aware padding method is a good candidate for retinal image background padding, however, this is a waste of computational power since the only aim of padding the background is to let the convolution kernel move smoothly, and the filled area will be removed once the convolution and subsequent tasks are finished. A simple and efficient algorithm to achieve black background padding is, therefore, in demand. According to our knowledge, there is no such report that is dedicated to reporting relevant background padding techniques for fundus images.

In order to fill this gap, in this research, we propose a simple and straightforward algorithm to achieve fundus image background padding. The remainder of this manuscript is structured as follows: Section 2 describes the background padding algorithm in detail. Experimental results are demonstrated in Section 3, where convolution results with and without background padding are compared. Section 4 is the closing remark.

2. Method details

2.1 Obtaining ROI mask

Assuming I denote the 3D matrix of a given color fundus image, which size is $M \times N \times 3$. M and N are the numbers of the pixels in the vertical and horizontal directions, respectively. A typical fundus image is divided into two areas – the centered retinal imaging area and the black background area. Only the centered retinal imaging area is the region of interest (ROI) where further analyses are applied to the ROI mask can be obtained by setting a threshold in the red channel of I , which is

$$\mathbf{M}(x, y) = \begin{cases} 1, & \mathbf{I}(x, y, 1) > t \\ 0, & \text{otherwise} \end{cases}, \quad (2-1)$$

where t is the threshold and is set to 10 empirically for 8-bit color fundus images [1]. As shown in Fig. 2-1 (b), the ROI mask (colored in gray) is obtained from the fundus images given in Fig. 2-1 (a). Based on the ROI mask, the reflective padding procedure is applied to the outside of the ROI mask. Furthermore, the shape of ROI for all fundus images is round or can be approximated by convex polygons due to the design of the fundus camera. This feature will largely simplify our background padding task.

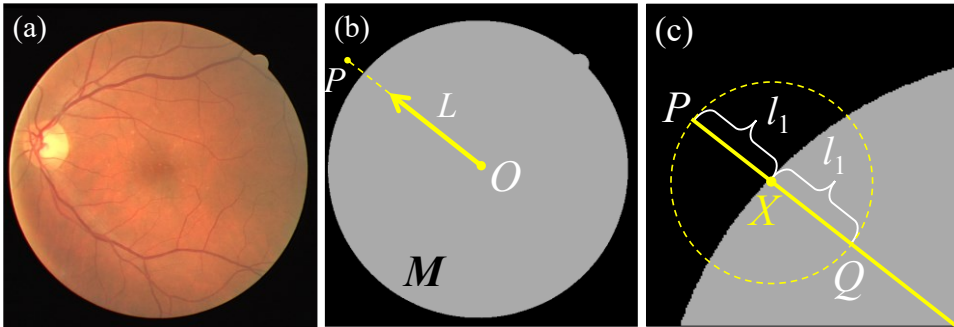


Fig. 2-1. The ROI for fundus image (a) is obtained using the threshold method and is shown in (b). (c) denotes the idea of how to find the Q that is on the line L and symmetric to P with respect to X.

2.2 Reflective background padding

A natural train of thought to achieve reflective padding has 4 detailed procedures. (1)First we draw a line, say L that connects the center point O of the image and the point P where the color needs to be padded; (2)Next, we pinpoint the intersection point X between the line and the edge of M as shown in Fig. 2-1 (c); (3)Then, we note the distance between P and X , and find the point Q which locates also on L and is symmetrical to P with respect to X ; (4)Finally we copy the color at point Q to that of point P in the raw image I .

These 4 procedures are straightforward and simple, but they should be applied to the M and S , which are discrete matrices. Therefore, line L should also be discretized. To do so, we first note the coordinates of points P and O on a 2D image by their pixels index which are (x_1, y_1) and (x_0, y_0) , respectively. $x_0, x_1, y_0,$ and y_1 are positive integers. Then we use Bresenham's line algorithm [2] to create a discrete version of L . The algorithm is modified for the application in this research and is summarized in Algorithm 1 by pseudocode. In this algorithm, the pixels for L are generated from points P to O in sequence and will make point P always the starting point of L . Note that if O is assigned to the center pixel of the image, $x_0 = \langle N/2 \rangle, y_0 = \langle M/2 \rangle$, and $\langle \rangle$ denotes round toward the nearest integer.

As shown in Fig. 2-2 (a), the segment of L inside the ROI is colored in green where the pixel values of the M are 1. The segment outside the ROI is marked in red, and the corresponding pixel values in the M are 0.

Next, the position of X and the distance between X and P are determined. Since the coordinates of all pixels in L are noted, if we search on M along L from point P to point O direction, the first pixel that equals 1 implies the first intersection point between L and M . Therefore the position of that pixel is noted to be X . As shown in Fig. 2-2 (b), X is marked in yellow.

The distance between X and P can be determined by counting the number of pixels shifting between their coordinates in both horizontal and vertical directions. In the case shown in Fig. 2-2 (b), the horizontal pixel-shifting between P and X is 12 pixels, while the vertical pixel-shifting is 6 pixels.

Algorithm 2-1: generating Bresenham's line

Input: pixel-coordinates of $O(x_0, y_0)$ and $P(x_1, y_1)$.

$k \leftarrow (|y_1 - y_0| > |x_1 - x_0|)$.

if k **then:** swap(x_0, y_0), swap(x_1, y_1).

$\Delta x \leftarrow |x_1 - x_0|$, $\Delta y \leftarrow |y_1 - y_0|$.

$\varepsilon \leftarrow \Delta x / 2$, $y \leftarrow y_1$, $\alpha \leftarrow \text{sign}(y_1 - y_0)$.

$L \leftarrow \Delta x + 1$.

$\text{count} \leftarrow 0$.

for $x = x_1 : \text{sign}(x_0 - x_1) : x_0$ **do**

$\text{count} \leftarrow \text{count} + 1$.

if k **then:** $xdata[\text{count}] = x$, $ydata[\text{count}] = y$.

else: $xdata[\text{count}] = y$, $ydata[\text{count}] = x$.

$\varepsilon \leftarrow \varepsilon - \Delta y$.

if $\varepsilon < 0$ **then:** $y \leftarrow y - \alpha$, $\varepsilon \leftarrow \varepsilon + \Delta x$.

end for

Output: $xdata$ and $ydata$ pixel-coordinates on line L , line length: L

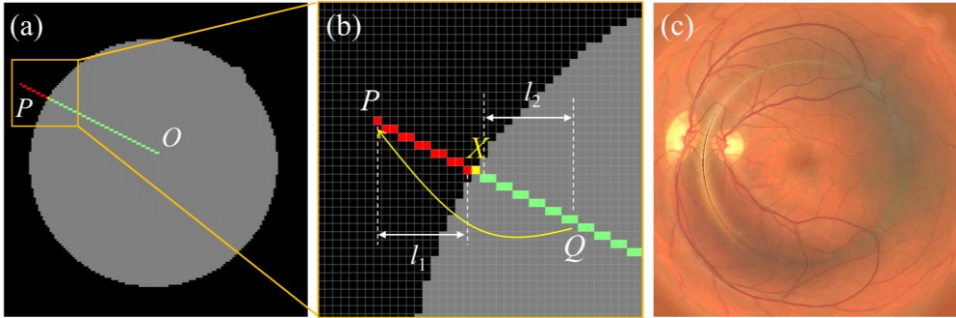


Fig. 2-2. Sketch of padding process. (a) Discrete L in M . (b) partially zoom-in for the area in the yellow box in (a). (c) Padding results.

The coordinates of Q can be determined in two ways. First, we can count the same pixels shifting with respect to X on L and find the position of Q . Second, since P is the first point of L , if X is the k -th point of L , Q must be the $(2k-1)$ -th point of L . In this way, the position of Q can be also determined. **Algorithm 2-2** summarizes the processes.

Algorithm 2-2: Background Padding

Input: I

M is obtained by Eq. (1)

for all points $P(x_1, y_1)$ **in** M **do**

if $M[x_1, y_1] == 0$ **then** :

$(xdata, ydata, L) \leftarrow \text{BresenhamLine}(P, O)$.

$k \leftarrow 1$.

for $i = 1 : L$ **do**

if $M[xdata[i], ydata[i]] == 1$ **then** : $k \leftarrow i$, **break**.

end for

$I[x_1, y_1] \leftarrow I[xdata[\min(2k-1, L)], ydata[\min(2k-1, L)]]$.

end for

Output: background padded I

Particularly, the pixel of Q should never exceed the pixel of O in order to avoid the case where the index of Q outside the I if P is very far from X . Finally, the color in pixel P is replaced by that of pixel Q in the original image I . Fig. 2-2 (c) shows the padding result of image Fig. 2-1(a).

3. Multi-threads accelerating

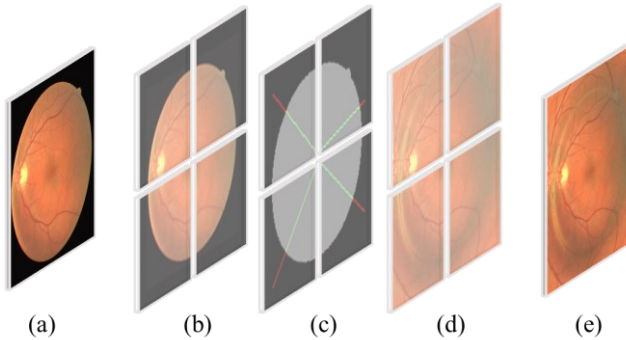


Fig. 2-3. Multithreads acceleration by dividing the Raw image into several subsets. (a) Raw image. (b) Four subsets of (a). The padding algorithm is applied simultaneously to each subset as shown in (c). (d) is the padded subset and their combination forms the final output (e).

The proposed algorithm can be accelerated using multi-threads. Fig. 2-3 demonstrates the idea of multi-threads. The image I and its ROI mask M are divided into four subsets as shown in Fig. 2-3 (b) and Fig. 2-3 (c), respectively. The background padding algorithm is applied to each subset (Fig. 2-3(c)) simultaneously, yielding the padded subsets as shown in Fig. 2-3 (d). Finally, all subsets are combined into a single image as shown in Fig. 2-3 (e).

Since the area of M in each subset has a fan shape in the case shown in Fig. 2-3, points O in each subset are assigned to the vertex of the fan. On the contrary, if the I and M are divided into 3 by 3 subsets, positions of point O in each subset are assigned in different ways. By using multi-threads accelerating, the average time that is needed to finish the entire background padding process can be improved by several times (depending on the subsets numbers)

4. Experimental results

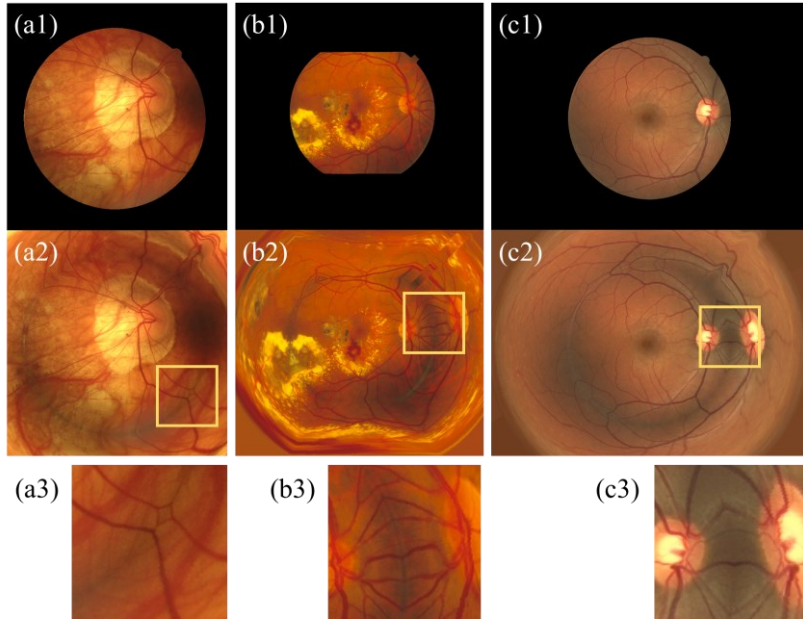


Fig. 2-4. Demonstration of background padding results. (a1) to (c1): raw images. (a2) to (c2): background padded results. (a3) to (c3) are zoomed-in of the region in yellow boxes.

Three fundus images were randomly chosen from the DRIVE [3] and STARE [4] databases. The raw images were first zero-padded into different sizes as shown in the first row of Fig. 2-4 and then filled by the proposed method as shown in the second row of Fig. 2-4. The black backgrounds were filled in a reflective manner along the radial direction.

Regions in the yellow boxes were enlarged to show detailed texture changing at the boundary of the ROI. As shown in the third row of Fig. 2-4, the color and textures are smoothly changing from inside toward the outside of ROI.

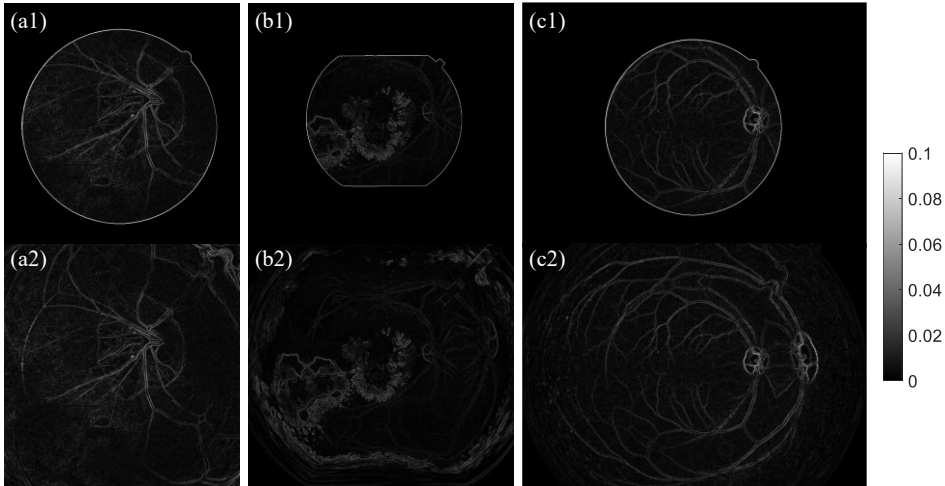


Fig. 2-5. Demonstration of background padding results. (a1) to (c1): $\|\nabla I\|$ of raw images of grayscale images. (a2) to (c2): background padded results.

The smoothness of an image can be described by its first-order derivative. Here we use $\|\nabla I\| = 0.5(|\partial_x I| + |\partial_y I|)$ of a grayscale image (the color fundus image is average in three color channels to obtain the grayscale image) to measure the smoothness of an image. The grayscale of the raw images is normalized within $[0, 1]$.

$\|\nabla I\|$ for fundus images in Fig. 2-4 (a1) to (c1) are shown in Fig. 2-5(a1) to (c1). Where $\|\nabla I\|$ has a high value (white circular pattern) at the boundary of the ROI, denoting the rapid changes of grayscale values at the regions. $\|\nabla I\|$ of fundus images after background padding is shown in Fig. 2-5 (a2) to (c2). The small values of $\|\nabla I\|$ at the boundary of the ROI denote slow- variations of the pixel value. According to this feature, the convolution kernel can smoothly move during the convolution process.

5. Retinex enhancement with and without background padding

Background padding directly benefits all convolution-based methods. The Retinex method, which is able to uniform the illumination of the image can be one of the

examples to show the advantage of background padding. A typical Retinex method for retinal image enhancement can be expressed by [5]

$$\mathbf{I}_{new} = \exp \left(\log(\mathbf{I} + \xi) - \log \left\{ \mathbf{I} \otimes \left[\frac{1}{2\pi\omega^2} \exp \left(-\frac{\mathbf{r}^2}{2\omega^2} \right) \right] + \xi \right\} \right) - \xi, \quad (2-2)$$

where ξ is a small value to avoid the situation of $\log(0)$, \mathbf{r} is the 2D position vector. ω is the width of the Gaussian kernel which should not be smaller than the largest structure of the fundus image (normally optical disk). In the following experiment $\xi = 0.01$ and $\omega = \langle \sqrt{MN}/20 \rangle$. For the background padding, the fundus images are first padded with zero elements according to the size of the kernel, and then the zero background is filled by our algorithm.

A max-min color correction procedure is applied to each color channel of \mathbf{I}_{new} to correct the color distortion.

$$\mathbf{I}_{new,colorcorrection} = [\max(\mathbf{I}) - \min(\mathbf{I})] \cdot \left[\frac{\mathbf{I}_{new} - \min(\mathbf{I}_{new})}{\max(\mathbf{I}_{new}) - \min(\mathbf{I}_{new})} \right] + \min(\mathbf{I}). \quad (2-3)$$

Fig. 2-6 compares the Retinex outputs without and with background padding. Results without background padding are shown in the second row of Fig. 2-6 from (a2) to (d2), where there are halo effects at the boundary of ROIs due to the sudden change of pixel values during the convolution process. Since the halo effects have high intensity than other parts of the image, the visual performance of the entire image is significantly decreased.

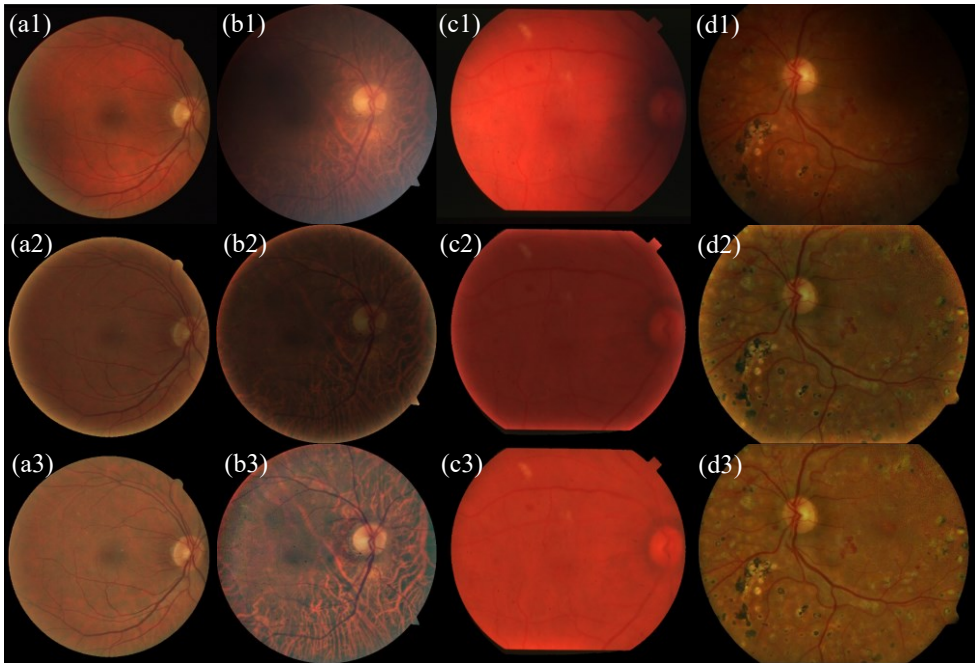


Fig. 2-6. Retinal image enhancement using the Retinex method. (a1) to (d1) Raw images. (a2) to (d2) Enhanced without background padding. (a3) to (d3) Enhanced with background padding.

On the contrary, the background padding allows the convolution kernel to smoothly moves on the image and won't lead to the halo effects, which largely increase the performance of enhanced images as shown in the third row of Fig. 2-6 from (a3) to (d3).

Currently, many retinal image enhancement methods are inspired by the Retinex theory, where the Gaussian convolution kernel is used to suppress the low-frequency components and enhance the structure of the retinal images. The low-pass filtering and alpha-root (LPAR) method [11], for example, is a state-of-the-art retinal image enhancement method that shares a similar convolution algorithm. Here we take the LPAR method as an example to show the importance of background padding.

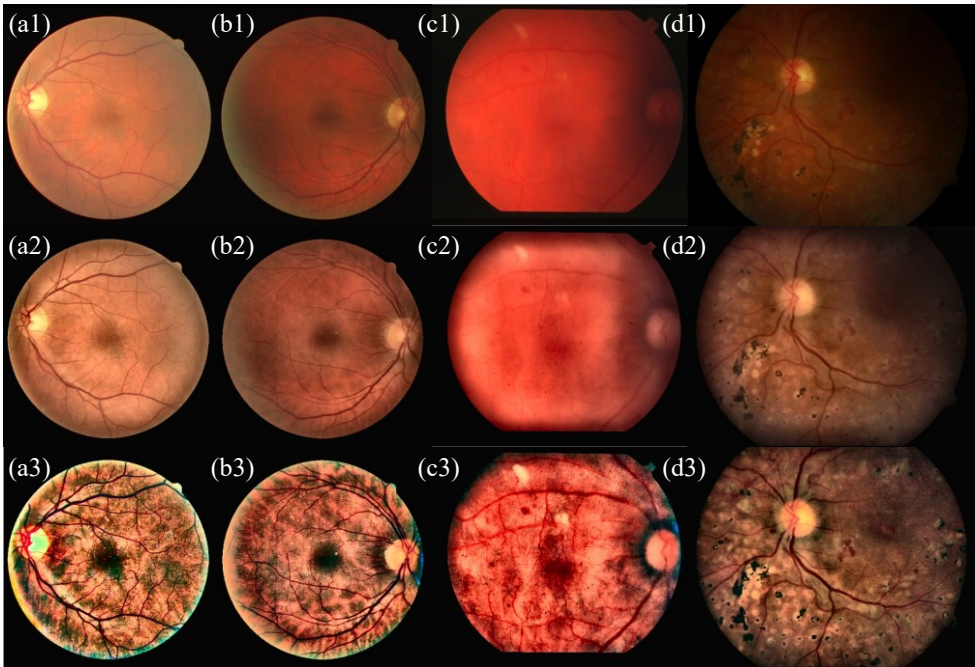


Fig. 2-7. Retinal image enhancement using the LPAR method. (a1) to (d1) Raw images. (a2) to (d2) Enhanced without background padding. (a3) to (d3) Enhanced with background padding.

The first row of Fig. 2-7 shows the raw images in this group of experiments. Enhancement results of LPAR outputs without and with background padding are shown in the second (a2) to (d2) and the third row (a3) to (d3) of Fig. 2-7, respectively. Without the background padding, the enhanced images show only mild contrast-enhanced results. This is mainly due to the halo effects during the convolution process. In LPAR, the low-frequency removal is performed adaptively on the raw image according to the average pixel intensity in the ROI. Due to the high intensity of halo effects, the algorithm will wrongly estimate how many low-frequency components should be removed, and thus influences all subsequent procedures.

The background padding prevents the convolution from the halo effects, and the algorithm can correctly and adaptively remove the low-frequency components, and successfully enhance the image quality as shown in the third row of Fig. 2-7.

Note that the halo effects in Fig. 2-7 (a3) and (b3) are not caused by our background padding algorithm, but due to the halo effects on the raw images shown in Fig. 2-7 (a1) and (b1), and further enhanced by the LPAR method.

Acknowledgments

Shuo Zhang and Shuhe Zhang acknowledge the China Scholarship Council for the support by State Scholarship Fund No. 201309110103 and No.201908340078, respectively.

References

1. L. Xiong, H. Li, and L. Xu, "An Approach to Evaluate Blurriness in Retinal Images with Vitreous Opacity for Cataract Diagnosis," *Journal of Healthcare Engineering*, vol. 2017 no. pp. 1-16, 2017.
2. W. E. Wright, "Parallelization of Bresenham's line and circle algorithms," *IEEE Computer Graphics and Applications* vol. 10, no. pp. 60-67, 1990.
3. "DRIVE: Digital Retinal Images for Vessel Extraction", retrieved <https://drive.grand-challenge.org/>.
4. "STructured Analysis of the Retina", retrieved <https://cecas.clemson.edu/~ahoover/stare/>.
5. E. Peli and T. Peli, "Restoration of retinal images obtained through cataracts," *IEEE Transactions on Medical Imaging*, vol. 8, no. 4, pp. 401-406, 1989.
6. P. Dai, H. Sheng, J. Zhang, L. Li, J. Wu, and M. Fan, "Retinal Fundus Image Enhancement Using the Normalized Convolution and Noise Removing," *International Journal of Biomedical Imaging*, vol. 2016 no. pp. 1-12, 2016.
7. B. Savelli, A. Bria, A. Galdran, C. Marrocco, M. Molinara, A. e. Campilho, and F. Tortorella, "Illumination correction by dehazing for retinal vessel segmentation," *IEEE 30th International Symposium on Computer-Based Medical Systems*, vol. no. pp. 2017.
8. L. Xiong, H. Li, and L. Xu, "An enhancement method for color retinal images based on image formation model," *Computer Methods and Programs in Biomedicine*, vol. 143, no. pp. 137-150, 2017.
9. A. Mitra, S. Roy, S. Roy, and S. K. Setua, "Enhancement and restoration of non-uniform illuminated Fundus Image of Retina obtained through thin layer of cataract," *Computer Methods and Programs in Biomedicine*, vol. 156, no. pp. 169-178, 2018.
10. L. Cao and H. Li, "Enhancement of blurry retinal image based on non-uniform contrast stretching and intensity transfer," *Medical & Biological Engineering & Computing*, vol. no. pp. 2020.
11. L. Cao, H. Li, and Y. Zhang, "Retinal image enhancement using low-pass filtering and α -rooting," *Signal Processing*, vol. 170, no. pp. 107445, 2020.
12. M. Foracchia, E. Grisan, and A. Ruggeri, "Luminosity and contrast normalization in retinal images," *Medical Image Analysis*, vol. 9, no. pp. 179-190, 2005.
13. R. Liu and J. Jia, "Reducing boundary artifacts in image deconvolution," in *2008 15th IEEE International Conference on Image Processing*, (IEEE, 2008), 505-508.
14. S. Wu, G. Wang, P. Tang, F. Chen, and L. J. a. p. a. Shi, "Convolution with even-sized kernels and symmetric padding," vol. no. pp. 2019.

Chapter 3

DPFR:

A double-pass fundus reflection model for efficient single retinal image enhancement

Shuhe Zhang, Carroll A. B. Webers, and Tos T. J. M. Berendschot

Signal Processing, 2022 (192), pp. 108400.

DOI: [10.1016/j.sigpro.2021.108400](https://doi.org/10.1016/j.sigpro.2021.108400).

Abstract

This study introduces a novel image formation model - the double pass fundus reflection (DPFR) model for retinal image enhancement (restoration). The DPFR model reveals the specific double pass fundus reflection feature that was hitherto neglected in modeling the light propagation of fundus imaging in all published reports on retinal image enhancement. Based on the DPFR model, the procedures of the proposed retinal image restoration algorithm are given. The failure of the dark channel prior on retinal images in RGB color space is clarified. While a solution about how to bypass the challenge is proposed. Each step of DPFR is tested experimentally with retinal images of different degraded situations to validate its robustness. Moreover, the DPFR method is tested on 906 images from five public databases. Six image quality matrixes including image definition, image sharpness, image local contrast, image multiscale contrast, image entropy, and fog density are used for objective assessments. The results are compared to the state-of-the-art methods, showing the superiority of DPFR over the others in terms of restoration quality and implementation efficiency.

Keywords

Ophthalmology; Retinal image; Image enhancement; Double pass

1. Introduction

The introduction of the ophthalmoscope by Helmholtz [1] allowed to obtain images of the retina and put ophthalmology on the map as a separate sub-area of medicine. Ever since, retinal images are widely used by ophthalmologists for early detection, diagnosis, and monitoring of ocular diseases and their progression. Morphologic changes due to eye diseases like diabetic retinopathy [2], glaucoma [3], and age-related macular degeneration [4-6] can be directly observed in these images. Moreover, neurological diseases such as stroke and cognitive dysfunction can be also diagnosed through retinal images [7]. Obviously, the higher the image clarity, the more detailed information can be observed from the image, and the better their diagnostic capabilities.

A typical retinal imaging platform can be regarded as two coupled imaging systems. One is the ocular system, and the other is a reflective imaging system that normally illuminates the fundus through the pupil and collects the reflected light from the retina, forming the image on the camera sensor. Retinal images may suffer from severe degenerations, not only because of uneven intraocular illumination, limited optical resolution, and aberrations of the eyes but also due to the blurry ocular media, including corneal edema, uveitis, vitreous hemorrhage, or cataracts. They all result in decreased image clarity.

To increase the image clarity again, several contrast-enhancement methods have been proposed, which can be divided into two major categories: data domain methods and restored model methods [8]. Data domain methods are further divided into two families based on their algorithms. The first family is known as the transform-domain algorithm which transforms a raw image into a new function of other parameters such as spatial frequency domain corresponding to the Fourier transform [9], or structure feature domain corresponding to the top-hat transform [10, 11]. The image is processed in the transformed domain and then transformed back resulting in a new image with enhanced contrast. The transform-domain algorithm enables us to globally or locally modify the weight for different structures within the image. However, due to its

significant cost of computation resources, the image-domain algorithm, which is the second family of data domain methods, is favored the most.

The core idea of the image-domain algorithm is the gray-level adjustment. Histogram equalization and its improved version, the contrast-limited adaptive histogram equalization (CLAHE), are usually used for quick and simple retinal image enhancement [12]. Other histogram modification method such as q -quantile [13] and gray-scale global spatial entropy method also shows promising results in improving the image's contrast [8, 14]. The global gray-level adjustment methods including the gamma map [15] and α -rooting [16] use a fixed function to convert the global gray-level distribution for adjusting the brightness of retina imaging.

Another group of image-domain algorithms uses filters to enhance contrast. These algorithms are similar to transform-domain algorithms but use a convolution kernel to separate the background and foreground information of an image [16, 17]. The foreground information usually corresponds to the detailed structure of an image. By modifying the weights between background and foreground, the contrast of the detailed structure can be enhanced. In general, data domain methods belong to pure signal (image) processing techniques that normally take a few considerations of the physical insight of the image formation and enhancement.

In order to obtain self-consistent methods for retinal imaging enhancement, restored model methods have been developed. These restored model methods share a similar idea of computational imaging, i.e. a physical model is built to describe the optical process of forming an image under the impact of degeneration agents which could be optical aberrations, unstable vibrations, or limited optical resolution. By directly or indirectly measuring the optical properties of these degeneration agents, one can compensate for the degeneration agents by digitally mimicking the propagation of the optical wave and modifying the wavefront of light [18-21]. Imaging through scattering media, for example, is a well-known application of computational imaging [20, 21].

Different from computational imaging, the restored models for imaging enhancement do not measure the optical properties of the degeneration agents but try to find the solutions corresponding to their statistical properties in optical or visual aspects. The solutions can be regarded as roughly estimated versions of those degeneration agents and can be also compensated by applying them to the image formation model, resulting in enhanced images.

Restored model methods are widely used for imaging de-hazing [22, 23], underwater image enhancement, and night image enhancement [24], while only a few studies have reported their use in retinal imaging enhancement. To our knowledge, the first publication about the application of restored model methods in retinal image enhancement can be traced back to 1989 [25], where the model for imaging the retina in photographs taken through intraocular scatter is considered similar to the model used to represent imaging of the earth from a satellite in the presence of light cloud cover. The scattering effect is removed (or suppressed) by using the Retinex theory.

Based on the image formation model, Xiong *et al.* [8] proposed to use of intensity correction and histogram adjustment to preprocess the image, and then a transmission map is generated according to the intensity of the preprocessed image in each color channel. The haze effect can be therefore suppressed through dehazing. Although the performance of their approach is good, it relies on statistical and empirical properties of the retina imaging database to determine the algorithm parameters, which makes it hard to be applied to different databases. A subsequent study [26] showed the results of using the illumination-reflectance model of image formation to correct the illumination of retinal images. In this research, the color-inversed dark-channel prior, also known as bright-channel prior [27], is employed which shows efficient illumination correction. Mitra *et al.* [28] proposed a model that includes a thin layer of cataracts. Gaudio *et al.* [29] demonstrated a pixel color amplification method for retina imaging enhancement which shows good performance in enhancing the detailed structure of retina images.

All these models show satisfactory experimental results, however, the imaging formation models they used have the common drawback that they do not take into account the physical principle underlying the fundus imaging, in particular the double-pass principle of the light path. Note that the restored models in all the above-mentioned four studies are modified versions of Koschmieder [30] and McCartney's [31] imaging formation formula, which applies to natural outdoor scenes where the illumination is provided by sunlight. Since the illumination light is far away from the scene and not in the same optical path as that of the camera, the transmission function of the haze or scatter is only applied to the reflected light from the scene which results in the exponential index of the transmission term to be one, however, this is rather different from retinal imaging.

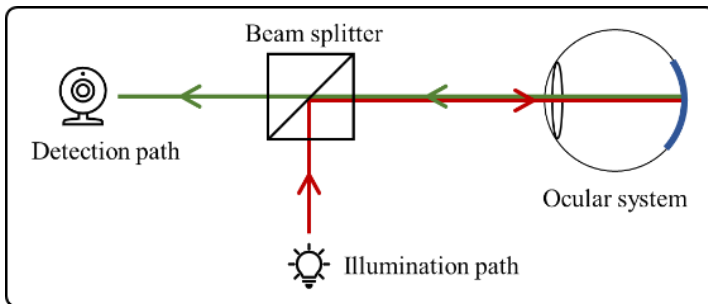


Fig. 3-1. The typical optical path in a fundus camera. Either a beam splitter or a mirror is used to deliver the illumination light into the ocular system.

As shown in Fig. 3-1, the fundus camera is a reflective/Epi-illumination imaging system where the incident illumination light is in the same optical path as that of the reflected light from the fundus. Therefore, the transmission function should be not only applied to the incident illumination light but also applied to the reflected (backscattering) light from the fundus resulting in a doubled exponential index for the transmission term [32-34].

In section 2 of this paper, we briefly review the formation of retinal images and derive the imaging formation model according to the actual optical process of retinal

imaging. The updated fundus reflection model is called the double-pass fundus reflection (DPFR) model. The inverse problem of retinal image restoration is also derived in section 3. Our derivation shows that the enhanced retinal image restoration can be simply achieved by one step of low-frequency removing using the Retinex theory, and two steps of the imaging de-hazing. Our DPFR model embeds illumination correction, stray light removal, and contrast enhancement using a single retinal image, without the need for statistic properties of the image database, and as such widens its applications.

In Section 4, we compare the performance of the DPFR model with different state-of-the-art methods on five different databases for both visual and objective assessment. The final section includes our concluding remarks. In the following analyses, all pixel values are normalized within the range of [0, 1].

2. Forward problem: double-pass fundus reflection model

The double-pass fundus reflection (DPFR) model is inspired by how a retinal image is formed, which is similar to a multi-layer model [35].

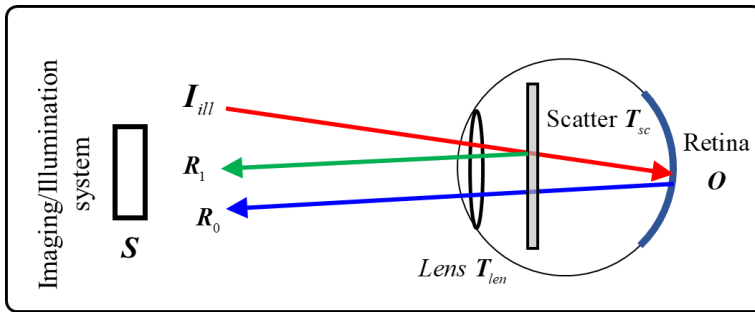


Fig. 3-2. The sketch of the double-pass fundus reflection model. S is the final output image of the imaging system.

In retinal imaging, intraocular scattering is very difficult to avoid, resulting in straylight that degenerates the retinal imaging quality. In the DPFR model as shown in Fig. 3-2, it is assumed that the intraocular scattering effect is caused by an intraocular

scatter, denoted by its transmission matrix $T_{sc}(\mathbf{r})$, which could be the cataract, vitreous opacities, or more complex cases. Here $\mathbf{r} = (x, y)$ is the 2D coordinate vector. Although there could also be a scatter outside the eye (in region A) for a typical imaging system, we assume that the imaging system is carefully maintained, and outside scatter can be ignored. Corneal and lens reflection is ignored since it is removed by the pupil mask in all commercial fundus cameras. The transmission matrix of the cornea and lens is together regarded as $T_{lens}(\mathbf{r})$.

Let \mathbf{S} be the final image on the camera sensor plane, it is mainly provided by two backscattering components $\mathbf{R}_0(\mathbf{r})$ and $\mathbf{R}_1(\mathbf{r})$. \mathbf{R}_0 denotes the backscattering from the retina $\mathbf{O}(\mathbf{r})$, while \mathbf{R}_1 is the backscattering from the intraocular scatter. Let $\mathbf{I}_{ill}(\mathbf{r})$ be the illumination from the outside of the eye and is in the same optical path as that of detection for all fundus camera systems. Our DPFR model, therefore, considers how \mathbf{I}_{ill} enters the eye and interacts with the lens, intraocular scatters, and retina. Accordingly, \mathbf{R}_0 and \mathbf{R}_1 is given by the product between the illumination and layers of scatter or retina, and can be separately expressed as

$$\mathbf{R}_0(\mathbf{r}) = \mathbf{I}_{ill}(\mathbf{r}) \cdot [\mathbf{T}_{lens}(\mathbf{r}) \cdot \mathbf{T}_{sc}(\mathbf{r})]^2 \cdot \mathbf{O}(\mathbf{r}), \quad (3-1)$$

and

$$\mathbf{R}_1(\mathbf{r}) = \mathbf{I}_{ill}(\mathbf{r}) \cdot \mathbf{T}_{lens}^2(\mathbf{r}) \cdot [\mathbf{I} - \mathbf{T}_{sc}(\mathbf{r})], \quad (3-2)$$

where \mathbf{I} denotes the unit matrix. The square terms in Eq. (3-1) and Eq. (3-2) show the core idea of the DPFR model in which the illumination \mathbf{I}_{ill} interacts twice with the scatter and lens. The final image \mathbf{S} in the camera sensor is the sum of two backscattering components, which is

$$\mathbf{S}(\mathbf{r}) = \mathbf{R}_0 + \mathbf{R}_1 = \mathbf{I}_{ill}(\mathbf{r}) \cdot \mathbf{T}_{lens}^2(\mathbf{r}) \cdot [\mathbf{T}_{sc}^2(\mathbf{r}) \cdot \mathbf{O}(\mathbf{r}) + \mathbf{I} - \mathbf{T}_{sc}(\mathbf{r})]. \quad (3-3)$$

The physical insight of Eq. (3-3) is clear, which shows that the degeneration of the retinal image is mainly due to three parts. (1) the uneven illumination condition \mathbf{I}_{ill} , (2) the filtering by the human lens \mathbf{T}_{lens} , and (3) the intraocular scattering \mathbf{T}_{sc} . Obviously,

the retinal image \mathbf{O} can be restored if the illumination matrix \mathbf{I}_{ill} , and both of the transmission matrix \mathbf{T}_{lens} and \mathbf{T}_{sc} can be measured as precisely as possible.

Although it is hard for us to measure the value for those matrixes based on a single retina image, we are able to get a rough estimation version of them according to some pre-known information about their statistical property. The pre-known information is also known as ‘prior’. With the estimations of the \mathbf{I}_{ill} , \mathbf{T}_{lens} and \mathbf{T}_{sc} , we can solve Eq. (3-3) and obtain an enhanced version for \mathbf{O} .

3. Inverse problem: image restoration.

Our method is divided into four steps, (1) the image pre-processing; (2) coarse illumination correction; (3) fine illumination boosting; and (4) scattering suppression.

3.1 Image pre-processing

A retinal image is composed of two major regions - a circular retina and a black background. Since the image will be filtered using a low-pass filter in the following section, unexpected results often occur when the convolution kernel slides through the retinal due to the sudden change of pixel values. Padding the black areas before an enhancement is an effective way to reduce an over-enhancement of the retinal boundary. The padding can be based on a mirror reflection or content-aware filling.

3.2 Coarse illumination correction

In the section, the coarse illumination correction algorithm applies to the red, green, and blue channels repeatedly. Here, we demonstrate the algorithm in a single-color channel.

We rewrite the illumination term, $\mathbf{I}_{ill} \cdot \mathbf{T}_{lens}^2$ as $\mathbf{I}_{ill} \cdot \mathbf{T}_{lens}^2 = \mathbf{I}_{coarse} \cdot \mathbf{I}_{fine}$, where \mathbf{I}_{coarse} , and \mathbf{I}_{fine} denote the coarse illumination pattern (slowly varying in the spatial domain) and fine illumination pattern (fast varying in the spatial domain) projected on the retina, respectively. Practically, the coarse illumination pattern denotes the overall illumination provided by the light source, while the fine illumination pattern can be due

to the intensity fluctuations or shadows of dust particles inside the optical path during the camera capturing and some noise. Eq. (3-3) is then rewritten as

$$\mathbf{S}(\mathbf{r}) = \mathbf{I}_{coarse}(\mathbf{r}) \cdot \mathbf{S}_0(\mathbf{r}), \quad (3-4)$$

with

$$\mathbf{S}_0(\mathbf{r}) = \mathbf{I}_{fine}(\mathbf{r}) \cdot \mathbf{S}_{ic}(\mathbf{r}), \quad (3-5)$$

and

$$\mathbf{S}_{ic}(\mathbf{r}) = \mathbf{T}_{sc}^2(\mathbf{r}) \cdot \mathbf{O}(\mathbf{r}) + \mathbf{1} - \mathbf{T}_{sc}(\mathbf{r}). \quad (3-6)$$

Eq. (3-4) is a typical illumination-reflectance model of image formation [26] where $\mathbf{I}_{coarse} \cdot \mathbf{I}_{fine}$ is regarded as the global illumination and $\mathbf{T}_{sc}^2 \cdot \mathbf{O} + \mathbf{1} - \mathbf{T}_{sc}$ is regarded as the reflectance object.

Since retinal images might have good illumination conditions or suffer from different levels of uneven illumination. We first need to correct the coarse illumination pattern, making the retinal images have even illumination conditions.

Here we take advantage of the Retinex theory on illumination correction [36]. Taking logarithm to both sides of Eq. (3-4), yielding

$$\log(\mathbf{S}) = \log(\mathbf{I}_{coarse}) + \log(\mathbf{S}_0), \quad (3-7)$$

Since \mathbf{I}_{coarse} is slowly varying in the spatial domain, a good estimation of \mathbf{I}_{coarse} can be generated by filtering \mathbf{S} with a low-pass filter which is

$$\mathbf{I}_{coarse} = \mathbf{S} \otimes \left[\frac{1}{2\pi\omega^2} \exp\left(-\frac{\mathbf{r}^2}{2\omega^2}\right) \right]. \quad (3-8)$$

Here, \otimes denotes convolution. ω controls the full width at half maximum (FWHM) of the Gaussian kernel. The size of ω is determined according to the size of the input image, where $\omega = \lfloor M/20 \rfloor$, M is the width (pixels) of the input image, and $\lfloor \cdot \rfloor$ denotes round-toward-zero. Substituting Eq. (3-8) to Eq. (3-7), we have

$$S_0(r) = \exp\left(\log(S + \varepsilon) - \log\left\{S \otimes \left[\frac{1}{2\pi\omega^2} \exp\left(-\frac{r^2}{2\omega^2}\right)\right] + \varepsilon\right\}\right) - \varepsilon. \quad (3-9)$$

Here ε is a small value to avoid the situation of $\log(0)$, $\varepsilon = 0.01$ in the following studies.

Eq. (3-9) corrects the illumination in a single-color channel. By applying Eq. (3-9) to Eq. (3-7) to each color channel, an illumination-corrected image is obtained. However, due to the subtracting operation in Eq. (3-9), the RGB color of the output images will become rather different from the raw image. Therefore, the color correction treatment should be involved for each channel. Here we use the min-max approach to correct the color difference, which is given by

$$S_{0,colorcorrection} = [\max(S) - \min(S)] \cdot \left[\frac{S_0 - \min(S_0)}{\max(S_0) - \min(S_0)}\right] + \min(S), \quad (3-10)$$

where S is the corresponding color channel of the raw images.

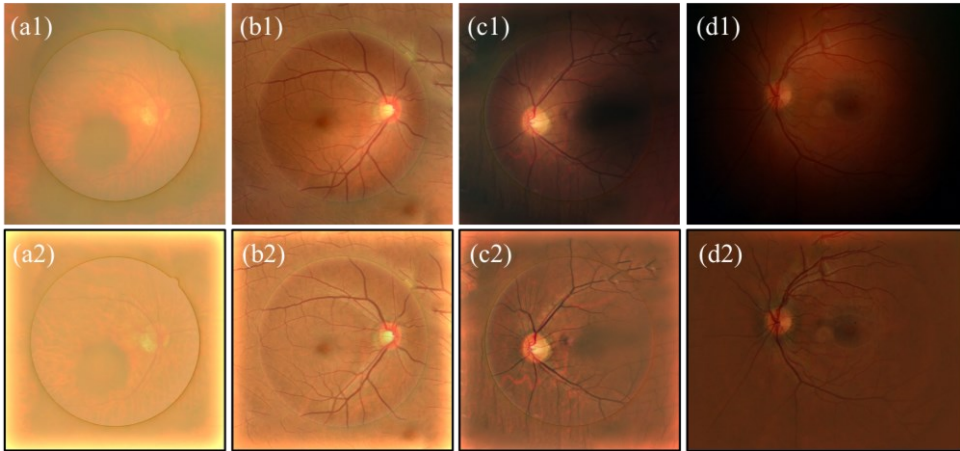


Fig. 3-3. Demonstration of step 2: coarse illumination correction. (a1) to (d1) are raw images after padding. (a2) to (d2) are outputs of coarse illumination correction.

To demonstrate the effect of the coarse illumination correction, we choose four retinal images with different degraded conditions as shown in Fig. 3-3 (a1) to Fig. 3-3

(d1). The retinal image in Fig. 3-3 (a1) denotes a good illumination condition but with a thick haze effect due to the cataract. The image in Fig. 3-3 (b1) is taken from a healthy subject with relatively good but still uneven illumination. Fig. 3-3 (c1) and (d1) denote retinal images with different uneven illumination patterns and levels of blurriness.

Images after the coarse illumination correction are shown in Fig. 3-3 (a2) to Fig. 3-3 (d2), where the uneven illumination pattern is corrected as shown in Fig. 3-3 (c2) and (d2), however, we can still find that the illuminating brightness for Fig. 3-3 (d2) is insufficient in compared with Fig. 3-3 (a2). Therefore, we enter the second step, the fine illumination boosting, to boost the overall brightness of the image.

In the following section, we will show that fine illumination boosting and scattering suppression can be unified into the same framework of dehazing and can be achieved by using the dark channel prior. However, before we continue, a background about the failure of dark channel prior on retinal images in RGB color space, and how can we overcome this challenge should be clarified.

3.2 Failure of dark channel prior on retinal images in RGB color space

The dark channel prior (DCP) and its improved version have been widely used for natural scene dehazing including underwater image enhancement and haze removal even for thick fog situations. The results of DCP dehazing are promising. Moreover, DCP can be used to correct the bad illumination condition by applying it to the color-inversed image. Applying the DCP to retinal images for dehazing seems to be natural, since (1) the visual effect of haze effect in natural scenes and retinal images are similar to each other. (2) The physical model for haze-scattering and intraocular scattering are consistent with each other. Both of them can be modeled as random phase perturbations [34, 37]. However, it is found that the performance of DCP on retinal image dehazing is limited, especially for thick cataracts. Before introducing our proposed method, the failure of the DCP in RGB retinal images should be clarified.

The principle of DCP tells us that in any haze-free image (in RGB color space), at least one pixel has zero intensity in at least one channel. This assumption does not hold for normal retinal images and retinal images with severe cataracts, due to the unique spectral property of the ocular lens. In order to protect the fundus from harmful blue-wavelength light, the human lens has high absorption for light rays of short wavelengths, which results in low intensities of the blue channel [1, 38, 39]. That is why most of the retinal images show reddish or yellowish in their photographs [40]. When the cataract is present, the absorption of short wavelengths will be even higher than the normal condition due to the presence of protein precipitation in the crystalline lens, and lens aging [41, 42].

The image formation model ignores the condition of light absorption. When DCP is applied to cataractous retinal images (searching for the minimum pixel value in red, green, and blue channels in a given small patch), the DCP will mostly extract the pixel value in the blue channel since it always tends to have a small intensity whether there is a cataract or not due to the absorption of short wavelength. As a result, the transmission maps will be over-estimated (under-estimate the haze thickness).

We perform a statistical analysis on pixel value distribution for both cataractous and normal retinal images. The pixel value of the raw images is normalized into [0,1] (divided by 255). One cataractous retinal image is shown in Fig. 3-4 (a), and Fig. 3-4 (b) shows the histogram of its blue channel, where most of the pixels in the blue channel have small gray values that are close to 0.2. Fig. 3-4 (c) shows the histogram of the blue channel for all 100 images in a public database of cataract retinal images of different degrees [43]. The pixel values in the blue channel ranged from 0 to 0.5, and most of the pixels' values are located between 0.25 to 0.5. That is mainly due to the short-wavelength absorption of the ocular lens. By applying DCP to those cataract retinal images, the pixel values in the dark channel are shown in Fig. 3-4 (d) with a range of [0, 0.5]. The distribution of the dark channel [Fig. 3-4 (d)] is similar to that of the blue

channel in Fig. 3-4 (c), as the DCP extracts the local minimum pixel value in RGB channels, and the local minimum pixel is located in the blue channels in this case.

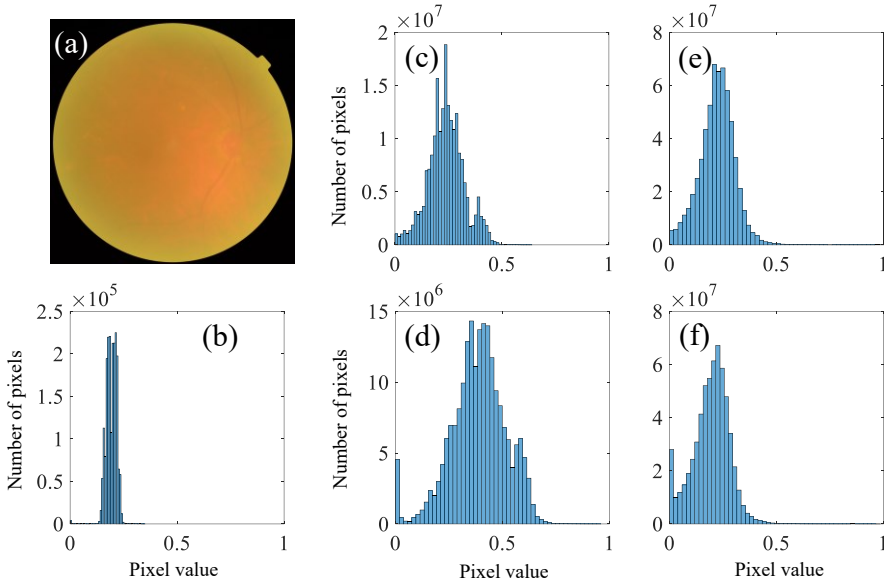


Fig. 3-4. Cataractous retinal image shown in (a) and its histograms of the blue channel shown in (b); (c) and (d) are blue channel and dark channel, respectively, for all images in the cataract retinal image database (100 images included); (e) and (f) are the blue and the dark channels for all images in the normal retinal image database (300 images included).

We have also investigated the pixel value distributions for retinal images of normal eyes. A total of 300 images are taken from the public normal retinal images database in [43]. Their blue channel and dark channel distribution are shown in Fig. 3-4 (e) and Fig. 3-4 (f), respectively. Apparently, this distribution does not agree with the DCP assumption for haze-free images, in which the pixel values should approach zero. Meanwhile, there is no distinct difference in the pixel values' distribution between cataractous and normal retinal images. As a result, the DCP will treat the cataractous retinal images as normal retinal images and over-estimates the transmission map

(under-estimates the haze density). A similar problem will happen when DCP is used to correct the illumination for retinal images in RGB color space.

However, one big advantage of the DCP is that it can be used for gray-scaled image dehazing and the result is also promising [44] if there are enough shadow regions in the image. Although this condition may not hold for retinal images, our following experimental results show its validity where the contrast is significantly improved. Accordingly, we can still take its advantage in the gray-scale image by applying DCP to a single channel of the retinal image to obtain a better estimation for the transmission map.

For expression convenience, we assume the single gray-scale channel of the retinal image is Q channel. We choose the Q channel based on three principles: (1) The Q channel makes full use of valid information of the raw image. (2) When applying DCP to the Q channel, the transmission map is obtained more properly than using DCP in RGB color space. (3) Enhancement of the color space should efficiently improve the contrast of the raw image, and won't lead to too many color distortions. Therefore, we are inspired by the color space transformation and let Q be the intensity or luminance channel of the retinal image in non-RGB color space, for example, the YUV, YCbCr, or CIE-LAB color space.

Here we choose to convert the color retinal image into YCbCr color space as its luminance channel $Y = (65.481 R + 128.553 G + 24.966 B + 16) / 255$, where Y is the linear combination of RGB channels with different weights. The green channel gains the most weight while the blue channel gains the least weight. This combination meets our requirement since that (1) it makes full use of valid information in the raw image as the red channel suffers less from the lens absorption, and the green channel includes important retinal structures like blood vessels [12]; (2) the blue channel gains the least weight which omits its impact on estimating the transmission map. Then we apply DCP to the Y channel for fine illumination boosting and dehazing, the output is transformed back to RGB color space yielding a contrast-enhanced image. Note that other color space

like CIE-LAB color space also provides promising enhancement results (see **Supplementary Note 3-2**). The HSV color space is not suitable for our requirements since the V component is given by the maximum value between RGB channels and it drops some information in one of the three color channels.

3.4 Fine illumination boosting

Equation (3-5) is rewritten as

$$\mathbf{I} - \mathbf{Y}_0(\mathbf{r}) = \mathbf{I}_{fine}(\mathbf{r}) \cdot [\mathbf{I} - \mathbf{Y}_{ic}(\mathbf{r})] + \mathbf{I} - \mathbf{I}_{fine}(\mathbf{r}), \quad (3-11)$$

where \mathbf{Y}_0 is the \mathbf{Y} channel of \mathbf{S}_0 in YCbCr color space. Equation (3-11) shows the same mathematical form as that of the hazed image formation model [26]. We therefore can apply the imaging de-hazing to Eq. (3-11) and obtain the illumination-corrected reflectance object \mathbf{Y}_{ic} . Here we use the gray-scale dark-channel prior for image de-hazing.

The estimation of \mathbf{Y}_{ic} is given by

$$\mathbf{Y}_{ic}(\mathbf{r}) = \mathbf{I} - \frac{[\mathbf{I} - \mathbf{Y}_0(\mathbf{r})] + \mathbf{I}_{fine}(\mathbf{r}) - \mathbf{I}}{\mathbf{I}_{fine}(\mathbf{r})}, \quad (3-12)$$

with

$$\mathbf{I}_{fine} \approx \mathbf{I} - \alpha \cdot (\mathbf{I} - \mathbf{Y}_0)_{dark}. \quad (3-13)$$

$(\mathbf{I} - \mathbf{Y}_0)_{dark}$ is the dark channel of $\mathbf{I} - \mathbf{Y}_0$. The dark channel \mathbf{J}_{dark} of an arbitrary gray-scale image \mathbf{J} estimated in a local neighborhood $\Omega(\mathbf{r})$ with the size of w pixels is equal to filtering the image using the local minimum filter [44]:

$$\mathbf{J}_{dark} = \min_{\rho \in \Omega(\mathbf{r})} \mathbf{J}(\rho). \quad (3-14)$$

w is the patch size, which is given by $w = \lfloor M/150 \rfloor$, according to the size of the input image. The parameter $\alpha \in [0, 1]$ is used to control the degree of illumination boosting. A large value of α will lighten the gray-scale image, and may also cause an overexposure effect where the pixel value is larger than 1. An appropriate value of α is important for

ensuring good enhancement results, we, therefore, use an adaptive manner to determine the value of α .

To determine the value of α with low complexity but high accuracy, an intensity-based global wise optimization function is designed as

$$\alpha = \arg \min \left\{ \left(\frac{\sum_{r \in \Gamma} [Y_{ic}(\mathbf{r})]}{\#pixels(\Gamma)} - 0.02 \right)^2 \right\}, \quad (3-15)$$

where the function $f[Y_{ic}(\mathbf{r})] = 1$ for $Y_{ic}(\mathbf{r}) > 1$, $f[Y_{ic}(\mathbf{r})] = 0$ for $Y_{ic}(\mathbf{r}) \leq 1$. Γ is the retinal area apart from the black background. $\#pixels(\Gamma)$ is the number of the total pixels in Γ . $\sum_r f[Y_{ic}(\mathbf{r})] / \#pixels(\Gamma)$ implies the percentage of overexposed pixels. Equation (3-15) ensures that the percentage of pixels that are overexposed during the illumination boosting should approach to 2%.

Considering that Eq. (3-15) is a one-dimensional optimization function. In this work, the Fibonacci method (FM) is adopted to solve Eq. (3-15) since it is able to gradually narrow the search interval for a one-dimensional optimization problem until the convergence condition is satisfied [45].

Continue with Fig. 3-3, the Y channel before and after illumination boosting is shown in Fig. 3-5. The first row of Fig. 3-5 is the Y channel of S_0 . The second row shows the illumination estimation, I_{fine} , generated by the DCP. As shown in Fig. 3-5 (a2) to (d2), the illumination near optical disks has a higher value than that of other areas. After the illumination boosting, images that originally have different illumination intensities are now in the same level of illumination conditions as shown in Fig. 3-5 (a3) to (d3).

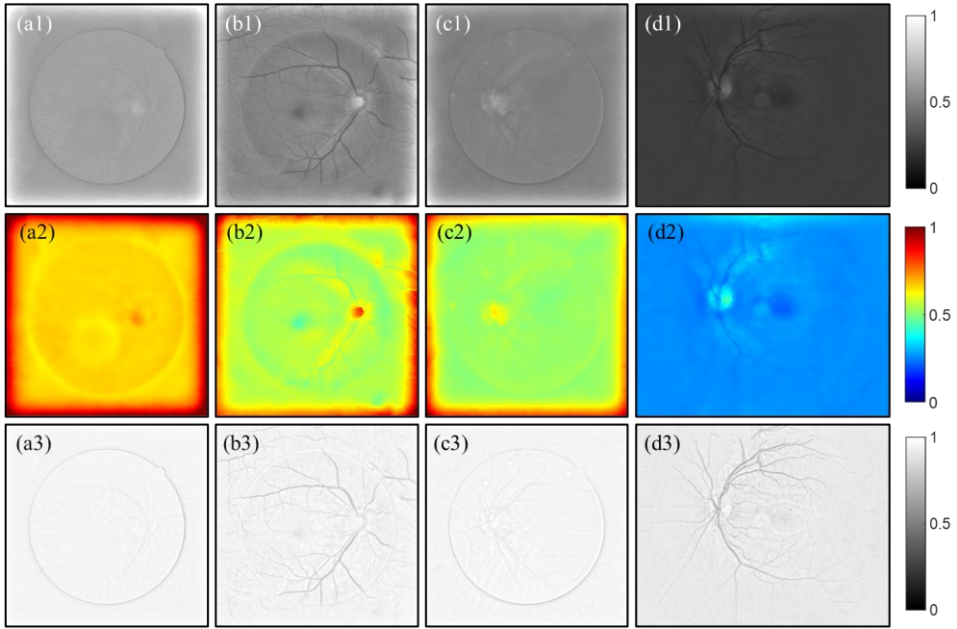


Fig. 3-5. Step 3: Fine illumination boosting in the Y channel. (a1) to (d1) are the Y channel of images in Fig. 3-3 (a2) to (d2). (a2) to (d2) are illumination estimations generated by the DCP. (a3) to (d3) are the Y channels after illumination boosting.

3.5 Scattering suppression

After the fine illumination boosting, we enter the last step to suppress the haze effect of the images. Substituting Eq. (3-12) to Eq. (3-6) yielding

$$Y_o(\mathbf{r}) = \frac{Y_{ic}(\mathbf{r}) - I}{T_{sc}^2(\mathbf{r})} + \min\left[\frac{I}{T_{sc}(\mathbf{r})}, I\right]. \quad (3-16)$$

again

$$T_{sc} \approx I - \beta \cdot \min_{\rho \in \Omega(\mathbf{r})}(Y_{ic}). \quad (3-17)$$

$Y_o(\mathbf{r})$ is the Y channel of the restored image. The minimum function in the second term of Eq. (3-16) is applied to each position of \mathbf{r} to prevent the image from overexposure effect due to the large value in $1/T_{sc}(\mathbf{r})$. The squared term in the denominator of Eq.

(3-16) shows the concept of double-pass. Due to the squared term of $T_{sc}(\mathbf{r})$, the DPFR model has a stronger haze suppression ability than that of the image formation model for natural scenes.

Parameter β is used to control the dehazing degree. A large value of β will have a less haze effect, but the intensity is also low. An appropriate value of β is also determined using an adaptive manner. Since the increase of β always decreases the value of $Y_o(\mathbf{r})$ for a given $Y_{ic}(\mathbf{r})$, a global wise optimization function for the determination of β is designed as

$$\beta = \arg \min \left\{ \left| 0.55 - \text{mean} [Y_o(\mathbf{r} \in \Gamma)] \right| \right\}, \quad (3-18)$$

Equation (3-18) ensures that the mean intensity of $Y_o(\mathbf{r})$ should approach 0.55. Again, the Fibonacci method (FM) is adopted to solve Eq. 18.

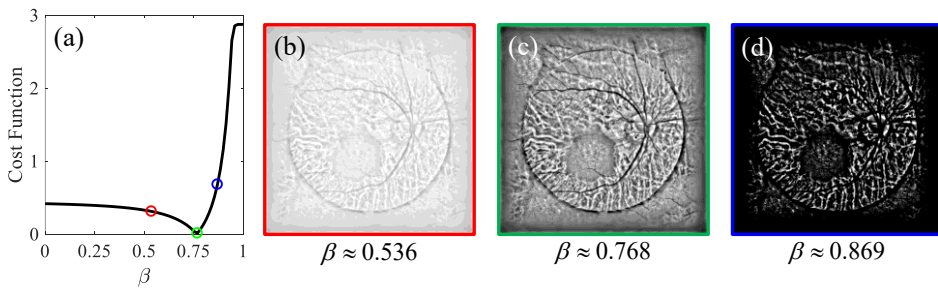


Fig. 3-6, Change of optimization function Eq. (3-18) with respect to β . (a) Value of the function. (b), (c) and (d) are $Y_o(\mathbf{r})$ patterns with β corresponding to the red, green, and blue circle in (a)

Fig. 3-6 is an example that shows how $\left| 0.55 - \text{mean} [Y_o(\mathbf{r} \in \Gamma)] \right|$ changes with respect to β , and the corresponding Y_o for the image in Fig. 3-5 (a3). As shown in Fig. 3-6 (c), when $\beta = 0.754$, Eq. (3-18) has the minimum value as shown in the green circle in Fig. 3-6 (a). A smaller or larger value of β will decrease the visual quality of the image as shown in Fig. 3-6(b) for $\beta = 0.478$, and in Fig. 3-6 (d) for $\beta = 0.841$.

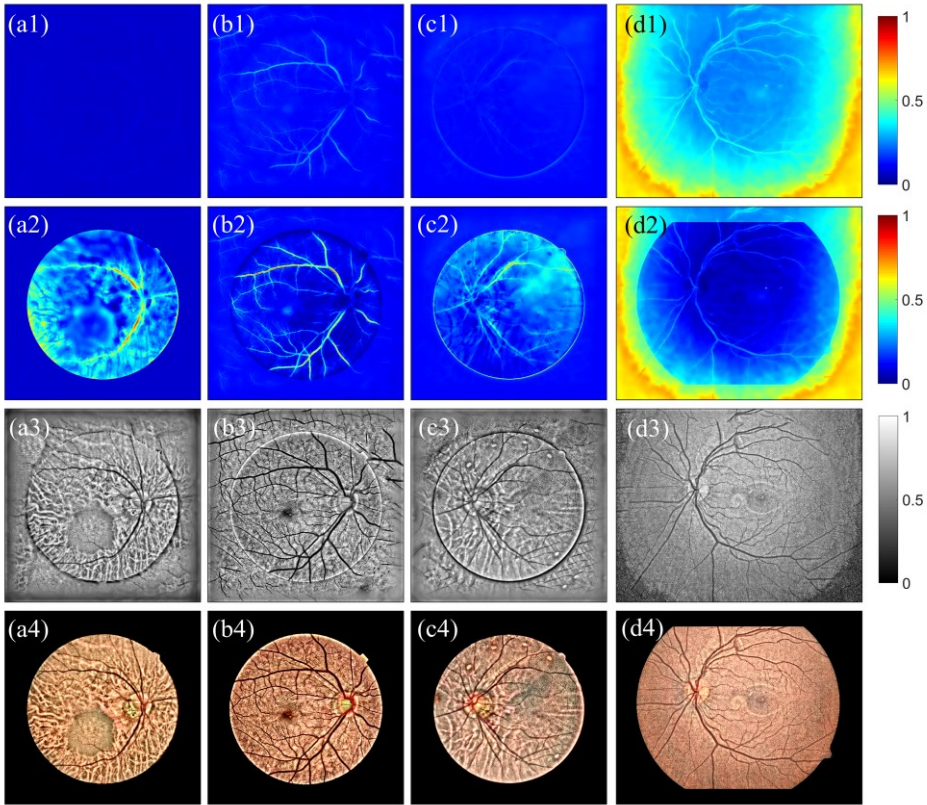


Fig. 3-7 Step 4: Scattering suppression in the Y channel. (a1) to (d1) are Y channel of images in Fig. 3-3 (a2) to (d2). (a1) to (d1) are estimations of T_{sc}^2 generated by the DCP. (a2) to (d2) are intensity normalized pattern of (a1) to (d1). (a3) to (d3) are Y channels after dehazing. (a4) to (d4) are corresponding outputs in RGB color space.

Followed by Fig. 3-5, results of scattering suppression are shown in Fig. 3-7. The first row in Fig. 3-7 is the estimation of T_{sc}^2 generated by the DCP. As shown in Fig. 3-7 (a1) and (c1), the images have low transmission maps due to the presence of cataracts. Since the intensity of the pattern of T_{sc}^2 is low for haze retinal images, in order to have a clear observation of the transmission pattern, we normalized the intensity in the area

Γ as shown in the second row of Fig. 3-7, where the blood vessels are regarded as having high transmission.

The Y channel after scattering suppression is shown in the second row of Fig. 3-7, where the contrast is significantly enhanced in comparison with that in Fig. 3-5. The YCbCr images are transformed back to the RGB color space giving us the final enhanced images as shown in the third row of Fig. 3-6. The padding area is clear.

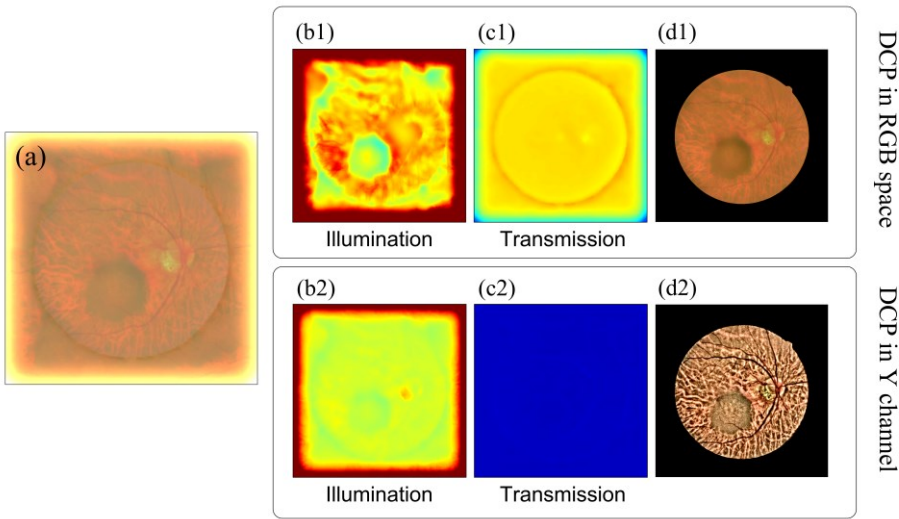


Fig. 3-8. Comparison of DCP in RGB color space and in Y channel. The intensity normalized map for (c2) is shown in Fig. 3-7 (a2).

Taking Fig. 3-3 (a2) as an example, we demonstrated the failure of DCP in RGB color space. As shown in Fig. 3-8 when the DCP is applied to the RGB color space, the DCP over-estimates the transmission illumination and the transmission maps (pixels value are large), as it mostly extracts the pixels in the blue channel of the image. The haze effect in the final output is still obvious as shown in Fig. 3-8 (d1). While the DCP works well in the Y channel of the image as shown in Fig. 3-8 (b2) to (d2). The transmission map has small pixel values denoting the heavy haze effect in the raw image of Fig. 3-8 (a).

As mentioned in section 3.2, theoretically the DCP might be invalid for a gray-scale image if there is not enough shadow region as shown in Fig. 3-5 (a3) to Fig. 3-5 (c3), however, our results show its validity. There are two inner mechanisms that work together to make the DCP valid in our case. (1) First, as shown in Fig. 3-5 (a3) to Fig. 3-5 (c3), the pixel values approach to the global atmosphere color (which is 1 in this case) and will lead to the under-estimation of DCP on T as $T = (1 - \beta \mathbf{I}_{dark}) \rightarrow 0$, in other words, the DCP over-estimating the haze density. Since the values in T are small, after applying Eq. (3-16) the contrast of the image is largely increased, but many over-enhanced pixels are also produced where they have negative gray values and decrease the image's quality. This effect can be seen in Fig. 3-6 (d) where $\beta = 0.869$. (2) Thanks to our intensity-based adaptive determination of parameter β , the impact of under-estimation of T will be canceled when a small value of β is assigned. These two inner mechanisms balance each other and make DCP valid in our case.

A flow chart for the proposed enhancement method is summarized in Fig. 3-9. In general, our DPFR method employs Retinex theory and DCP image dehazing to achieve illumination correction and haze suppression, simultaneously.

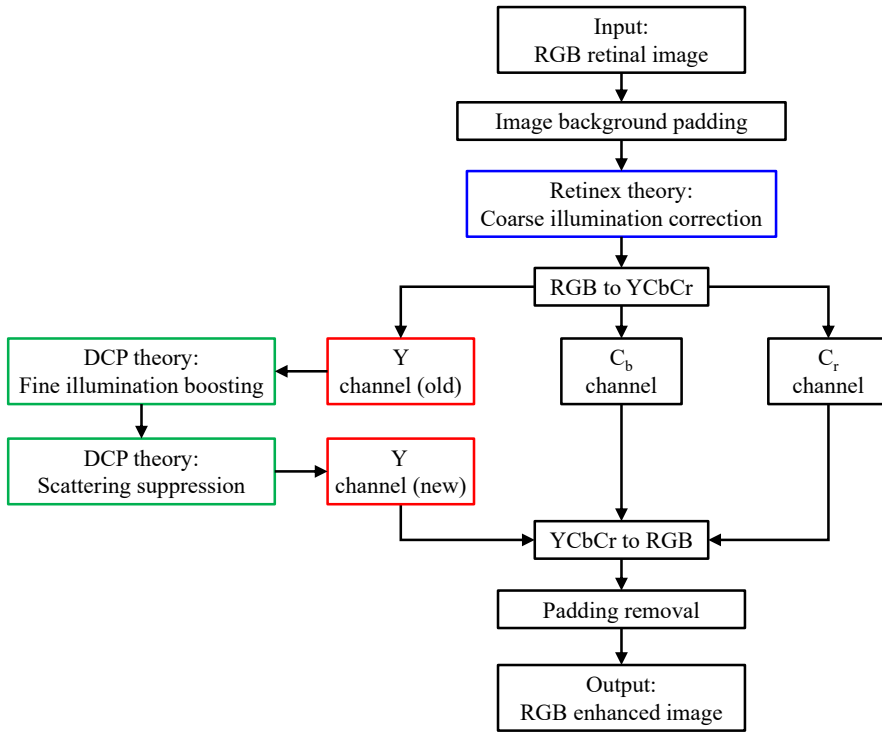


Fig. 3-9. Flow chart for DPF method.

4. Experimental results

This section compares the enhancement results from conventional methods and the proposed method qualitatively and quantitatively to verify the enhancement performance of the proposed method. We applied the proposed method (DPFR) and three start-of-art including the luminosity and contrast adjustment method (LCA) [15], the low-pass filtering and α -rooting (LPA) [16], and the pixel color amplification method (PCA) [29]. In [16], it is proved that the enhancement performance of LPA is better than methods in [8] and [28].

In Section 4.1, we compared the results to assess the performance of the proposed method in terms of removing the haze effect and enhancing retinal structure visibility

with images in Fig. 3-3 (a1) to Fig. 3-3 (d1). In Section 4.2, the performance of the proposed method is validated with various metrics for retinal images in Fig. 3-3 (a1) to 3(d1). Lastly, we compared the enhancement results of the proposed algorithm and the conventional methods with a total of 906 retinal images from the Digital Retinal Images for Vessel Extraction (DRIVE-) [46], STructured Analysis of the Retina (STARE-) [47, 48], Standard Diabetic Retinopathy Database Calibration level 1 (DiaRetDB1-) [49, 50], Cataract- [43], and Normal- images databases.

4.1 Qualitative evaluations with images in Fig. 3-3(a1) to 3-3(d1)

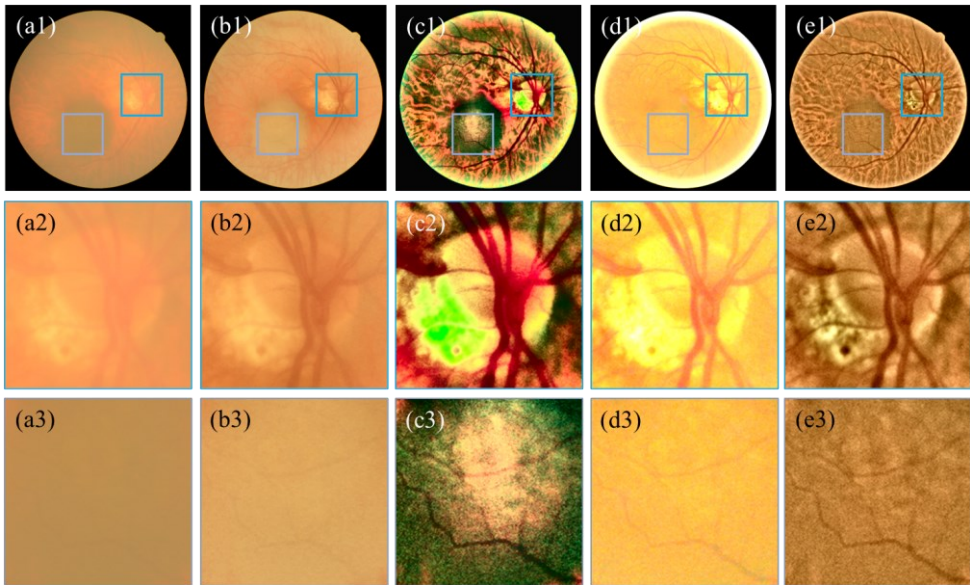


Fig. 3-10. Enhancement results for the image in Fig. 3-3(a1). Images in the first row are raw image, enhanced results of LCA, LPAR, PCA, and DPFR methods, respectively. The second third rows are enlarged parts in the corresponding yellow and magenta boxes.

Enhancement results for Fig. 3-3 (a1) with different methods are shown in Fig. 3-10, where Fig. 3-10 (a1) to Fig. 3-10 (e1) are raw images, the results of LCA, LPAR, PCA, and DPFR method, respectively. The image is obtained from the cataract database, and shows blurriness due to cataract [Fig. 3-10 (a1)]. An area near the optical disk is

enlarged to show the detailed structure as shown in Fig. 3-10 (a2). The blood vessels can be hardly observed due to the haze effect of cataracts.

The results of LCA method are shown in Fig. 3-10 (b1). As shown in Fig. 3-10 (b2) and Fig. 3-10 (b3), the image contrast is slightly improved, and some blood vessels in Fig. 3-10 (b3) can be observed faintly. The haze effect is still dominant.

The results of LPAR method are shown in Fig. 3-10 (c1), where the haze effect is removed as the blood vessels in Fig. 3-10 (c2) and Fig. 3-10 (c3) can be clearly observed. However, the LPAR causes an overexposure effect on the bright area of the optical disk as shown in Fig. 3-10 (c2).

Fig. 3-10 (d1) to Fig. 3-10 (d3) are the results of PCA method. Since it uses the DCP in RGB space, the haze effect is not removed, and there is also an overexposure effect on the bright area of the optical disk as shown in Fig. 3-10 (d2).

Results of DPFR method are shown in Fig. 3-10 (e1) to Fig. 3-10 (e3). The haze effect is efficiently suppressed as shown in Fig. 3-10 (e2) and (e3). Blood vessels that are hidden behind the cataract can be now observed. Meanwhile, the DPFR method won't lead to overexposure effect, as it can restore a clear optical disk area as shown in Fig. 3-10 (e2).

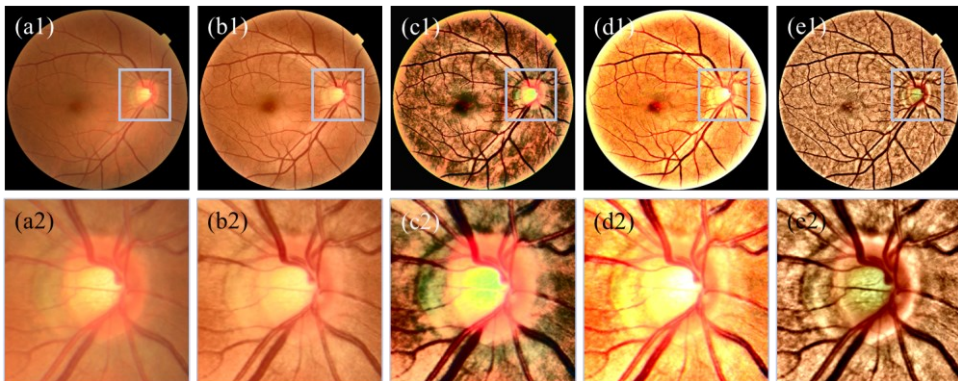


Fig. 3-11. Enhancement results for the image in Fig. 3-3(b1). (a1) to (e1) are raw images, results of LCA, LPAR, PCA and DPFR, respectively. (a2) to (e2) are enlarged parts in the optical disk areas in corresponding white boxes.

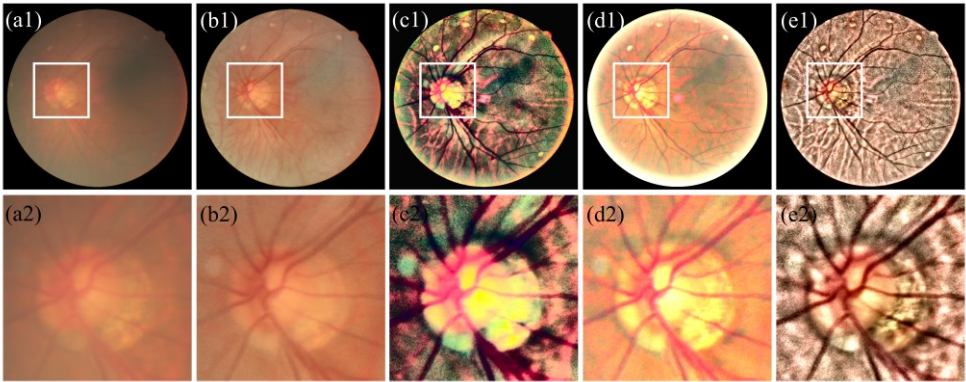


Fig. 3-12. Enhancement results for image in Fig. 3-3(c1). (a1) to (e1) are raw image, results of LCA, LPAR, PCA and DPFR, respectively. (a2) to (e2) are enlarged parts in the optical disk areas in corresponding white boxes.

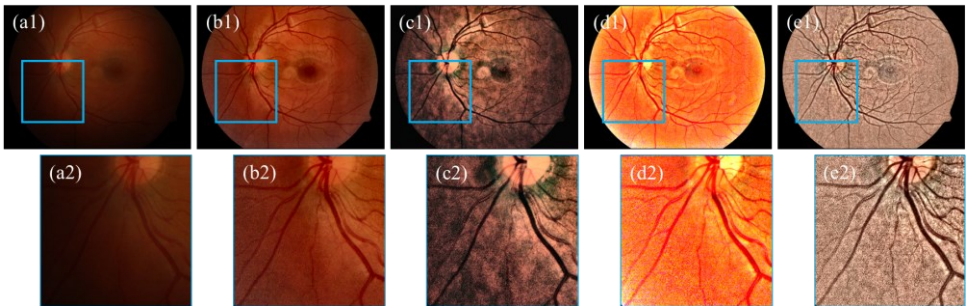


Fig. 3-13. Enhancement results for the images in Fig. 3-3(d1). (a1) to (e1) are raw images, results of LCA, LPAR, PCA and DPFR, respectively. (a2) to (e2) are enlarged parts in the optical disk areas in corresponding gray boxes.

Enhancement results for Fig. 3-3 (a2) are shown in Fig. 3-11. The image is obtained from a healthy subject from the NORMAL database [43]. We also enlarge the optical disk area as shown in Fig. 3-11 (a2), where the retinal structure can be clearly observed. Both four methods show their ability in contrast enhancement, while the LPAR and PCA methods lead to overexposure effect at the bright area as shown in Fig. 3-11 (c2) and Fig. 3-11 (d2). The LPAR seems to intensify the dark area of the retinal image (near the macular area) as shown in Fig. 3-11 (c1) and the illumination of the entire image is still

uneven. Our proposed DPFR method enhances image contrast and some small blood vessels within the optical disk can be observed. The uniform illumination is also achieved.

Enhancement results for Fig. 3-3 (c1) and Fig. 3-3 (d1) are shown in Fig. 3-12 and Fig. 3-13, respectively. The performance of LCA method is limited since it uses only the CLAHE method to enhance the contrast as shown in Fig. 3-12 (b1) and Fig. 3-13 (b1). Note that the LPAR also employs the CLAHE for contrast refinement. It uses the Retinex method to suppress the haze effect, however, its illumination correct ability is limited and can lead to an overexposure effect at the bright optical disk area, similar results can be also found in [16].

The PCA uses DCP in RGB color space to correct the illumination and suppress the haze effect as shown in Fig. 3-12(d1) and Fig. 3-12 (d1). However, due to the failure of DCP in RGB color space for retinal images, the PCA method shows limited performance. Meanwhile, the image formation model of PCA ignores the optical truth of double pass fundus reflection, which further limited its performance of enhancement. Accordingly, the DPFR method shows its ability on illumination correction and haze effect suppression as shown in Fig. 3-12 (e1) and Fig. 3-13(e1).

4.2 Quantitative evaluations for experimental results

In this sub-section, we conducted quantitative evaluations of the enhancement results in terms of (1) image definition, (2) image sharpness, (3) image local contrast, (4) image entropy, (5) image multiscale contrast, and (6) image fog density. All images are formatted in 8-bit unsigned integers (uint8, the gray-value is ranged in [0, 255]).

The image definition implies the richness of texture information [28, 45] which is given by

$$DE = \sum_{c \in \{R, G, B\}} \lambda_c \cdot \frac{1}{(M-1)(N-1)} \sum_{x=1}^M \sum_{y=1}^N \sqrt{\frac{|\nabla I_c|^2}{2}}, \quad (3-19)$$

where I_c is the c channel of input images. ∇ is the gradient operator. The DEs are evaluated in each RGB channel and are combined linearly with coefficients λ_c since the different color channel has different visual response. $\lambda_R = 0.299$, $\lambda_G = 0.587$ and $\lambda_B = 0.114$ are used according to the relative visual responses of the red, green, and blue channels [51].

Sharpness is the attribute related to the preservation of fine details and edges. For images captured under the water, severe blurring occurs due to the forward scattering. This blurring effect causes degradation of image sharpness.

The image sharpness is adapted from the underwater image sharpness measure (UISM) [51], since severe blurring occurs to both underwater image and retinal images due to the forward scattering. Moreover, since the UISM does not rely on the statistical property of images, it can be applied to retinal images, regardless of the statistical difference between retinal images and underwater images.

To measure the sharpness on edges, the Sobel edge detector is first applied on each RGB color channel. The resultant edge map is then multiplied with the original image to get the grayscale edge map in the corresponding color channel. By doing this, only the pixels on the edges of the original retinal image are preserved. The UISM is given by

$$UISM = \sum_{c \in \{R, G, B\}} \lambda_c \cdot \left[\frac{2}{k_1 k_2} \sum_{l=1}^{k_1} \sum_{k=1}^{k_2} \log \left(\frac{I_{\max, k, l}}{I_{\min, k, l}} \right) \right], \quad (3-20)$$

where the image is divided into $k_1 k_2$ blocks, $I_{\max, k, l} / I_{\min, k, l}$ indicates the relative contrast ratio within each block. Images are divided into 10×10 blocks, viz $k_1 = k_2 = 10$.

For retinal images, contrast degradation is usually caused by intraocular scattering, which has the same visual effect as that of fog scenes, or underwater images. Therefore, the Underwater Image Contrast Measure (UIConM) [51] can be also employed for evaluating the contrast of the retinal image. Also, the UIConM does not rely on the statistical property of images. The UIConM is given by

$$UIConM = \sum_{c \in \{R, G, B\}} \lambda_c \cdot \log \left[\frac{1}{k_1 k_2} * \sum_{l=1}^{k_1} \sum_{k=1}^{k_2} \frac{I_{\max, k, l} \ominus I_{\min, k, l}}{I_{\max, k, l} \oplus I_{\min, k, l}} \log \left(\frac{I_{\max, k, l} \ominus I_{\min, k, l}}{I_{\max, k, l} \oplus I_{\min, k, l}} \right) \right], \quad (3-21)$$

where an image is divided into blocks, and \oplus , $*$, and \ominus are the PLIP operations in which $A \oplus B = A + B - A \cdot B / \gamma(M)$, $A * B = \gamma(M) - \gamma(M) \cdot [1 - B / \gamma(M)]^A$, and $A \ominus B = \kappa(M) \cdot (A - B) / [\kappa(M) - B]$. $\gamma(M) = 1024$ and $\kappa(M) = 600$ are chosen in this study. Images are divided into 10×10 blocks, viz $k_1 = k_2 = 10$.

Image entropy (IE) describes the randomness distribution of the image and its value denotes the amount of image information [52, 53], which is given by

$$IE = \sum_{c \in \{R, G, B\}} \lambda_c \cdot \sum_{g=0}^{255} P(x_g) \log(x_g), \quad (3-22)$$

where $P(x_g)$ is the probability of the appearance of the pixels that have gray-value g in the gray-scaled image. Image entropy can be used to characterize the texture of the image [40] and determine the amount of image information. Images affected by haze tend to have low IE values due to the biased brightness distribution. In contrast, haze-free images have a relatively high IE.

The multiscale-contrast of the image, C_{RAMM} , is calculated with a pyramidal multi-resolution representation of luminance [54]. C_{RAMM} is defined as

$$C_{RAMM} = \frac{1}{8(\# \text{ pixel})(\# \text{ levels})} \sum_{\forall \text{ level}} \left[\sum_{\forall \text{ pixel}} \left(\sum_{8\text{-neigh}} |P_i - P_j| \right) \right], \quad (3-23)$$

where $\# \text{ pixel}$ denotes the total pixel numbers of the image, $\forall \text{ level}$ denotes the total level of down-sampling, in each level the image is halved without pre-filtering. In this paper, we use 6 levels of the down-sampling. The pixel numbers of retinal images in the databases are large enough for 6 levels of down-sampling. During the calculation, the down-sampling process is stopped if either the row pixel numbers, or the column pixel numbers of the next down-sampled image are less than 3.

Lastly, the fog aware density evaluator (FADE) [55] is used to numerically predict perceptual hazy density. MATLAB code for the FADE is available in [56].

Tab. 3-1. Quantitative assessment of the enhanced images among Fig. 3-3(a1) to 3-3(d1) with six metrics.

Raw image	Matrixes	Raw	Methods			
			LCA	LPAR	PCA	DPFR
Figure 3(a1)	DE	0.7457	1.3930	8.5798	2.4048	11.3595
	UISM	3.3387	3.5016	7.1968	3.8122	6.3865
	UIConM	0.0870	0.1137	0.3605	0.2009	0.39273
	IE	3.7783	3.9032	5.0146	2.9855	5.0237
	C _{RAMM}	2.0623	2.6328	8.2115	3.9342	8.476
	FADE	1.0455	0.7261	0.2143	0.5133	0.17273
Figure 3(b1)	DE	1.2298	2.8213	8.4657	5.6587	13.6305
	UISM	4.4821	5.2004	7.2436	7.1333	7.0683
	UIConM	0.1025	0.1990	0.3563	0.3558	0.39856
	IE	4.3621	4.6008	5.0771	3.8008	5.0585
	C _{RAMM}	2.2474	3.7691	8.7034	6.5540	9.9654
	FADE	0.4719	0.3182	0.1703	0.2292	0.13138
Figure 3(c1)	DE	0.6909	1.7769	12.4526	3.7974	14.1229
	UISM	3.0368	3.7516	7.3951	4.3700	5.2604
	UIConM	0.0558	0.1050	0.3722	0.2513	0.38931
	IE	4.3999	4.1133	5.2492	4.2078	5.154
	C _{RAMM}	1.5730	2.4297	8.1420	4.5733	8.3505
	FADE	1.1586	0.6600	0.2329	0.4872	0.19138
Figure 3(d1)	DE	0.9049	4.0778	10.1221	10.1379	18.2453
	UISM	2.9964	5.1093	4.9589	5.5718	5.4324
	UIConM	0.0164	0.0900	0.2105	0.3087	0.37906
	IE	4.0066	4.3095	5.0064	3.6781	5.0386
	C _{RAMM}	0.6300	2.0243	5.0777	4.1120	6.6639
	FADE	0.4850	0.2389	0.1494	0.1496	0.17417
Average	DE	0.8928	2.5172	9.9051	5.4997	13.8615
	UISM	3.4635	4.3907	6.6986	5.2218	6.0576
	UIConM	0.0654	0.1269	0.3249	0.2792	0.3837
	IE	4.1367	4.2317	5.0868	3.6681	5.05362
	C _{RAMM}	1.6281	2.7140	7.5337	4.79338	8.1211
	FADE	0.7903	0.4858	0.1917	0.3448	0.1673

Tab. 3-1 demonstrates the enhancement results of the quantitative evaluations using the objective metrics. Here, High values of DE, UISM, UIConM, IE, and C_{RAMM}

indicate high image quality, while low values of FADE indicate a better dehazing performance.

As listed in Tab. 1, the value of FADE is large for retinal images with severe haze effects as shown in the raw images of Fig. 3-3(a1) and Fig. 3-3 (c1). Here FADE = 1.05 for Fig. 3-3 (a1) and FADE = 1.16 for Fig. 3-3(c1). For the normal retinal image as shown in Fig. 3-3 (b1), the FADE is low as FADE = 0.47.

Numbers marked in blue denote the best performance, while numbers marked in green denote the second-best performance. Among the four retinal images in Fig. 3-3 (a1) to Fig. 3-3 (d1), the proposed DPFR method shows the best performance for DE, UIConM, C_{RAMM} , and FADE, on average. It also shows the second-best performance for UIISM, IE. The LPAR method has the highest values for UIISM, and IE among the four retinal images.

Additional experimental results for both qualitative and quantitative assessments are available in the first section of the **Supplementary Note 3-1**.

4.3 Quantitative evaluations for public databases

Tab. 2 demonstrates the enhancement results of the quantitative evaluation for a total of 906 images from five databases. In general, the DPFR model is able to achieve better image quality including image definition, image sharpness, and image contrast than current state-of-art methods. The haze density of the retinal image is efficiently suppressed as the results of DPFR method have the smallest value of FADE for all five databases.

Note that the DPFR method has not yet used the CLAHE method to improve the contrast, it can be imaged that the performance of the DPFR method can be further improved by combined with the CLAHE.

Tab. 3-2. Quantitative assessment of the enhanced images for five databases with six metrics.

Databases	Matrixes	Raw	Methods			
			LCA	LPAR	PCA	DPFR
DRIVE (20 images)	DE↑	1.7826	3.6015	12.3326	7.3283	16.9964
	UISM↑	4.7703	5.3879	9.0585	8.4717	9.6539
	UIConM↑	0.1375	0.2027	0.3679	0.3719	0.4055
	IE↑	3.9502	4.2225	4.8263	3.4884	5.0333
	C _{RAMM} ↑	2.9338	4.1783	8.6381	7.0552	9.2348
	FADE↓	0.446	0.2956	0.1819	0.2804	0.1206
STARE (397 images)	DE↑	1.4395	2.8692	8.9126	4.9638	12.9468
	UISM↑	5.4539	6.2693	8.4133	7.3555	8.5576
	UIConM↑	0.1499	0.2108	0.3549	0.3529	0.4006
	IE↑	4.1664	4.2398	4.4953	3.0822	4.8957
	C _{RAMM} ↑	2.7915	4.0455	8.29	6.4792	9.4374
	FADE↓	0.4716	0.2855	0.1843	0.2746	0.1299
DiaRet DB1 (89 images)	DE↑	0.9119	2.9236	10.8005	7.0287	16.2699
	UISM↑	2.8892	4.4872	5.7884	5.4441	5.6445
	UIConM↑	0.03	0.1008	0.2653	0.2857	0.3824
	IE↑	3.8249	4.1409	4.8895	3.0856	4.9728
	C _{RAMM} ↑	0.9539	2.1664	7.0132	4.105	7.5133
	FADE↓	0.4715	0.2775	0.1604	0.159	0.1442
Cataract (100 images)	DE↑	0.7878	1.7825	10.6005	3.8732	13.0668
	UISM↑	3.4133	3.8432	7.1361	4.9715	7.0112
	UIConM↑	0.075	0.121	0.355	0.2623	0.38377
	IE↑	3.8983	3.9858	5.0023	3.6359	4.8861
	C _{RAMM} ↑	1.8511	2.6709	8.0554	4.7462	7.9459
	FADE↓	0.9234	0.5757	0.2098	0.4029	0.16594
Normal (300 images)	DE↑	1.0183	2.3513	8.6007	4.9219	12.815
	UISM↑	3.9204	4.6155	7.388	6.2853	7.6355
	UIConM↑	0.0859	0.1601	0.3542	0.3323	0.40141
	IE↑	4.1601	4.323	5.0108	3.7073	5.0015
	C _{RAMM} ↑	2.0443	3.2959	8.4536	5.9017	9.335
	FADE↓	0.6569	0.4308	0.1862	0.3112	0.13741
Average	DE↑	1.1880	2.7056	10.2494	5.6232	14.4190
	UISM↑	4.0894	4.9206	7.5569	6.5056	7.7005
	UIConM↑	0.0957	0.1591	0.3345	0.3210	0.3947
	IE↑	4.0000	4.1824	4.8448	3.3999	4.9579
	C _{RAMM} ↑	2.1149	3.2714	8.0901	5.6575	8.6933
	FADE↓	0.5939	0.3730	0.1845	0.2856	0.1396

4.4 Enhancement of cataractous retinal images

Due to the ability of scattering suppression of the DPFR model, it is able to enhance the image contrast for retinal images with thick haze effects. To demonstrate this, we apply the DPFR model to two groups of cataractous retinal images and show the results before and after the restoration.

Fig. 3-14 shows the first group of restoration results for cataractous retinal images. Where the raw images in Fig. 3-14 (a1) to Fig. 3-14 (c1) have different levels of haze effect. As shown in Fig. 3-14 (c1), the raw image suffers from both uneven illumination and haze blurriness, where the blood vessels far from the optical disk can be hardly observed. Fig. 3-14 (a2) to Fig. 3-14 (c2) show the corresponding restoration results. The blood vessels can be clearly observed, and the restored images have better illumination conditions than that of raw images.

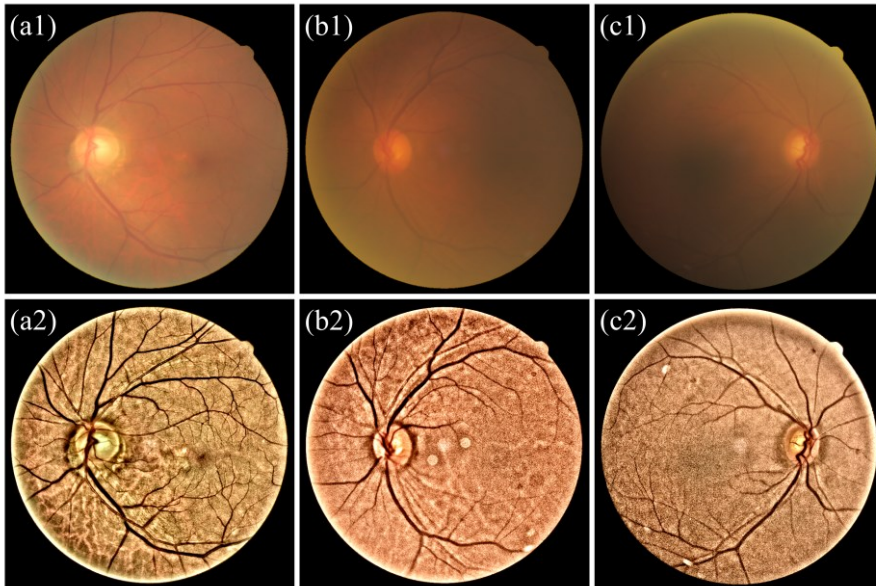


Fig. 3-14. Restoration of cataractous retinal images. First row: raw images. Second row: restored images.

Tab. 3-3. FADE for cataractous retinal images before and after restoration in Fig. 3-14.

FADE	Raw image			Restoration		
	(a1)	(b1)	(c1)	(a2)	(b2)	(c2)
	0.8484	0.8408	0.8799	0.1452	0.1526	0.1706

Tab. 3-3 lists the FADE value for raw and enhanced images, where the DPFR model decreases the fog density as the FADE value of the enhancement image is smaller than that of raw images.

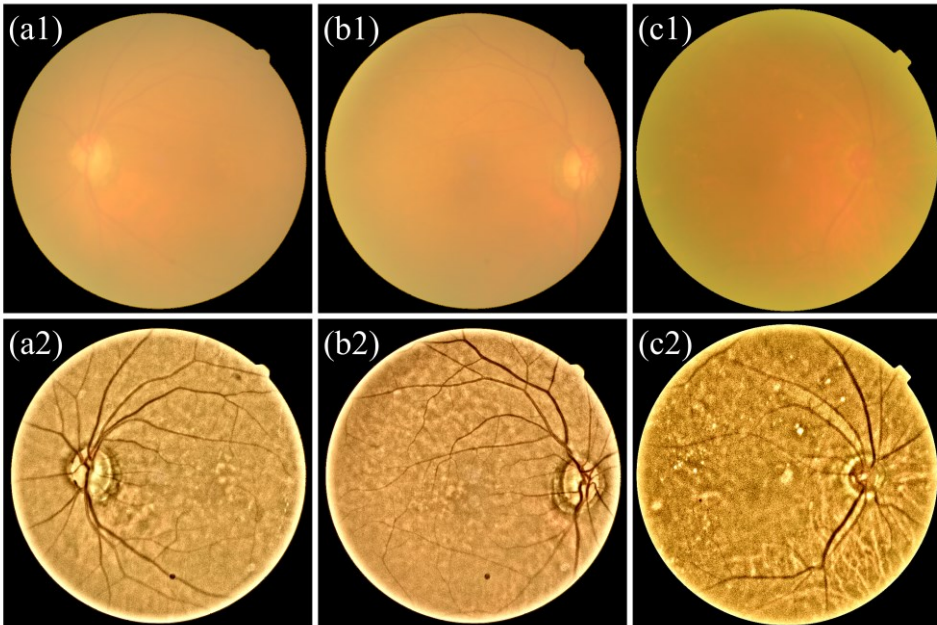


Fig. 3-15. Restoration of cataractous retinal images. First row: raw images. Second row: restored images.

Fig. 3-15 shows the second group of restoration results for cataractous retinal images. All images in this group have good illumination conditions, however, the blurriness is even worse than that of the first group. As shown in Fig. 3-15 (b1) and Fig. 3-15 (c1), only the blood vessels near the optical disk can be faintly observed. After the

restoration, as shown in the second row of Fig. 3-15, the haze effects are efficiently suppressed, as the blood vessels and their small branches can be clearly observed.

Tab. 3-4 lists the FADE value for raw and enhanced images in Fig. 3-15. In general, the DPFR model can reduce the fog density of cataractous images.

Tab. 3-4. FADE for cataractous retinal images before and after restoration in Fig. 3-15.

	Raw image			Enhancement		
	(a1)	(b1)	(c1)	(a2)	(b2)	(c2)
FADE	1.2509	1.1669	0.8827	0.2245	0.2239	0.1589

5. Concluding remarks

In this paper, we briefly review the formation of retinal images and derive the imaging formation model, particularly for retinal image enhancement, according to the light propagation in the fundus camera system. The updated model is called the double-pass fundus reflection (DPFR) model, in which the property of the special fundus reflection is taken into consideration.

The DPFR model embedded the illumination correction and image de-hazing into a single image formation formula Eq. (3-3), and the mathematical derivation shows that the retinal image restoration can be simply achieved by the combination of the Retinex theory and two steps of the dark channel prior de-hazing process. The Retinex theory is to correct the coarse uneven illumination pattern. After that, the resultant images are converted into the YCbCr color space in order to bypass the failure of DCP in RGB color space. During the DCP dehazing process, the first de-hazing process is to boost the illumination of the entire image, and the second de-hazing is to reduce the haze-effect caused by intraocular scattering.

Unlike previous restoration methods for retinal images where the natural scene image formation model is directly employed, our DPFR is physically self-consistent which obeys the wave propagation within a fundus imaging system. All parameters in

the DPFR model are automatically determined in an adaptive manner, which makes it efficient and simple to be applied to different databases.

Our DPFR model significantly improves the contrast of the retinal images and reveals retinal structures that might be ignored by clinical doctors due to insufficient contrast of the raw images. Under the guidance of the raw retinal images, the enhanced image may improve the diagnosing accuracy for ophthalmologists, and facilitate diagnoses, such as drusen, hard exudates, bleeding hemorrhages, or microaneurysm detection, as these tasks may use more color or local detail information. The early detection of these findings may improve the outcome of treatment of retinal diseases like age-related macular degeneration, diabetic retinopathy, and venous occlusions. We plan to pilot test this in a follow-up study using clinical data.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding. This research is supported by China Scholarship Council (CSC) (201908340078)

References

1. T. T. J. M. Berendschot, P. J. DeLint, and D. v. Norren, "Fundus reflectance—historical and present ideas," *Progress in retinal and eye research* **22**, 171-200 (2003).
2. J. Zhang, B. Dashtbozorg, F. Huang, T. T. J. M. Berendschot, and B. M. ter Haar Romeny, "Analysis of Retinal Vascular Biomarkers for Early Detection of Diabetes," in *VipIMAGE 2017*, J. M. R. S. Tavares and R. M. Natal Jorge, eds. (Springer International Publishing, Cham, 2018), pp. 811-817.
3. M. C. V. S. Mary, E. B. Rajsingh, and G. R. Naik, "Retinal Fundus Image Analysis for Diagnosis of Glaucoma: A Comprehensive Survey," *IEEE Access* **4**, 4327-4354 (2016).
4. M. Miura, M. Yamanari, T. Iwasaki, A. E. Elsner, S. Makita, T. Yatagai, and Y. Yasuno, "Imaging Polarimetry in Age-Related Macular Degeneration," *Investigative Ophthalmology Visual Science* **49**, 2661-2667 (2008).
5. M. Trieschmann, F. J. van Kuijk, R. Alexander, P. Hermans, P. Luthert, A. C. Bird, and D. Pauleikhoff, "Macular pigment in the human retina: histological evaluation of localization and distribution," *Eye* **22**, 132-137 (2008).
6. T. Theelen, T. T. J. M. Berendschot, C. B. Hoyng, C. J. Boon, and B. J. Klevering, "Near-infrared reflectance imaging of neovascular age-related macular degeneration," *Graefes Arch Clin Exp.Ophthalmol* **247**, 1625-1633 (2009).
7. S. Zafar, J. McCormick, L. Giancardo, S. Saidha, A. Abraham, and R. Channa, "Retinal Imaging for Neurological Diseases: "A Window into the Brain"," **59**, 137-154 (2019).
8. L. Xiong, H. Li, and L. Xu, "An enhancement method for color retinal images based on image formation model," *Computer Methods and Programs in Biomedicine* **143**, 137-150 (2017).
9. J. Mukherjee and S. K. Mitra, "Enhancement of Color Images by Scaling the DCT Coefficients," *IEEE Transactions on Image Processing* **17**, 1783-1794 (2008).
10. X. Bai, F. Zhou, and B. Xue, "Image enhancement using multi scale image features extracted by top-hat transform," *Optics & Laser Technology* **44**, 328-336 (2012).
11. M. Liao, Y.-q. Zhao, X.-h. Wang, and P.-s. Dai, "Retinal vessel enhancement based on multi-scale top-hat transformation and histogram fitting stretching," *Optics & Laser Technology* **58**, 56-62 (2014).
12. A. W. Setiawan, T. R. Mengko, O. S. Santoso, and A. B. Suksmono, "Color Retinal Image Enhancement using CLAHE," *International Conference on ICT for Smart Society* (2013).
13. B. Gupta and M. Tiwari, "Color retinal image enhancement using luminosity and quantile based contrast enhancement," *Multidimensional Systems and Signal Processing* **30**, 1829-1837 (2019).
14. T. Celik, "Spatial Entropy-Based Global and Local Image Contrast Enhancement," *IEEE Transactions on Image Processing* **23**, 5298 - 5308 (2014).
15. M. Zhou, K. Jin, S. Wang, J. Ye, and D. Qian, "Color Retinal Image Enhancement Based on Luminosity and Contrast Adjustment," *IEEE Transactions on Biomedical Engineering* (2018).
16. L. Cao, H. Li, and Y. Zhang, "Retinal image enhancement using low-pass filtering and α -rooting," *Signal Processing* **170**, 107445 (2020).
17. P. Dai, H. Sheng, J. Zhang, L. Li, J. Wu, and M. Fan, "Retinal Fundus Image Enhancement Using the Normalized Convolution and Noise Removing," *International Journal of Biomedical Imaging* **2016** 1-12 (2016).
18. N. D. Shemonski, FredrickA.South, Y.-Z. Liu, S. G. Adie, P. S. Carney, and S. A. Boppart, "Computational high-resolution optical imaging of the living human retina," *Nat. Photonics* **9**, 1-5 (2015).
19. J. Chung, G. W. Martinez, K. C. Lencioni, S. R. Sadda, and C. C. Yang, "Computational aberration compensation by codedaperture- based correction of aberration obtained from optical Fourier coding and blur estimation," *Optica* **6**, 647-661 (2019).

20. A. Arias and P. Artal, "Wavefront-shaping-based correction of optically simulated cataracts," *Optica* **7**, 22-27 (2020).
21. R. Dutta, S. Manzanera, A. Gambín-Regadera, E. Irlles, E. Tajahuerce, J. Lancis, and P. Artal, "Single-pixel imaging of the retina through scattering media," *Biomedical Optics Express* **10**, 4159-4167 (2019).
22. D. Singh and V. Kumar, "A Comprehensive Review of Computational Dehazing Techniques," *Archives of Computational Methods in Engineering* (2018).
23. K. P. Senthilkumar and P. Sivakumar, "A Review on Haze Removal Techniques," *Computer Aided Intervention and Diagnostics in Clinical and Medical Images*, 113-123 (2019).
24. S. Banerjee and S. S. Chaudhuri, "Nighttime Image-Dehazing: A Review and Quantitative Benchmarking," *Archives of Computational Methods in Engineering* (2020).
25. E. Peli and T. Peli, "Restoration of retinal images obtained through cataracts," *IEEE Transactions on Medical Imaging* **8**, 401-406 (1989).
26. B. Savelli, A. Bria, A. Galdran, C. Marrocco, M. Molinara, A. e. Campilho, and F. Tortorella, "Illumination correction by dehazing for retinal vessel segmentation," *IEEE 30th International Symposium on Computer-Based Medical Systems* (2017).
27. Y. Wang, S. Zhuo, D. Tao, J. Bu, and N. Li, "Automatic local exposure correction using bright channel prior for under-exposed images," *Signal Processing* **93**, 3227-3238 (2013).
28. A. Mitra, S. Roy, S. Roy, and S. K. Setua, "Enhancement and restoration of non-uniform illuminated Fundus Image of Retina obtained through thin layer of cataract," *Computer Methods and Programs in Biomedicine* **156**, 169-178 (2018).
29. A. Gaudio, A. Smailagic, and A. e. Campilho, "Enhancement of Retinal Fundus Images via Pixel Color Amplification," *International Conference on Image Analysis and Recognition* (2020).
30. H. Koschmieder, "Luftlicht und sichtweite," *Naturwissenschaften*, **26**, 521-528 (1938).
31. E. J. McCartney, "Optics of the Atmosphere: Scattering by molecules and particles," New York, 408 (1976).
32. P. Artal, I. Iglesias, and N. L'opez-Gil, "Double-pass measurements of the retinal-image quality with unequal entrance and exit pupil sizes and the reversibility of the eye's optical system," *J. Opt. Soc. Am. A* **12**, 2358-2366 (1995).
33. P. Artal, S. Marcos, R. Navarro, and D. R. Williams, "Odd aberrations and double-pass measurements of retinal image quality," *J. Opt. Soc. Am. A* **12**, 195-201 (1995).
34. D. Christaras, H. Ginis, A. Pennos, and P. Artal, "Intraocular scattering compensation in retinal imaging," *Biomed. Opt. Express* **7**, 3996-4006 (2016).
35. S. Wang and G. Luo, "Naturalness Preserved Image Enhancement Using a priori Multi-Layer Lightness Statistics," *IEEE transactions on image processing* **27**, 938-948 (2017).
36. D. J. Jobson, Z.-u. Rahman, and G. A. Woodell, "Properties and Performance of a Center/Surround Retinex," *IEEE Trans. Image Process.* **6**, 451-462 (1997).
37. A. Arias, H. Ginis, and P. Artal, "Light scattering in the human eye modelled as random phase perturbations," *Biomed. Opt. Express* **9**, 2664-2670 (2018).
38. T. T. J. M. Berendschot, J. v. d. Kraats, M. J. Kanis, and D. v. Norren, "Directional model analysis of the spectral reflection from the fovea and para-fovea," *Journal of Biomedical Optics* **15**, 065005 (2010).
39. X. Hadoux, F. Hui, J. K. H. Lim, C. L. Masters, A. Pebay, S. Chevalier, J. Ha, S. Loi, C. J. Fowler, C. Rowe, V. L. Villemagne, E. N. Taylor, C. Fluke, J. P. Soucy, F. Lesage, J. P. Sylvestre, P. Rosa-Neto, S. Mathotaarachchi, S. Gauthier, Z. S. Nasreddine, J. D. Arbour, M. A. Rheaume, S. Beaulieu, M. Dirani, C. T. O. Nguyen, B. V. Bui, R. Williamson, J. G. Crowston, and P. van Wijngaarden, "Non-invasive in vivo hyperspectral imaging of the retina for potential biomarker use in Alzheimer's disease," *Nat Commun* **10**, 4227 (2019).

40. H. S. Ginis, G. M. Perez, J. M. Bueno, A. Pennos, and P. Artal, "Wavelength dependence of the ocular straylight," *Invest Ophthalmol Vis Sci.* **54**, 3702-3708 (2013).
41. M. Ao, X. Li, W. Qiu, Z. Hou, J. Su, and W. Wang, "The impact of age-related cataracts on colour perception, postoperative recovery and related spectra derived from test of hue perception," *BMC Ophthalmology* **19**, 1-9 (2019).
42. U. Mehta, A. Diep, K. Nguyen, B. Le, C. Yuh, C. Frambach, J. Doan, A. Wei, A. M. Palma, M. Farid, S. Garg, S. Kedhar, M. Wade, K. A. Marshall, K. A. Jameson, M. C. Kenney, and A. W. Browne, "Quantifying Color Vision Changes Associated With Cataracts Using Cone Contrast Thresholds," *Translational Vision Science & Technology* **9**, 1-10 (2020).
43. "cataract dataset", retrieved <https://www.kaggle.com/jr2ngb/cataractdataset>.
44. K. He, J. Sun, and X. Tang, "Single Image Haze Removal Using Dark Channel Prior," (2009).
45. M. Ju, C. Ding, Y. J. Guo, and D. Zhang, "IDGCP: Image Dehazing Based on Gamma Correction Prior," *IEEE Trans. Image Process.* **29**, 3104 - 3118 (2021).
46. "DRIVE: Digital Retinal Images for Vessel Extraction", retrieved <https://drive.grand-challenge.org/>.
47. "STructured Analysis of the Retina", retrieved <https://cecas.clemson.edu/~ahoover/stare/>.
48. A. Hoover, V. Kouznetsova, and M. Goldbaum, "Locating Blood Vessels in Retinal Images by Piece-wise Threshold Probing of a Matched Filter Response," *IEEE Transactions on Medical Imaging* **19**, 203-210 (2000).
49. "DIARETDB1 - Standard Diabetic Retinopathy Database Calibration level 1", retrieved <http://www2.it.lut.fi/project/imageret/diaretdb1/>.
50. T. Kauppi, V. Kalesnykiene, J.-K. Kamarainen, L. Lensu, I. Sorri, A. Raninen, R. Voutilainen, H. Uusitalo, H. Kälviäinen, and J. Pietilä, "DIARETDB1 diabetic retinopathy database and evaluation protocol," in *Medical Image Understanding and Analysis*, (Citeseer, 2007), p. 61.
51. K. Panetta, C. Gao, and S. Agaian, "Human-Visual-System-Inspired Underwater Image Quality Measures," *IEEE Journal of Oceanic Engineering* **41**, 1-11 (2016).
52. J. Wang, K. Lu, J. Xue, N. He, and L. Shao, "Single Image Dehazing Based on the Physical Model and MSRCR Algorithm," *IEEE Transactions on Circuits and Systems for Video Technology* **29**, 2190-2199 (2018).
53. S. Hong, M. Kim, and M. Kang, "Single image dehazing via atmospheric scattering model-based image fusion," *Signal Processing* **178**, 107798 (2021).
54. A. Rizzi, T. Algeri, G. Medeghini, and D. Marini, "A proposal for Contrast Measure in Digital Images," in *Conference on colour in graphics, imaging, and vision*, (2004), pp. 187-192.
55. L. K. Choi, J. You, and A. C. Bovik, "Referenceless prediction of perceptual fog density and perceptual image defogging," *IEEE Trans. Image Process.* **24**, 3888-3901 (2015).
56. L. K. Choi, J. You, and A. C. Bovik, "LIVE Image Defogging Database" (2015), retrieved https://live.ece.utexas.edu/research/fog/fade_defade.html.

Supplementary Notes for Chapter 3

DPFR: A double-pass fundus reflection model for efficient single retinal image enhancement

Supplementary Note 3-1: Experimental results for visual and quantitative assessments

Tab. s3-1. Quantitative assessment of the enhanced images among Fig. s3-1(a1) to s3-2(d1) with six metrics.

Raw image	Matrixes	Raw	Methods			
			LCA	LPAR	PCA	DPFR
Group 1 (Fig. s3-1)	DE	0.7033	1.6829	11.6208	3.1715	10.5461
	UISM	3.2169	3.6772	6.8555	4.5054	6.5344
	UIConM	0.0663	0.1112	0.3673	0.2330	0.3794
	IE	3.7828	3.7605	5.2418	3.7321	5.0104
	CRAMM	1.7218	2.5363	7.9630	4.3047	6.9873
	FADE	1.1502	0.7293	0.2123	0.5127	0.1705
Group 2 (Fig. s3-2)	DE	0.7035	1.7578	10.2321	4.2726	10.6374
	UISM	2.9919	3.4943	7.5659	4.7161	6.6337
	UIConM	0.0520	0.0989	0.3536	0.2617	0.3723
	IE	4.2933	4.2000	5.1948	3.9056	4.8937
	CRAMM	1.5881	2.3981	7.7575	4.6301	6.8675
	FADE	0.8029	0.5088	0.2219	0.3503	0.1689
Group 3 (Fig. s3-3)	DE	1.7946	4.2453	13.3563	8.2246	16.3367
	UISM	5.2235	6.0137	9.0230	9.2007	9.6236
	UIConM	0.1147	0.2060	0.3683	0.3896	0.4052
	IE	4.0018	4.5398	5.1823	3.7721	5.1266
	CRAMM	2.5071	4.2325	8.5827	7.4953	8.5843
	FADE	0.4434	0.2724	0.1793	0.2786	0.1317
Group 4 (Fig. s3-4)	DE	0.9598	3.4012	7.9515	1.6917	14.0354
	UISM	3.6441	4.7619	6.2515	3.8831	6.8982
	UIConM	0.0365	0.1339	0.2973	0.2650	0.3358
	IE	3.8867	4.5507	5.2396	4.5239	5.2381
	CRAMM	0.9922	2.7305	5.7523	4.6707	7.2670
	FADE	0.5351	0.2097	0.1875	0.3604	0.1609
Average	DE	1.0403	2.7718	10.7902	4.3401	12.8889
	UISM	3.7691	4.4868	7.4240	5.5763	7.4223
	UIConM	0.0674	0.1375	0.3466	0.2873	0.3732
	IE	3.9912	4.2626	5.2146	3.9834	5.0672
	CRAMM	1.7023	2.9744	7.5139	5.2752	7.4265
	FADE	0.7329	0.4301	0.2003	0.3755	0.158

In this section of the Supplementary Materials, additional four groups of experimental results are shown for visual and objective assessment. Tab. s3-1 lists the quantitative assessment of the four groups.

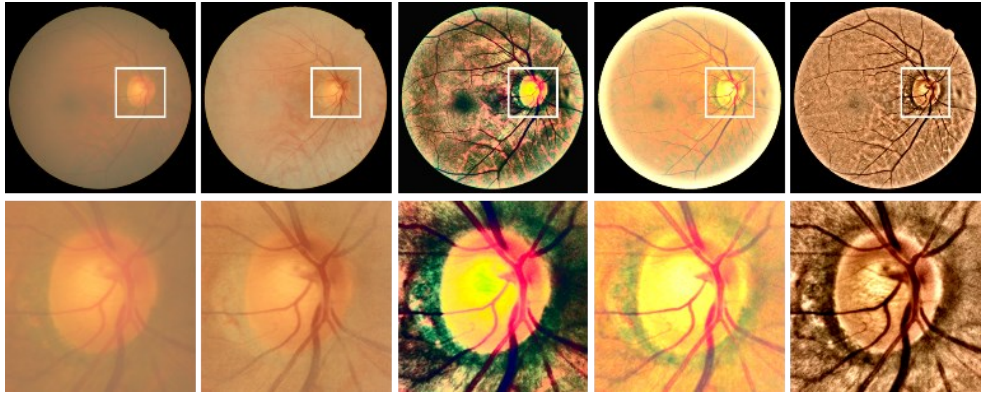


Fig. s3-1. Group 1 for visual assessment. Left to right are raw image, results of LCA, LPAR, PCA, and DPFR methods.

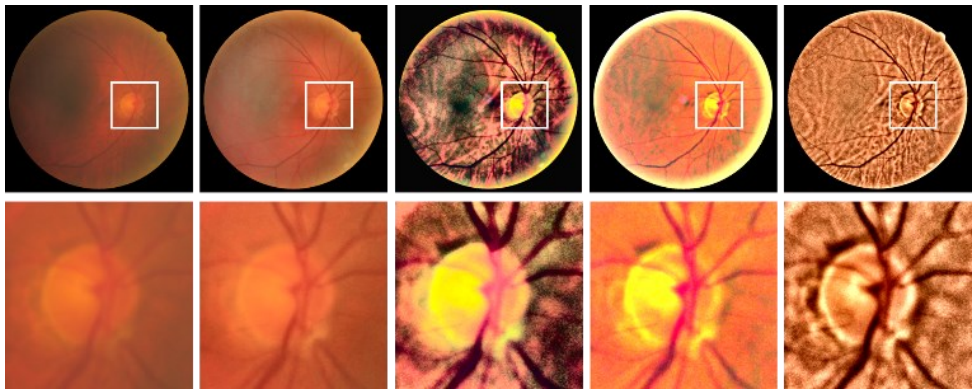


Fig. s3-2. Group 2 for visual assessment. Left to right are raw image, results of LCA, LPAR, PCA, and DPFR methods.

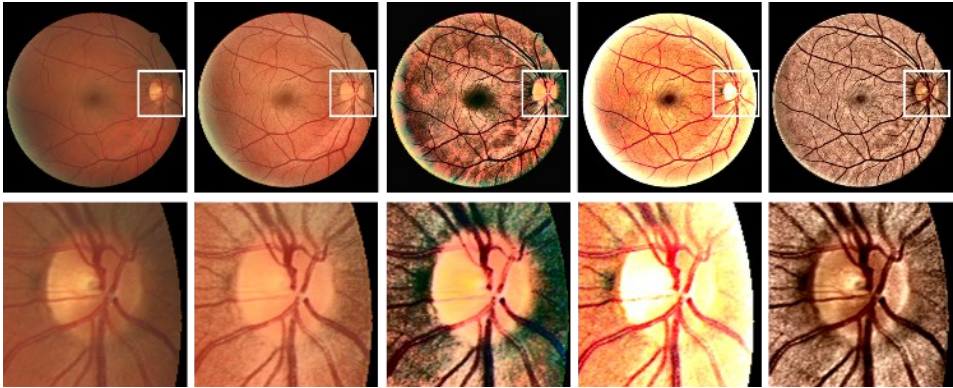


Fig. s3-3. Group 3 for visual assessment. Left to right are raw image, results of LCA, LPAR, PCA, and DPFR methods.

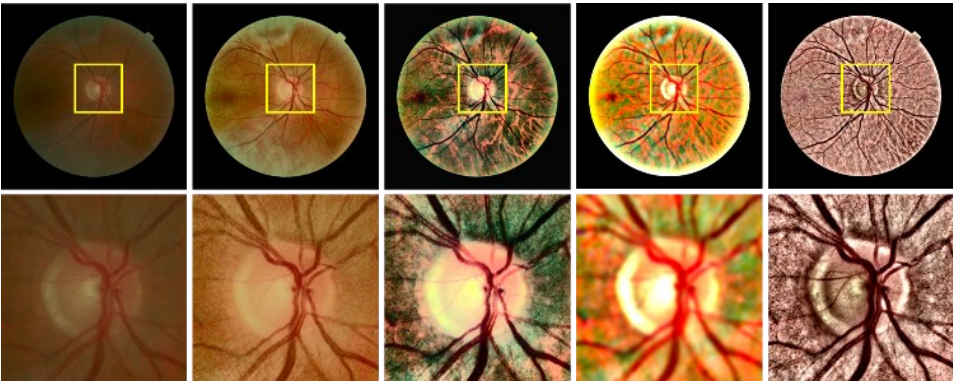


Fig. s3-4. Group 4 for visual assessment. Left to right are raw image, results of LCA, LPAR, PCA, and DPFR methods.

Supplementary Note 3-2: Experimental results in different color space

In this section, we demonstrate the enhancement by converting the coarse illumination corrected image into different color space. Results are shown in Fig. s3-5. The CIE-Lab results have similar visual effect as that of in YCbCr space. While the intensity of the entire image enhanced in CIE-Lab color space is slightly dimmer than that in YCbCr space. Enhanced results in HSV color space shows loose of texture information (near the optical disk) as the V component is given by the maximum value in RGB channels, and didn't meet our first requirement where the intensity should make full use of valid information of raw images.

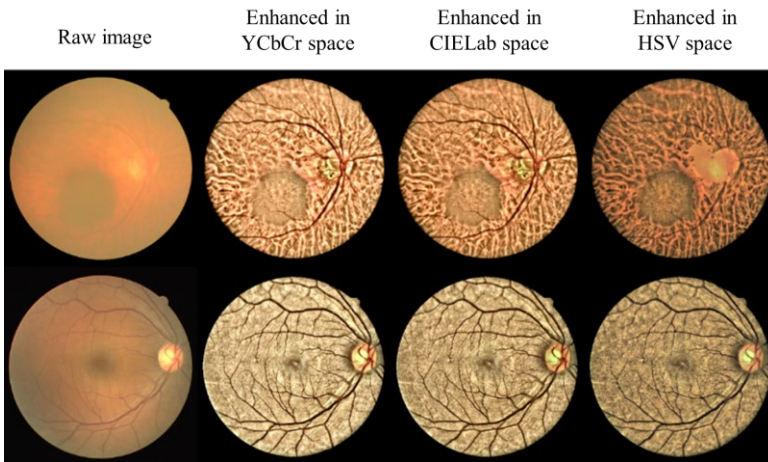


Fig. s3-5, Enhancement results in different color space. Columns from left to right are raw images, enhanced in the YCbCr color space, enhanced in the CIE Lab color space, and in the HSV color space, respectively.

Chapter 4

MUTE:

A multilevel-stimulated denoising strategy for single cataractous retinal image dehazing

Shuhe Zhang, Ashwin Mohan, Carroll A. B. Webers, and
Tos T. J. M. Berendschot

Medical Image Analysis, 2023 (88), pp. 102848.
DOI: 10.1016/j.media.2023.102848.

Abstract

In this research, we studied the duality between cataractous retinal image dehazing and image denoising and proposed that the dehazing task for cataractous retinal images can be achieved with the combination of image denoising and sigmoid function. To do so, we introduce the double-pass fundus reflection model in the YPbPr color space and developed a multilevel stimulated denoising strategy termed MUTE. The transmission matrix of the cataract layer is expressed as the superposition of denoised raw images of different levels weighted by pixel-wise sigmoid functions. We further designed an intensity-based cost function that can guide the updating of the model parameters. They are updated by gradient descent with adaptive momentum estimation, which gives us the final refined transmission matrix of the cataract layer. We tested our methods on cataract retinal images from both public and proprietary databases, and we compared the performance of our method with other state-of-the-art enhancement methods. Both visual assessments and objective assessments show the superiority of the proposed method. We further demonstrated three potential applications including blood vessels segmentation, retinal image registrations, and diagnosing with enhanced images that may largely benefit from our proposed methods.

Keywords

Ophthalmology; Retinal image; Cataract; Dehazing

1. Introduction

Retinal images provide noninvasive measurement of fundus morphology and are widely used by ophthalmologists for early detection, diagnosis, and monitoring of ocular diseases and their progression [1-7]. However, imaging of cataract patients' retina using fundus cameras is challenging since the image quality is severely degraded by light scattering of the turbid cataract layers. Dependent on the severity, many of the subtle retinal structures are hidden by the haze effect of cataract which largely limits the diagnosing accuracy and reliability for ophthalmologists, as a result, diagnosing according to cataractous retinal images is error-prone.

Improving the cataractous retinal image quality by modifying the optical system of a fundus camera, for example, using laser source or confocal imaging technique is expensive and also inconvenient for quick and universal retinal image checking. Therefore, image processing methods draw more attention to cataractous retinal image enhancement as they are economical and practical [8-12]. Various algorithms were investigated for cataractous retinal image enhancement, within which the image formation model-based methods show their priority in experimental results.

Given the similarity between cataractous retinal image and hazy image for natural scenes, many studies have adopted natural scene dehazing methods to achieve cataractous retinal image dehazing [8, 9, 13, 14], however, their results are limited. The main reason is that the spectral properties of retinal images are different from nature scenes due to the existence of intraocular lens, and absorption by ocular tissue [15]. Therefore, the natural scene dehazing methods suffer from the problem where the transmission matrix of cataract layers is underestimated (**Supplementary Note 4-1**).

In our previous study [16] we reviewed the historical image formation model of retinal imaging and developed the double-pass fundus reflection model (DPFR) dedicated to retinal image enhancement. It incorporates a specific double-pass fundus reflection feature [15, 17, 18] that had been neglected hitherto in modeling the light propagation of fundus imaging in all reports on retinal image enhancement. In this

study, we continued this research and show that the dehazing of retinal images can be achieved in a framework of image denoising or smoothing. We developed a multilevel-stimulated denoising strategy, called MUTE. In MUTE, the transmission matrix of the cataract layer is expressed as the superposition of denoised raw retinal images of different levels weighted by pixel-wise sigmoid functions. We further designed an intensity-based cost function that guides and stimulates the updating of the transmission matrix. A demonstration of experimental results of MUTE on cataractous retinal image dehazing is shown in Fig. 4-1.

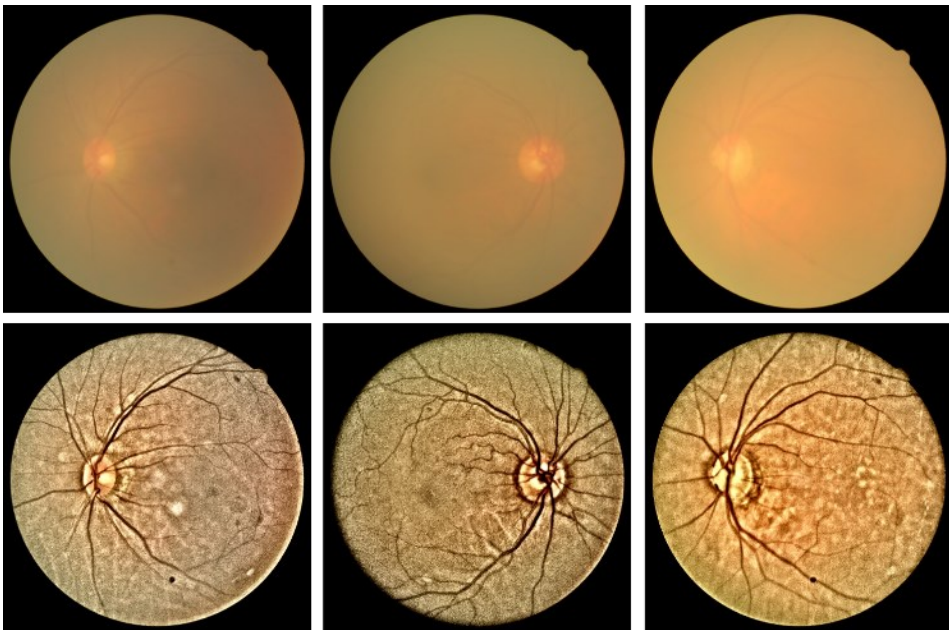


Fig. 4-1. Demonstrations of MUTE on cataractous retinal image dehazing. Images in the first row are cataractous retinal images. Images in the second row are outputs of MUTE.

The subsequent manuscript is organized as follows. Section 2 briefly reviews related work. Section 3 introduces the basic formulas for our retinal image formation model and the preprocessing procedures that are applied to the retinal image before retinal image dehazing. Section 4 presents and analyzes the proposed multilevel-stimulated

denoising (MUTE) strategy in detail. Section 5 shows the experimental results for both visual and objective assessment in comparison to state-of-the-art (SOTA) algorithms. Demonstrations of some potential clinical applications of MUTE are also included in Section 5. Section 6 contains detailed ablation study of each module of MUTE. Section 7 provides a Hessian denoising strategy to suppress the noise signals during the MUTE dehazing, and section 8 is discussions and the concluding remarks.

2. Related works

2.1 Retinal imaging formation/structuring model

The early-stage of retinal image formation model was directly adopted from Koschmieder & McCartney's model [19, 20] of hazy nature scenes, which is given by

$$\mathbf{S}_c(\mathbf{r}) = \mathbf{T}(\mathbf{r}) \circ \mathbf{O}_c(\mathbf{r}) + A_c [1 - \mathbf{T}(\mathbf{r})], \quad (4-1)$$

where \mathbf{O} is the haze-free image, \mathbf{T} is the transmission matrix of the haze medium describing the portion of the light that is not scattered and reaches the camera. A is the global atmospheric light, \mathbf{S} is the observed image. $\mathbf{r} = (x, y)$ is the vector of spatial coordinates and $c \in \{R, G, B\}$ is the index of the color channels. \circ is the pixel-wise multiplication.

Based on Eq. (4-1), many studies have been proposed to enhance the visual quality of hazy and underwater images, achieving promising results [21-24]. Some studies also introduced Eq. (4-1) for retinal image enhancement. In these applications, \mathbf{T} was regarded as the transmission matrix of the cataract layer, as cataract also scatters light and imposes haze-effect on captured retinal images. Xiong *et al.* proposed to use intensity correction and histogram adjustment to preprocess the image, generating a transmission map according to the intensity of the preprocessed image in each color channel [9]. Haze effects were then suppressed through dehazing. Gaudio *et al.* demonstrated a pixel color amplification method for retinal imaging enhancement that showed good performance in enhancing the detailed structure of retinal images [14]. Cao *et al.* proposed the detail-richest-channel to estimate the transmission matrix \mathbf{T}

[13]. However, Eq. (4-1) was developed for natural scenes and may not be the optimal choice for fundus imaging, ignoring its double-pass property.

Peli *et al.* [8] developed an optical model for imaging the retina through cataracts which is

$$\mathbf{S}_c(\mathbf{r}) = \alpha L \circ \mathbf{T}(\mathbf{r}) \circ \mathbf{O}_c(\mathbf{r}) + L \circ [1 - \mathbf{T}(\mathbf{r})], \quad (4-2)$$

L is the flash illumination of the fundus camera and α is the attenuation of retinal illumination due to the cataract. Both L and α are considered as constant. Following Eq. (4-2), many methods have been proposed to enhance the cataractous retinal image. Mitra *et al.* proposed to use the Retinex theory to remove the haze effect [10], which is analogous to Peli's work. Luo *et al.* developed an unpaired generative adversarial network to achieve haze removal [12].

Different from Eq. (4-1), Eq. (4-2) reveals that the illumination pattern also impacts the quality of retinal imaging. However, as L is constant, Eq. (4-2) loses the ability to correct the uneven (spatially varying) illumination of retinal imaging. In addition, the existing parameter α shows the basic idea of the double-pass property where the illumination light interacts twice with cataract layer. (When the light goes inside the eye, and when it is reflected out from the fundus).

In our previous study [16], we proposed the double-pass fundus reflection (DPFR) that deals with image formation in retinal imaging. This DPFR model is given by

$$\mathbf{S}_c(\mathbf{r}) = \mathbf{L}(\mathbf{r}) \circ \mathbf{T}_{lens}^2(\mathbf{r}) \circ [\mathbf{T}_{sc}^2(\mathbf{r}) \circ \mathbf{O}_c(\mathbf{r}) + 1 - \mathbf{T}_{sc}(\mathbf{r})], \quad (4-3)$$

where $\mathbf{L}(\mathbf{r})$ is the illumination from the outside of the eye and is delivered by the illumination system of the fundus camera. We consider that the retina is illuminated by white light (identical value in R, G and B channels) which may have an uneven and insufficient illumination pattern. $\mathbf{T}_{lens}(\mathbf{r})$ is the transmission matrix of the combination of cornea and lens. $\mathbf{T}_{sc}(\mathbf{r})$ is the transmission matrix of intraocular scatter including cataractous. Eq. (4-3) reveals that the degeneration of the retinal image is mainly due to three parts: (1) an uneven illumination condition, (2) filtering by the human lens, and

(3) intraocular scattering. Based on Eq. (4-3), we now can correct the illumination problem and suppress the haze effect.

Besides the image formation model, there are also image structure models used for retinal image enhancement [11, 25-27], and the image structure models can be summarized as

$$\mathbf{S}_c(\mathbf{r}) = \mathbf{S}_{c,background}(\mathbf{r}) + \mathbf{S}_{c,details}(\mathbf{r}), \quad (4-4)$$

where $\mathbf{S}_{background}$ is the background information of the observed image which corresponds to the low-frequency components, while $\mathbf{S}_{details}$ denotes the detailed information implying the detailed structures and textures of the image. By giving a large weight to $\mathbf{S}_{details}$ and suppressing the $\mathbf{S}_{background}$ one can obtain a contrast-enhanced image. $\mathbf{S}_{background}$ can be obtained by low pass filtering of \mathbf{S} [11, 26] and total variation regularization [27]. While $\mathbf{S}_{details}$ can be obtained by high pass filtering of \mathbf{S} or subtracting $\mathbf{S}_{background}$ from \mathbf{S} . Note that Eq. (4-4) is not based on the optical process of how the image is formed and the physical insight is different from Eqs. (4-1) to (4-3).

2.2 Retinal image illumination correction

Since retinal images may suffer from uneven and insufficient illumination conditions, correction of the uneven illumination is necessary to enhance its visual quality. Methods for illumination correction are categorized into model-based and non-model-based.

For non-model-based methods, Zhou *et al.* proposed to use gamma correction to the V channel of retinal images in HSV color space [28]. This process provides a luminance gain matrix, which can later be applied to RGB channels to correct the uneven illumination. Gupta *et al.* modified Zhou's method by combining the adaptive gamma correction to refine the luminance gain matrix [29].

For model-based methods, the Retinex theory and its variations are used in many reports about correcting the uneven illumination of retinal images. Based on the assumption where the illumination pattern of an image is the low-frequency

component, an estimation of the illumination pattern can be obtained by low-pass filtering the given retinal images and removed in the logarithm domain [11, 26]. Besides, based on the observation that the color reversed image which has insufficient/uneven illumination conditions is similar to the image with hazed effects, the bright channel prior (color inverted version of dark-channel prior) also provides an illumination correction method [13, 30], and they can be combined with the Retinex theory [16].

2.3 Cataractous retinal image dehazing

Dehazing of cataractous retinal image seems to be similar to a natural scene. Dehazing of a single natural scene image is ill-posed, therefore, many prior knowledges were proposed to obtain the transmission matrix T_{sc} from known only of raw haze image S . For example, in the framework of the dark channel prior [31], T_{sc} can be obtained by local minimum filtering of S_Y^I among RGB color channel. In the color attenuation prior [32], T_{sc} can be estimated from the difference between image brightness and saturation. In the gamma correction prior [33], T_{sc} is given by the intensity difference before and after the gamma correction of the raw image.

However, as we show in **Supplementary Note 4-1**, prior knowledges for natural hazed image fail to obtain the appropriate T_{sc} in the case of retinal image dehazing. For example, the dark channel prior fails because there are insufficient shadow areas in retinal image S . The color attenuation prior fails because the spectral statistical property of retinal images is different from nature images due to the existence of the ocular lens. While the failure of gamma correction prior fails is due to that Eq. (4-3) cannot be simplified to meeting the derivation of gamma correction prior due to the squared term of T_{sc} . In general, new dehazing algorithms that are dedicated to retinal image dehazing tasks need to be designed.

3. The DPFR model and image preprocessing

3.1 The DPFR in the YPbPr color space

As shown in Fig. 4-2 (a), the DFPR model for retinal image formation in RGB color space is described by Eq. (4-3). Let's consider the linear transform matrix Ψ that converts RGB color space to YPbPr color space, Ψ is given by

$$\Psi = \begin{pmatrix} K_R & K_G & 1 - K_R - K_G \\ \frac{K_R}{2(K_R + K_G)} & \frac{K_G}{2(K_R + K_G)} & \frac{1}{2} \\ \frac{1}{2} & -\frac{K_G}{2(1 - K_R)} & -\frac{1 - K_R - K_G}{2(1 - K_R)} \end{pmatrix}, \quad (4-5)$$

and $\Psi(1,1,1)^T = (1,0,0)$. Applying Eq. (4-5) to both sides of Eq. (4-3), we obtained

$$\begin{pmatrix} \mathbf{S}_Y \\ \mathbf{S}_{P_B} \\ \mathbf{S}_{P_R} \end{pmatrix} = \mathbf{L} \cdot \mathbf{T}_{lens}^2 \cdot \left[\mathbf{T}_{sc}^2 \cdot \begin{pmatrix} \mathbf{O}_Y \\ \mathbf{O}_{P_B} \\ \mathbf{O}_{P_R} \end{pmatrix} + \begin{pmatrix} \mathbf{I} - \mathbf{T}_{sc} \\ 0 \\ 0 \end{pmatrix} \right]. \quad (4-6)$$

Equation (4-6) reveals the mathematical insight that the hazed effect only happens to the Y channel of the retinal image as the Y channel has the haze term $\mathbf{I}_{ill} \cdot \mathbf{T}_{lens}^2 \cdot (1 - \mathbf{T}_{sc})$ which denotes the backscattering of light from intraocular scatters as shown in Fig. 4-2 (b). For P_B and P_R channels, the existence of \mathbf{T}_{sc} works similar to \mathbf{I}_{ill} which only decreases the intensity of corresponding color channels of the haze-free image. According to ITU-R BT.601 standard, $K_R = 0.299$ and $K_G = 0.587$.

According to Eq. (4-6), our proposed algorithm converts the color retinal image in $Y P_B P_R$ color space, and three channels are treated differently. Both intensity correction and dehazing are applied to the Y channel, while only intensity correction is applied to the P_B and P_R channels. After then, the retinal image is converted back to RGB color

space. In our proposed method, the grayscale is normalized between $[0, 1]$ by dividing the loaded images (uint8) by 255.

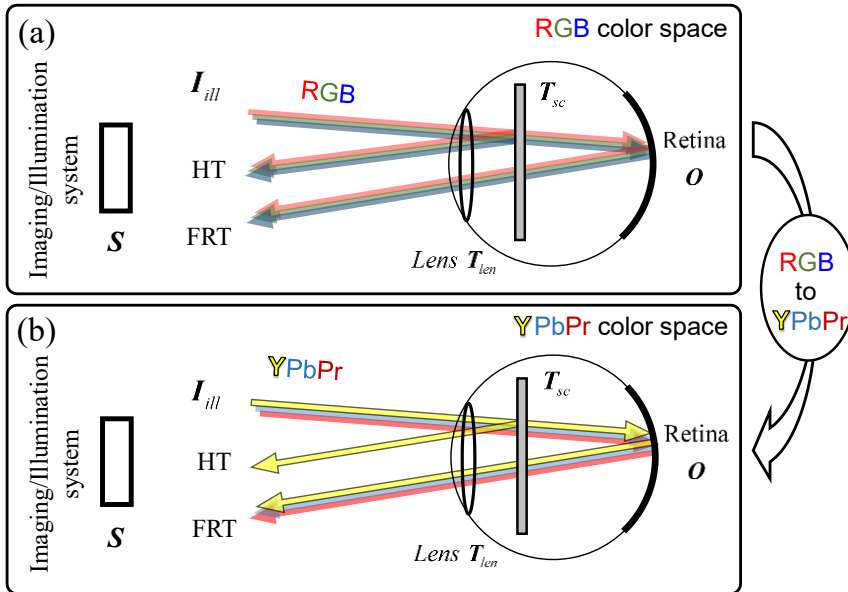


Fig. 4-2. The sketch of the double-pass fundus reflection model. S is the final output image of the imaging system. (a) in RGB color space. (b) in YPbPr color space. HT: haze term; FRT: fundus reflection term.

3.2 Image preprocessing

Before entering the dehazing process, preprocessing including background padding and illumination correction is applied to the raw images. A retinal image is composed of two major regions - a circular retinal region of interest (ROI) and a black background. Let T be all pixels in the circular ROI. Since the image will be filtered by Gaussian kernels during the following process, unexpected enhancement results often occur when the convolution kernel slides through the retinal image due to the sudden change of pixel values. Padding the black areas before enhancement is an effective way to reduce the boundary effect during the convolution. Here we use mirror reflection padding to fill the black background.

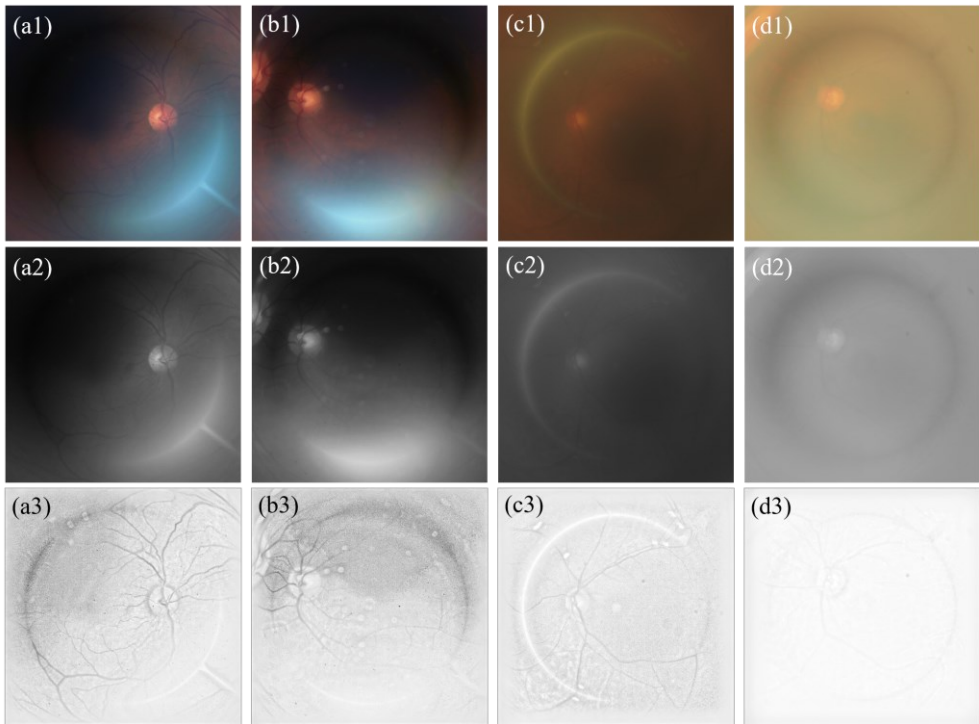


Fig. 4-3. Image preprocessing. (a1) to (d1) are background padded retinal image. (a2) to (d2) are their Y channels, respectively. (a3) to (d3) show the illumination correction results.

Hereafter, we apply our previously reported adaptive illumination correction method to the Y channel of the retinal image to correct its uneven and insufficient illumination pattern. The illumination correction can unify the intensity distribution of input images, and benefit following dehazing process. Since this research focuses on the dehazing process, we refer to Ref [16] or **Supplementary Note 4-2** for detailed information on illumination correction. Moreover, the ablation study for illumination correction will be discussed in Section 6.2 and 6.3. Fig. 4-3 shows the raw cataractous retinal images and corresponding Y channels before and after illumination correction treatment.

4. Dehazing using pixel-wise stimulated multilevel denoising

After retinal image illumination correction, the Y channel becomes

$$\mathbf{S}_Y^{IC} = \mathbf{T}_{sc}^2 \circ \mathbf{O}_Y + \mathbf{1} - \mathbf{T}_{sc}, \quad (4-7)$$

\mathbf{S}_Y^{IC} denotes the Y channel of the source image after illumination correction. The first term $\mathbf{T}_{sc}^2 \circ \mathbf{O}_Y$ of Eq. (4-7) is the scene term denoting the light reflected from the fundus and attenuated by the cataract layer. The second term $\mathbf{1} - \mathbf{T}_{sc}$ is the haze term denoting the backscattering light from the cataract layer. Once \mathbf{T}_{sc} is known, \mathbf{O}_Y can be retrieved by

$$\mathbf{O}_Y = \frac{\mathbf{S}_Y^{IC} - \mathbf{1}}{\mathbf{T}_{sc}^2} + \min\left(\frac{\mathbf{1}}{\mathbf{T}_{sc}}, \mathbf{1}\right). \quad (4-8)$$

Considering that (I) the retinal image \mathbf{O}_Y contains blood vessels, and many speckle-like textures due to the scattering of biological tissues, and (II) the haze term is spatially slow-varying since the cataract lays on the pupil plane which is severely out-of-focus from the retinal plane [8, 34]. (III) Eq. (4-8) is analog to Eq. (4-4) as $\mathbf{T}_{sc}^2 \circ \mathbf{O}_Y$ can be regarded as the structure term and $\mathbf{1} - \mathbf{T}_{sc}$ can be regarded as the background term. We, therefore, are able to regard $\mathbf{T}_{sc}^2 \circ \mathbf{O}_Y$ as the “noise” imposed on $\mathbf{1} - \mathbf{T}_{sc}$ and can be removed by image denoising, yielding

$$\text{Denoising}(\mathbf{S}_Y^{IC}) \approx \mathbf{1} - \mathbf{T}_{sc} \Rightarrow \mathbf{T}_{sc} \approx \mathbf{1} - \text{Denoising}(\mathbf{S}_Y^{IC}), \quad (4-9)$$

Eq. (4-9) implies that the transmission matrix \mathbf{T}_{sc} can be obtained by denoising (or smoothing) the input hazy image \mathbf{S}_Y^{IC} .

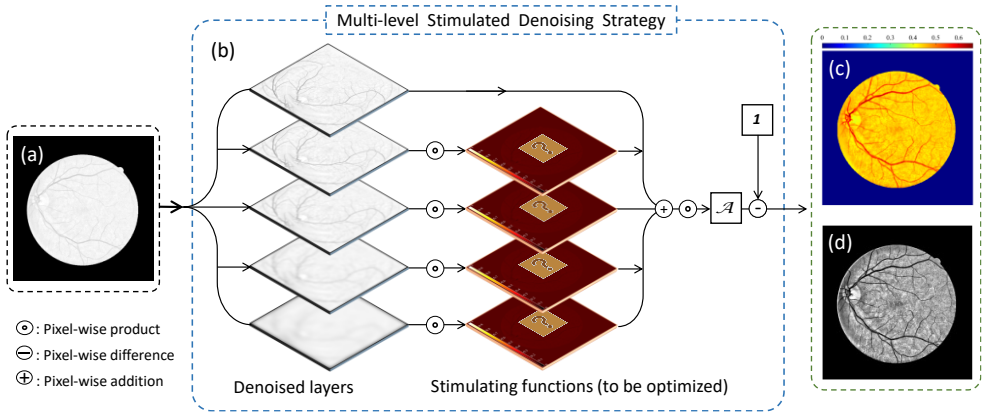


Fig. 4-4. Multilevel-stimulated denoising strategy for retinal image dehazing. (a) input Y channel. (b) Multilevel denoised image controlled by stimulating functions. (c) the final output of estimation of T_{sc} . (d) final output Y channel of the restored image.

Since off-the-shelf denoising algorithms are not designed for cataractous retinal image dehazing tasks, the denoised image needs to be further refined to form a good estimation of T_{sc} . Inspired by multi-scale treatment, we consider the T_{sc} as a linear combination of multiple denoised images of different denoising strengths. Images with strong denoising can be regarded as coarse estimation of T_{sc} since they contain large-scale features of T_{sc} . While images with weak denoising can be regarded as fine estimation of T_{sc} but corrupted by the reflected components $T_{sc}^2 \circ \mathbf{O}_Y$. We are further inspired by stimulating functions, and we assign stimulated functions to different levels of denoised images, so that we can design a cost function that guides the model on which pixel should contain more details and which pixel shouldn't. (For the detailed relationship between image dehazing and image denoise, please refer to **Supplementary Note 4-3**.)

Based on this idea, here we proposed the adaptively multilevel stimulated denoising strategy (MUTE) to obtain a delicate estimation of T_{sc} . The structure of MUTE is shown in Fig. 4-4. Eq (4-9) is rewritten as

$$\mathbf{T}_{sc} = 1 - \mathcal{A} \cdot \left\{ \mathcal{S}_0[\mathbf{S}_Y^{IC}(\mathbf{r})] + \sum_{l=1}^{\mathcal{L}} \theta_l(\mathbf{r}) \circ \mathcal{S}_l[\mathbf{S}_Y^{IC}(\mathbf{r})] \right\}, \quad (4-10)$$

where \mathcal{A} is a global amplitude term that controls the magnitude of all denoising terms. \mathcal{S}_l , ($l=0, 1, 2 \dots \mathcal{L}-1$) is the denoising of \mathbf{S}_Y^{IC} at l -th level of denoising strength (a total of \mathcal{L} levels, $\mathcal{L} \geq 2$). $\theta_l(\mathbf{r})$ is the pixel-wise stimulated (sigmoid) function which is given by

$$\theta_l(\mathbf{r}) = \frac{1}{\mathcal{L}-1} \cdot \left\{ \frac{1}{1 + \exp[-\varphi_l(\mathbf{r})]} \right\} \quad (4-11)$$

that assigns different weights at point \mathbf{r} for l -th denoising level. $\varphi_l(\mathbf{r})$ is the parameter that controls the stimulated weight for $\theta_l(\mathbf{r})$. We choose the sigmoid function due to its simple expression of its first-order derivative and the sigmoid function gets saturated even if the value of $\varphi_l(\mathbf{r})$ gets very large, which makes the gradient descent won't go too far from the optimal points in each iteration. The remaining tasks are to design a cost function that guides the automatic adjustment of $\varphi_l(\mathbf{r})$ and the amplitude \mathcal{A} .

Recall that an underestimation of \mathbf{T}_{sc} darkens the intensity of \mathbf{O}_Y but enriches its texture information [35], while the overestimation of \mathbf{T}_{sc} brightens the intensity of \mathbf{O}_Y but decreases the texture information due to the haze effect. To optimize the parameters $\varphi_l(\mathbf{r})$ and \mathcal{A} , we design an intensity-based cost function

$$E = \sum_{\mathbf{r}} (\mathbf{O}_Y - \bar{I})^2 - \lambda \sum_{\mathbf{r}} |\nabla \mathbf{O}_Y|^2, \quad (4-12)$$

where $\bar{I} = I_0 \times \text{mean}_{\mathbf{r} \in I}(\mathbf{S}_Y^{IC})$, and I_0 is the intensity ratio that controls the dehazing strength. A small value of I_0 leads to a strong dehazed effect while also darkening the image. The first term ensures that the output \mathbf{O}_Y has a proper intensity distribution,

and the second term ensures the image has rich texture information. λ ($\lambda < 1$, or $\lambda = 0$) is a small value to ensure the first term contributes most to the gradient.

The MUTE can be regarded as a one-layer network, with only sigmoid functions. The derivative of Eq (4-12) with respect to parameters $\varphi_l(\mathbf{r})$ and \mathcal{A} can be either calculated by automatic differentiation, or calculated in closed-form as written in **Supplementary Note 4-4**.

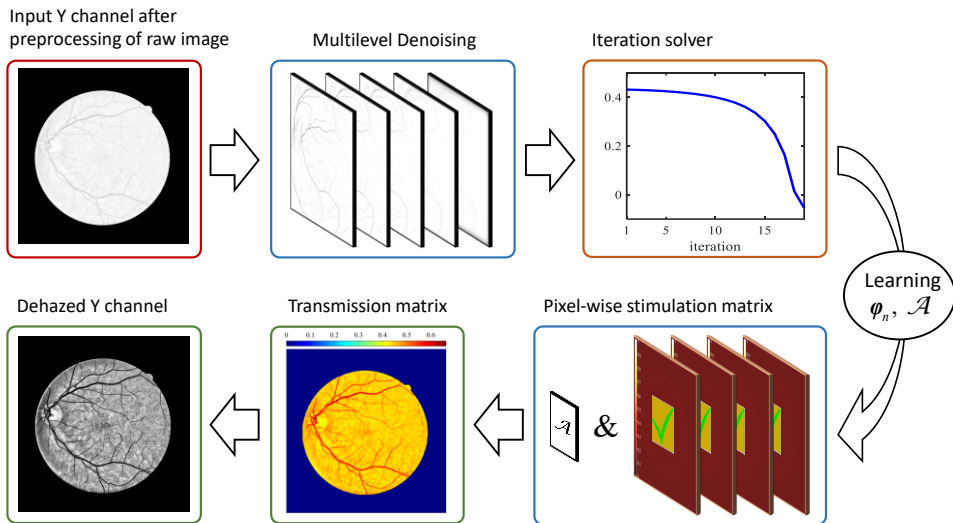


Fig. 4-5. Flow chart of proposed Y channel dehazing method.

In MUTE, we use the Adam optimizer to update model parameters [36], $\eta = 0.01$ (step size), $\alpha = 0.9$ (first-order decay rate) and $\beta = 0.99$ (second-order decay rate) in all following experiments. Moreover, Let $R_n = \text{mean}_{r \in I} (\mathbf{O}_Y - \bar{I})^2$ be part of the residual of cost function in the n -th iteration and we stop the iteration process when $R_n < 0.005$. Fig. 4-5 shows the flow chart of the proposed dehazing algorithm.

Many off-the-shelf image denoising algorithms can be used for computing \mathcal{S}_l in Eq. (4-10), here we choose Gaussian filter for simplicity which are given by

$$\mathcal{S}_l(\mathbf{r}) = \mathcal{S}_Y^{IC}(\mathbf{r}) \otimes \left[\frac{1}{2\pi w_l^2} \exp\left(-\frac{\mathbf{r}^2}{2w_l^2}\right) \right], \quad (4-13)$$

and the size of the gaussian kernel at l -th level is

$$w_l = \left(\frac{\sqrt{M \cdot N}}{20} \right) \times \left(\frac{1}{2} \right)^{\mathcal{L}-l}. \quad (4-14)$$

Figure 6 shows \mathbf{O}_Y and corresponding \mathbf{T}_{sc} when $\lambda = 1$, $\mathcal{L} = 4$ and $I_0 = 0.50$ for images in Figs. 3(a3) to 3(d3), respectively. As shown in Fig. 6 (a1) to (d1), the patterns of \mathbf{T}_{sc} have low transmission maps due to the presence of cataracts. Since the intensity of the pattern is low for haze retinal images, in order to have a clear observation of the transmission pattern, we normalized the intensity in area I as shown in the second row of Fig. 4-6, where the areas near the blood vessels are regarded as having higher transmission than other places.

Comparing images in Fig. 4-3 (a3-d3) to Fig. 4-6 (a3-d3), the haze effect is significantly suppressed as the retinal structures including optical disks and blood vessels can be clearly observed. Since the iteration could be fast and converged in only a few steps, we insert additional frames in the videos for better demonstration.

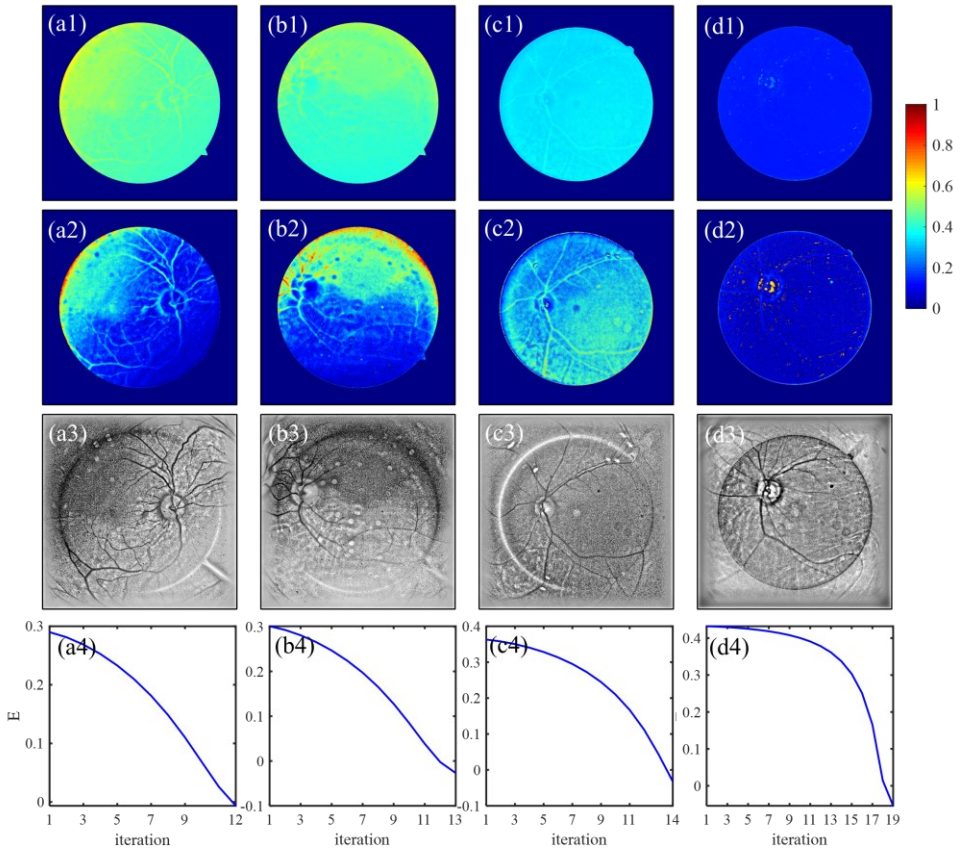


Fig. 4-6. Results for Y channel dehazing. (a1) to (d1) are obtained T_{sc} for images in Figs. 4-3(a3) to 3(d3). The patterns are truncated by the mask of the region of interest. (a2) to (d2) are intensity normalized pattern of (a1) to (d1). (a3) to (d3) are dehazed Y channel. (a4) to (d4) are changes of the cost function with respect to iterations.

In general, a flow chart for the proposed enhancement method is summarized in Fig. 4-7. The Hessian denoising is an optional process that will be discussed in Section 7 in detail.

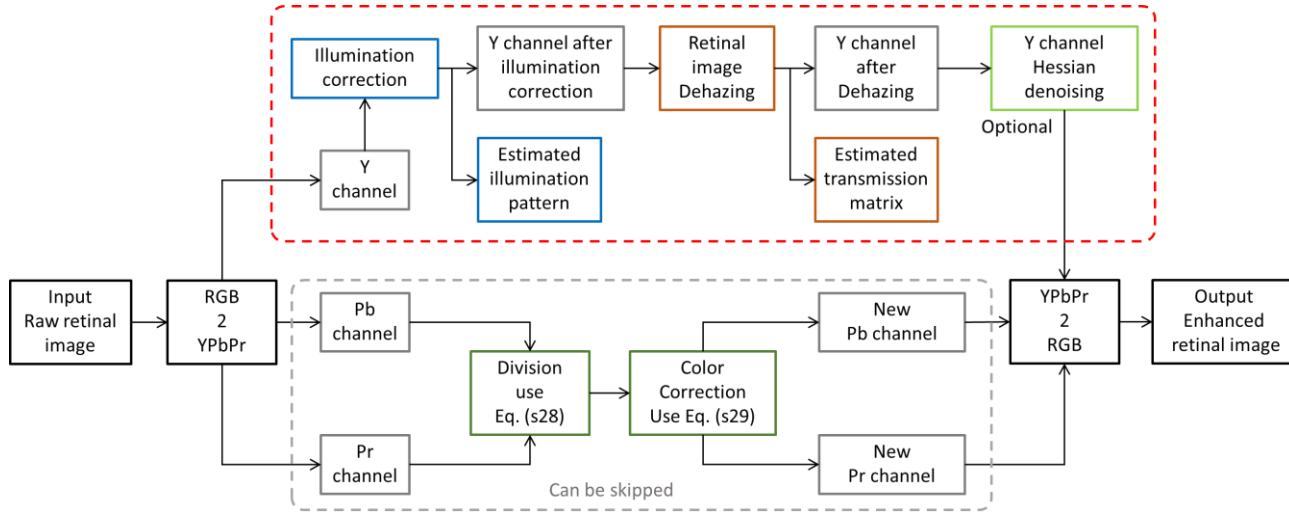


Fig. 4-7. A flow chart for MUTE method. Eq. (s21) and Eq. (s22) are in the **Supplementary Note 4-5**.

5. Experimental results

To show its performance, this section compares, both qualitatively and quantitatively, MUTE with that of three recently published state-of-art methods, namely the decomposition and visual adaptation (DVA) [27], the detail-richest-channel method (DRC) [13], and the low-pass filtering and α -rooting (LPA) [26]. Note that in [13] and [26] it was proven that the enhancement performances of DRC and LPA are better than methods presented in [9] and [10]. Image black background padding was applied to all raw retinal images before retinal image enhancement. The treatment in Pb/Pr channel is not applied.

In actual implementation, the kernel parameters for DRC and LPA are determined adaptively according to the size of input images. Other parameters for DRC and LPA models such as intensity correction strength can be automatically determined by the models. While for DVA method, it has two parameters, λ to determine the frequency band, and α to control the local contrast for the enhancement. In our implementation, $\lambda = 0.3$, $\alpha = 600$ based on [27]. For MUTE, we choose $I_0 = 0.52$ in following experiments. Explanation of this selection will be discussed in Section 6.2.

5.1 Visual assessments

First of all, we compared the results to assess the performance of the proposed method in terms of removing the haze effect and enhancing retinal structure visibility with images in a public, and our proprietary cataractous retinal image databases. Four groups of raw images and corresponding enhanced images are shown in Fig. 4-8. The first row shows raw cataract retinal images, where each image has a severe haze effect as most of the structures including blood vessels and optical disks are hidden behind the cataract layer.

From visual assessment, DVA corrects the haze effect for some images as shown in Fig. 4-8 (b2) and Fig. 4-8 (c2), as the blood vessels near the optical disks are enhanced. However, the performance of DVA is not stable as it failed to correct the haze effect for

images in Fig. 4-8 (a2) and Fig. 4-8 (d2), especially for the region that is far from the optical disks.

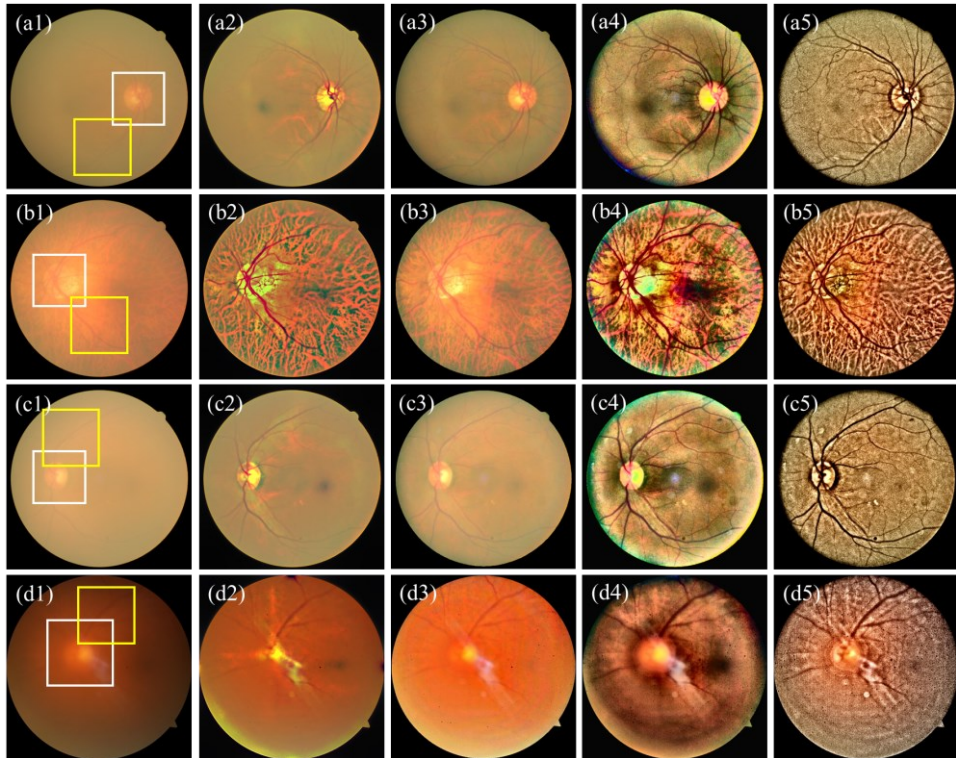


Fig. 4-8. Cataractous retinal image enhancement results. Columns from left to right are raw images, enhanced results of DVA, DRC, LPAR, and MUTE methods.

The results of DRC are shown in Fig. 4-8 (a3) to Fig. 4-8 (d3), where the haze effects are still significant, although the detailed structures on the retinal image are enhanced. Results of LPAR are shown in Fig. 4-8 (a4) to Fig. 4-8 (d4), where the haze effect is removed as the blood vessels can be clearly observed. However, the LPAR causes an overexposure effect on the bright area of the optical disk as shown in Fig. 4-8 (b4) and Fig. 4-8 (c4).

Note that some DVA results are similar to that of LPAR method as shown in Fig. 4-8 (b2) and Fig. 4-8 (b4), the main reason being that both belong to variations of Eq. (4-4). DVA uses TV regularization while the LPAR uses the Gaussian filter to obtain the background components of the retinal image. Meanwhile, the LPAR uses an adaptive strategy to determine the parameters for retinal image enhancement, thus, the LPAR has better performance on retinal images of different conditions than that of DVA method whose parameters still need hand-tuning.

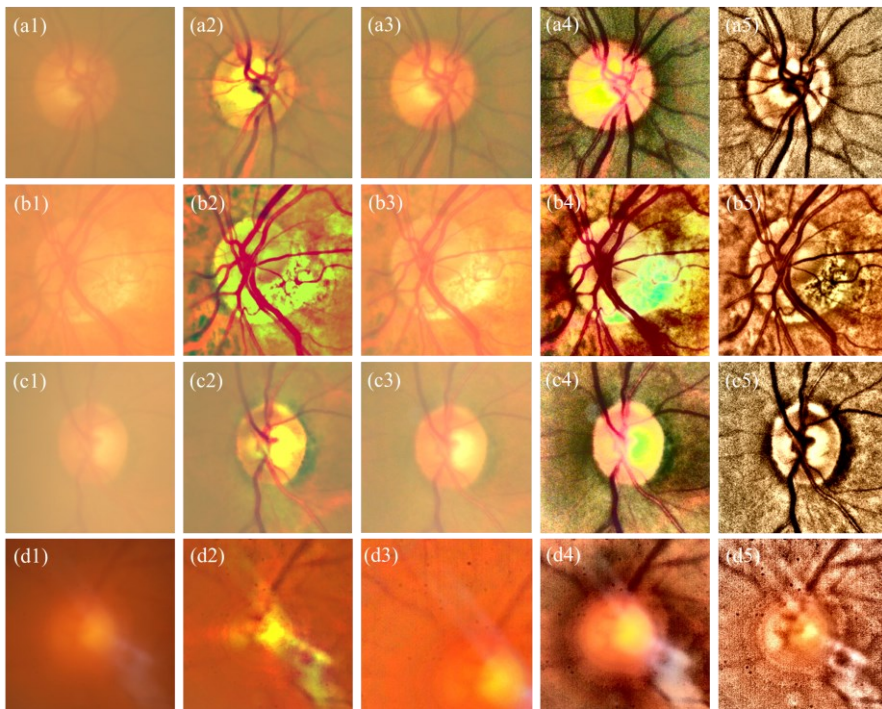


Fig. 4-9. Partially enlarged areas of raw images in the white boxes. Columns from left to right are raw images, enhanced results of DVA, DRC, LPAR, and MUTE methods.

Results of MUTE are shown in the last column of Fig. 4-8. The haze effect is efficiently suppressed and blood vessels that are hidden behind the cataract can be now observed. Meanwhile, MUTE also uses an adaptive manner to determine model parameters. It

does not lead to an overexposure effect, as it can restore a clear optical disk area as shown in Fig. 4-8 (b5) and Fig. 4-8 (c5).

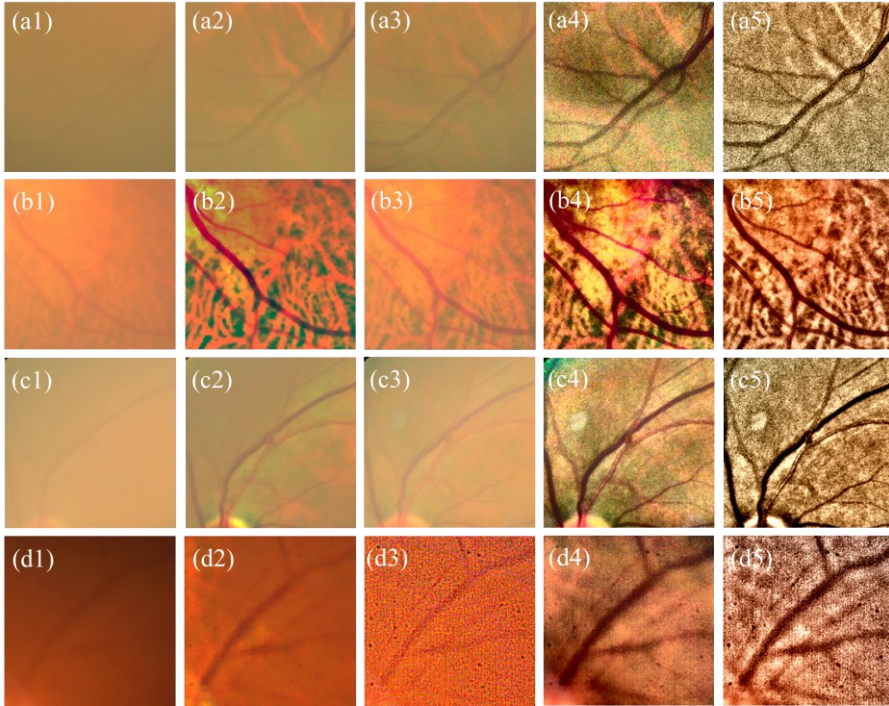


Fig. 4-10. Partially enlarged areas of raw images in the yellow boxes. Columns from left to right are raw images, enhanced results of DVA, DRC, LPAR, and MUTE methods.

Partially enlarged images for the areas in the white and yellow boxes are shown in Fig. 4-9 and Fig. 4-10, respectively. All four methods show their ability in contrast enhancement, while the DVA and LPAR methods lead to an overexposure effect at the bright area as shown in Fig. 4-9 (b2), (c2), (b4), and (c4). MUTE enhances image contrast and some small blood vessels within the optical disk can be observed as shown in Fig. 4-9 (b5) and Fig. 4-10 (a5) to Fig. 4-10 (d5). Some blood vessels that are hard to be observed in the raw images are now can be clearly found, especially in Fig. 4-9 (d5) and Fig. 4-10 (a5).

According to the experimental results, MUTE suppresses the haze effect of the cataractous retinal image and gains better contrast-enhanced results than other SOTA methods.

5.2 Objective assessments

To evaluate the performance of the four methods, we also conducted a quantitative evaluation in terms of (1) image sharpness, (2) image local contrast, (3) image entropy, (4) image multiscale contrast, and (5) image fog density. All images were reformatted in 8-bit unsigned integers (uint8, the gray-value is ranged in [0, 255]).

We adapted the Underwater Image Sharpness Measure (UISM) and Underwater Image Contrast Measure (UIConM) for image sharpness and image local contrast since severe blurring occurs in both underwater images and retinal images due to forward scattering [37]. More importantly, both the UISM and the UIConM do not rely on the statistical property of images and thus can be applied to retinal images, regardless of the statistical difference between retinal images and underwater images. The UISM and the UIConM are calculated by

$$UISM = \sum_{c \in \{R, G, B\}} \lambda_c \cdot \left[\frac{2}{k_1 k_2} \sum_{l=1}^{k_1} \sum_{k=1}^{k_2} \log \left(\frac{I_{\text{edge, max, } k, l}}{I_{\text{edge, min, } k, l}} \right) \right], \quad (4-15)$$

and

$$UIConM = \sum_{c \in \{R, G, B\}} \lambda_c \cdot \log \left[\frac{1}{k_1 k_2} \otimes \sum_{l=1}^{k_1} \sum_{k=1}^{k_2} \frac{I_{\text{max, } k, l} \ominus I_{\text{min, } k, l}}{I_{\text{max, } k, l} \oplus I_{\text{min, } k, l}} \log \left(\frac{I_{\text{max, } k, l} \ominus I_{\text{min, } k, l}}{I_{\text{max, } k, l} \oplus I_{\text{min, } k, l}} \right) \right]. \quad (4-16)$$

In Eq. (4-15), I_{edge} denotes the product between the Sobel edge map and the original image in the corresponding color channel. λ_c is the weight that is assigned to each color channel according to the visual response. $\lambda_R = 0.299$, $\lambda_G = 0.587$ and $\lambda_B = 0.114$ are used according to the relative visual responses of the red, green, and blue channels [37].

The image is divided into $k_1 k_2$ blocks, $I_{\max,k,l} / I_{\min,k,l}$ indicates the relative contrast ratio within each block.

Image entropy (IE) describes the randomness distribution of the image and its value denotes the amount of image information [38, 39], which is given by

$$IE = \sum_{c \in \{R, G, B\}} \lambda_c \sum_{g=0}^{255} P(x_g) \log(x_g), \quad (4-17)$$

where $P(x_g)$ is the probability of the appearance of the pixels that have gray-value g in the gray-scaled image. Image entropy can be used to characterize the texture of the image [40] and determine the amount of image information. Images affected by haze tend to have low IE values due to the biased brightness distribution. In contrast, haze-free images have a relatively high IE.

The multiscale contrast of the image, C_{RAMM} , was calculated with a pyramidal multi-resolution representation of luminance [40]. C_{RAMM} is defined as

$$C_{RAMM} = \frac{1}{8(\# \text{ pixel})(\# \text{ levels})} \sum_{\forall \text{ level}} \left[\sum_{\forall \text{ pixel}} \left(\sum_{8\text{-neigh}} |P_i - P_j| \right) \right], \quad (4-18)$$

where $\# \text{ pixel}$ denotes the total pixel numbers of the image, $\forall \text{ level}$ denotes the total level of down-sampling, in each level the image is halved without pre-filtering. In this paper, we use 6 levels of down-sampling. The pixel numbers of retinal images in the databases are large enough for 6 levels of down-sampling.

Lastly, the fog-aware density evaluator (FADE) [41] [42] was used to numerically predict perceptual hazy density.

Tab. 4-1 demonstrates the enhancement results of the quantitative evaluations using objective metrics. Here, High values of UISM, UIConM, IE, and CRAMM indicate high image quality, while low values of FADE indicate a better dehazing performance.

Numbers marked in blue denote the best performance, while numbers marked in green denote the second-best performance. As listed in Tab. 4-1, MUTE performed best

in aspects of UISM, UIConM, IE, CRAMM, and FADE, as it gained better scores than the other three methods. The DRC, LPAR, and MUTE decrease the FADE value denoting good dehazing results.

Tab. 4-1. Quantitative assessment of the enhanced images among Fig. 4-8 (a1) to 4-8 (d1) with five metrics.

Raw image	Matrixes	Raw	Methods			
			DVA	DRC	LPAR	MUTE
Figure 8 (a1)	UISM	0.7782	0.9069	2.9260	4.6992	6.7863
	UIConM	0.0578	0.0938	0.0780	0.2909	0.3851
	IE	5.2354	5.1692	5.6260	7.4339	7.5100
	CRAMM	1.1207	1.5072	1.4048	4.7888	7.5293
	FADE	0.5127	1.1202	0.6086	0.2214	0.1554
Figure 8 (b1)	UISM	0.8098	2.2411	3.2962	7.0178	5.9409
	UIConM	0.0739	0.1950	0.1129	0.3534	0.3759
	IE	5.8075	6.2273	6.0271	7.2154	7.3425
	CRAMM	1.3705	3.7879	1.9448	7.1645	8.8383
	FADE	0.4347	0.6151	0.4590	0.1746	0.1191
Figure 8 (c1)	UISM	0.7715	0.9615	2.9473	4.4087	7.1608
	UIConM	0.0745	0.0968	0.0914	0.2970	0.3876
	IE	5.0473	4.9388	5.4631	7.3920	7.4646
	CRAMM	1.2978	1.5133	1.5083	4.9293	8.0858
	FADE	0.6148	1.3645	0.7896	0.2260	0.1427
Figure 8 (d1)	UISM	1.1770	1.5114	3.3452	4.3085	8.1445
	UIConM	0.0274	0.0805	0.1279	0.2186	0.4020
	IE	6.1477	6.2052	6.3507	7.3024	7.5447
	CRAMM	0.7426	1.4215	1.9617	3.8497	8.4174
	FADE	0.7134	0.7112	0.2980	0.2129	0.1183
Average	UISM	0.8841	1.4052	3.1287	5.1086	7.0081
	UIConM	0.0584	0.1165	0.1026	0.2900	0.3877
	IE	5.5595	5.6351	5.8667	7.3359	7.4655
	CRAMM	1.1329	2.0575	1.7049	5.1831	8.2177
	FADE	0.5689	0.9528	0.5388	0.2087	0.1339

Tab. 4-2. Averages and standard deviations of enhancement results from both public and proprietary cataractous retinal image datasets.

Databases mean \pm std.	Matrixes	Raw	Methods				
			DVA	DRC	LPAR	MUTE	
Cataract (100 images)	UISM	0.9002	2.2829	3.4599	5.5883	6.7884	
		± 0.1432	± 0.9623	± 1.4614	± 0.7461	± 0.7869	
	UIConM	0.0573	0.1752	0.1097	0.3257	0.3890	
		± 0.0123	± 0.0532	± 0.0437	± 0.0294	± 0.0093	
	IE	5.8443	6.1823	6.2841	7.4230	7.4151	
		± 0.4449	± 0.4350	± 0.3819	± 0.1960	± 0.1338	
	CRAMM	1.2013	2.9870	1.9515	6.2711	8.1483	
		± 0.1741	± 1.0950	± 0.4257	± 0.8769	± 0.6035	
	FADE	0.4715	0.6769	0.4502	0.2031	0.1469	
		± 0.1155	± 0.2548	± 0.1388	± 0.0321	± 0.0228	
	Proprietary (94 images)	UISM	1.2065	2.2897	3.8239	3.9118	6.8936
			± 0.3028	± 0.9169	± 0.7211	± 0.9212	± 1.1603
UIConM		0.0446	0.1477	0.1983	0.2140	0.3811	
		± 0.0133	± 0.0549	± 0.0640	± 0.0511	± 0.0646	
IE		6.8377	6.8649	6.7774	7.3770	7.4889	
		± 0.5775	± 0.4550	± 0.4714	± 0.1914	± 0.4049	
CRAMM		1.0378	2.4631	2.8145	3.8878	7.9087	
		± 0.1741	± 0.9972	± 0.7723	± 1.2095	± 1.5078	
FADE		1.0455	0.9345	0.2070	0.3351	0.1857	
		± 0.3905	± 0.5244	± 0.2112	± 0.1118	± 0.1177	
Average (194 images)		UISM	1.0486	2.2862	3.6363	4.7760	6.8393
			± 0.2798	± 0.9381	± 0.6049	± 1.1832	± 0.9844
	UIConM	0.0512	0.1619	0.1527	0.2716	0.3852	
		± 0.0143	± 0.0557	± 0.0657	± 0.0695	± 0.0455	
	IE	6.3164	6.5057	6.5231	7.4045	7.4568	
		± 0.7181	± 0.5639	± 0.4929	± 0.1952	± 0.2996	
	CRAMM	1.1221	2.7332	2.3696	5.1163	8.0322	
		± 0.1920	± 1.0785	± 0.7532	± 1.5891	± 1.1388	
	FADE	0.7496	0.8017	0.4528	0.2671	0.1657	
		± 0.4038	± 0.4272	± 0.1772	± 0.1045	± 0.0856	

Tab. 4-2 compares the average and standard deviation among five quality matrixes for raw and enhanced images from both public and proprietary cataractous retinal image datasets. In general, the performance of MUTE is better than those of the other three methods, as it has the best average scores among all four methods. The robustness

of MUTE is also better than the other three SOTA methods which are implied by the small standard deviation of MUTE.

Apart from the cataractous retinal image, we test the performance of MUTE on non-cataractous retinal images from public databases of DiaRet- , [43, 44] STARE-, MESSIDOR- [45], and NORMAL- [46] datasets, adding up to a total of 2116 retinal images.

The results are listed in Tab. 4-3. For non-cataractous retinal images, DVA had the best score on IE, while MUTE showed the best score on UISM, UICoM, C_{RAMM} , and FADE. In general, MUTE increases the contrast of retinal images by 7-fold, while decreasing their haze density by more than about 3-fold. The small standard deviation value of MUTE denotes its algorithm robustness. More experimental results for visual assessments are available in **Supplementary Note 4-6** and **Supplementary Note 4-7**. Comparison results on the cataractous dataset between MUTE and SOTA deep-learning methods are available in **Supplementary Note 4-12** for both visual and objective assessments.

In the following subsection, we present demonstrates of MUTE's potential contribution to three clinical applications including blood vessel segmentations, retinal image registration, and retinopathy diagnosis.

Tab. 4-3. Averages and standard deviations of enhancement results from non-cataractous retinal image databases

Databases mean \pm std.	Matrixes	Raw	Methods			
			DVA	DRC	LPAR	MUTE
DiaRetDB0+DB1 (219 images)	UISM	3.2857 \pm 0.2107	3.5914 \pm 0.5870	4.2797 \pm 0.3261	5.3378 \pm 0.8601	6.7373 \pm 0.3610
	UIConM	0.0396 \pm 0.0140	0.1793 \pm 0.0393	0.0989 \pm 0.0389	0.2419 \pm 0.0471	0.3977 \pm 0.0068
	IE	5.6481 \pm 0.3389	6.2941 \pm 0.2314	5.7972 \pm 0.3434	5.8912 \pm 1.8016	6.6910 \pm 0.1443
	C _{RAMM}	0.7169 \pm 0.1332	2.4855 \pm 0.7205	1.4165 \pm 0.4100	3.7898 \pm 1.2053	8.0577 \pm 0.5667
	FADE	0.4752 \pm 0.0395	0.4188 \pm 0.0648	0.3094 \pm 0.0461	0.2495 \pm 0.1344	0.1278 \pm 0.0156
STARE (397 images)	UISM	2.6542 \pm 0.6460	4.1758 \pm 1.0861	4.6212 \pm 0.5306	5.5971 \pm 1.0953	6.5124 \pm 0.8229
	UIConM	0.0879 \pm 0.0195	0.2440 \pm 0.0458	0.1930 \pm 0.0444	0.2863 \pm 0.0499	0.3693 \pm 0.0180
	IE	5.4185 \pm 0.3468	6.3837 \pm 0.2481	5.5470 \pm 0.3198	5.5085 \pm 1.2873	5.8185 \pm 0.2760
	C _{RAMM}	1.9010 \pm 0.2998	5.1454 \pm 1.4516	3.3505 \pm 0.7190	5.4806 \pm 1.2965	7.8314 \pm 0.7522
	FADE	0.4744 \pm 0.1274	0.3947 \pm 0.1296	0.2367 \pm 0.0617	0.2739 \pm 0.1453	0.1331 \pm 0.0236
MESSIDOR (1200 images)	UISM	1.8419 \pm 0.3224	4.0235 \pm 0.7075	5.0035 \pm 0.4089	6.2763 \pm 0.6584	7.0831 \pm 0.4933
	UIConM	0.0746 \pm 0.0185	0.2383 \pm 0.0318	0.1608 \pm 0.0418	0.2777 \pm 0.0280	0.3660 \pm 0.0073
	IE	5.1908 \pm 0.3454	6.2726 \pm 0.1881	5.6377 \pm 0.3846	5.9916 \pm 0.7165	6.0725 \pm 0.2111
	C _{RAMM}	1.1810 \pm 0.1639	3.9134 \pm 0.7737	2.2784 \pm 0.4610	5.7179 \pm 0.7411	7.6652 \pm 0.5981
	FADE	0.4652 \pm 0.0673	0.4361 \pm 0.0842	0.2320 \pm 0.0410	0.1812 \pm 0.0726	0.1252 \pm 0.0136
NORMAL (300 images)	UISM	1.5512 \pm 0.2604	3.4903 \pm 0.7928	4.7259 \pm 0.3344	5.7581 \pm 0.6043	6.9285 \pm 0.4592
	UIConM	0.0768 \pm 0.0152	0.2389 \pm 0.0342	0.1700 \pm 0.0336	0.3092 \pm 0.0267	0.3733 \pm 0.0047
	IE	5.5313 \pm 0.2497	6.4120 \pm 0.1878	5.8949 \pm 0.2886	6.3338 \pm 0.7880	6.3062 \pm 0.1898
	C _{RAMM}	1.2468 \pm 0.1495	3.9861 \pm 0.8219	2.5810 \pm 0.4295	6.3165 \pm 0.8719	8.3668 \pm 0.4882
	FADE	0.6671 \pm 0.1226	0.5629 \pm 0.1132	0.3143 \pm 0.0622	0.2096 \pm 0.0876	0.1497 \pm 0.0186
Average (2116 images)	UISM	2.1025 \pm 0.6595	3.9318 \pm 0.8274	4.8175 \pm 0.4819	5.9783 \pm 0.8516	6.9183 \pm 0.5981
	UIConM	0.0738 \pm 0.0219	0.2334 \pm 0.0404	0.1618 \pm 0.0478	0.2801 \pm 0.0390	0.3709 \pm 0.0137
	IE	5.3291 \pm 0.3735	6.3154 \pm 0.2130	5.6736 \pm 0.3731	5.9391 \pm 1.0404	6.1220 \pm 0.3221
	C _{RAMM}	1.2774 \pm 0.3847	4.0071 \pm 1.1709	2.4333 \pm 0.7364	5.5587 \pm 1.1420	7.8364 \pm 0.6612
	FADE	0.4966 \pm 0.1128	0.4445 \pm 0.1094	0.2525 \pm 0.0599	0.2097 \pm 0.1062	0.1304 \pm 0.0188

5.3 Potential clinical applications: Retinal image blood vessels segmentations

Retinal image blood vessels segmentation provides the shape distributions of blood vessels, which is important for clinical diagnosis as the morphological changes of blood vessels are biomarkers for diseases such as lacunar stroke [3], cognitive dysfunction [47], cardiovascular risk [5], diabetes [48] and glaucoma [49]. Blood vessel segmentation can be achieved by either human specialists or computer software. The former provides accurate results but is time-consuming. The latter option provides fast segmentation results but is less accurate than human specialists. Moreover, due to poor image contrast of the cataractous retinal image, hand-based segmentation is even more time-consuming, and automatic segmentation for hazy retinal images can be error-prone.

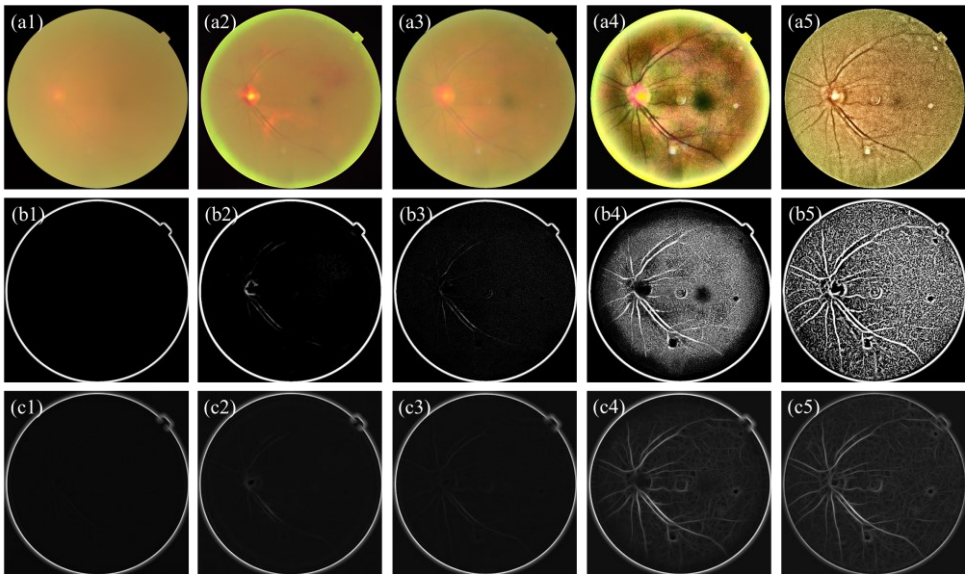


Fig. 4-11. Demonstration of retinal blood vessels segmentation for the cataractous retinal image. (a1) to (a5) are raw, and enhanced images by DVA, DRC, LAPR, and MUTE, respectively. (b1) to (b5) are segmentation results of the Frangi filter. (c1) to (c5) are results for the Match filter.

Fig. 4-11 (a1) to Fig. 4-11 (c1) demonstrates the raw image and segmentation results given by two filtering-based methods including the Frangi filter [50] [Fig. 4-11 (b1)] and the match filter [51] [Fig. 4-11 (c1)]. As shown in Fig. 4-11 (b1) and Fig. 4-11 (c1), the blood vessels have a very weak response to the filters due to the hazy effect, resulting in the failure of blood vessels segmentation.

Without other image preprocessing, enhancing the raw image using the DVA, DRC, LPAR, and MUTE methods before segmenting increases the performance of segmentation results in different degrees. The results for DVA and DRC are still limited as shown in Fig. 4-11 (b2) to Fig. 4-11 (c3), while the two methods are able to detect the vessels' branches for image enhanced by LPAR and MUTE as shown in Fig. 4-11 (b4) to Fig. 4-11 (c5).

The response of blood vessels to the Frangi filters in Fig. 4-11 (a5) is higher than that of Fig. 4-11 (a5) as the intensity of blood vessels in Fig. 4-11 (b5) are higher than that in Fig. 4-11 (b4). This implies that the segmentation performance in Fig. 4-11 (a5) is better than Fig. 4-11 (a4). Note that the Frangi filter fails to detect the blood vessels near the edge of the ROI for LPAR enhanced image due to the halo effect as shown in Fig. 4-11 (a4).

Another clinical application is shown in Fig. 4-12, reporting a cataractous retinal image with poor illumination conditions. Both Frangi filter and match filter are able to detect more blood vessels' branches for Fig. 4-12 (a5) than other enhanced images shown in Fig. 4-12 (b1) to Fig. 4-12 (c4).

As there is no ground-truth blood vessel map for real cataractous retinal images since manually blood vessel segmentation is extremely hard due to the lack of image contrast for cataractous retinal images, we can only present the visual assessment to compare the segmentation results. In order to quantitatively present that MUTE benefits blood vessel segmentation, we perform simulation experiments based on simulated cataractous retinal images with well-defined ground-truth blood vessel maps in **Supplementary Note 4-8** in detail. In general, our experimental and simulation

results show that MUTE improves the blood vessel segmentation in both visual and objective aspects.

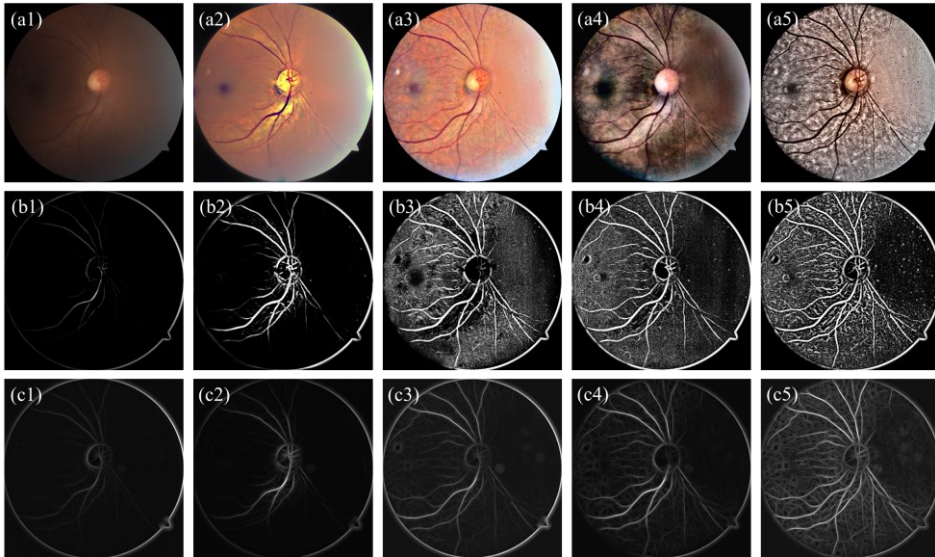


Fig. 4-12. Demonstration of retinal blood vessels segmentation for cataractous retinal image with poor illumination condition. (a1) to (a5) are raw, and enhanced images by DVA, DRC, LAPR, and MUTE, respectively. (b1) to (b5) are segmentation results of the Frangi filter. (c1) to (c5) are results for the Match filter.

5.4 Potential clinical applications: Retinal image registration

Image registration is an important application as it provides in the fields of computer vision, pattern recognition, and medical image analysis. It aligns two or more retinal images together in the same spatial axis to provide an overall comprehensive understanding. A promising retinal image registration relies on precise feature detecting and matching for images to be registered. Registration of cataractous retinal images can be failed due to poor feature paring results as the features are hidden by the haze effect of cataracts.

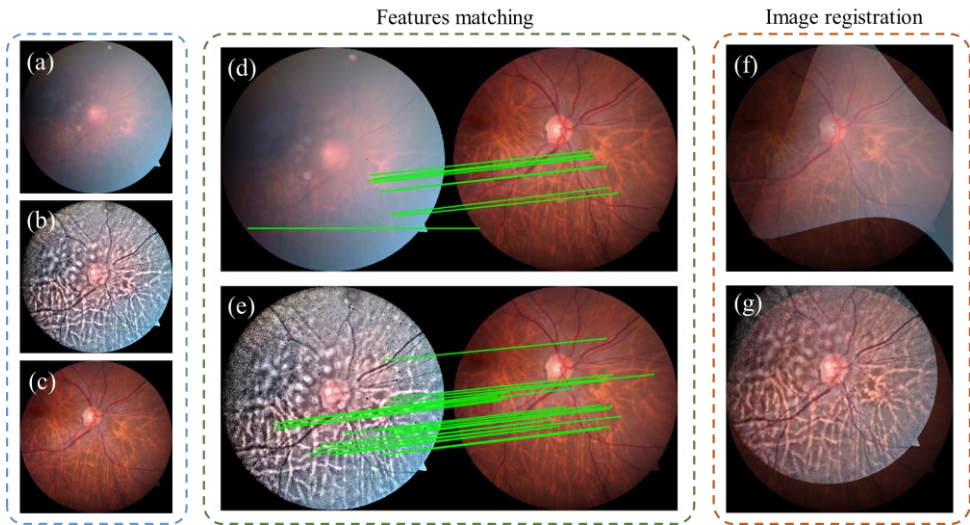


Fig. 4-13. Demonstration of retinal image registration with and without enhanced by MUTE. (a) to (c) are raw images, enhanced images, and post-surgery images, respectively. (d) indicates the matched features for (a) and (c), while (e) is the matched features for (b) and (c). (f) and (g) are montages of retinal image registration.

Fig. 4-13 compares the registration results before and after being processed by MUTE. Fig. 4-13 (a) is the retinal image before cataract surgery, while Fig. 4-13 (b) is the enhanced image. Fig. 4-13 (c) is the retinal image after cataract surgery. We registered retinal image using Gaussian Field Estimator with Manifold Regularization (GFMR) method [52], Fig. 4-13 (d) shows the feature matching results between images in Fig. 4-13 (a) and Fig. 4-13 (c). The feature matching results are problematic as the algorithm failed to find the correct paired features in both images (indicated by green lines), resulting in a distorted registration as shown in Fig. 4-13 (f). Fig. 4-13 (e) shows the feature matching results between images in Fig. 4-13 (b) and (c) where the algorithm pinpoints the paired features as indicated by green lines. Successfully registered images are shown in Fig. 4-13 (g).

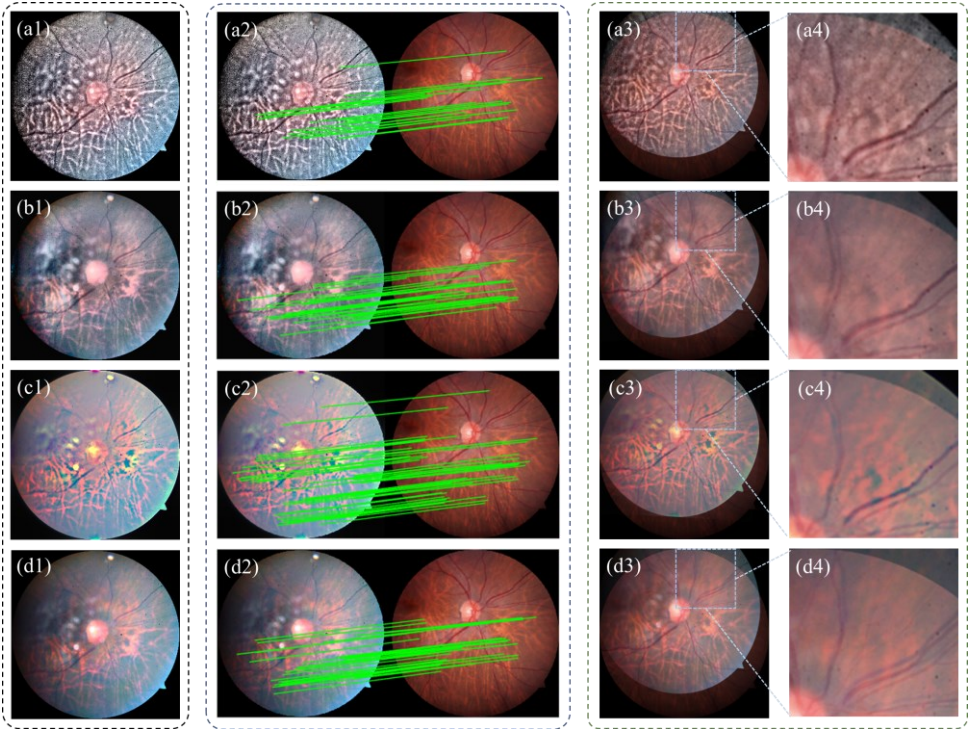


Fig. 4-14. Comparison of registration results for (a1-a3) MUTE, (b1-b3) LPAR, (c1-c3) DVA and (d1-d3) DRC methods. (a1), (b1), (c1), and (d1) are enhancements. (a2), (b2), (c2), and (d2) are paired features (a3), (b3), (c3), and (d3) are registration results. (a4) to (d4) are zoomed-in part in the blue boxes.

Fig. 4-14 shows the comparison of registration results for images enhanced by different methods. The GFMR can find sufficient paired features to perform registration based on MUTE, LPAR, and DVA outputs, however, according to the visual assessment in Fig. 4-14 (a3) to Fig. 4-14 (d4) the registration accuracy for MUTE is better than other methods. For example, as shown in Fig. 4-14 (a4) to Fig. 4-14 (c4), the blood vessels are correctly registered for MUTE [Fig. 4-14 (a4)] where the blood vessels overlapped in both enhanced image and post-surgery images, while the blood vessel is not completely overlapped for LPAR, DVA, and DRC as shown in Fig. 4-14 (b4) to (d4), respectively.

As there is no ground-truth registration image to quantitatively evaluate the registration quality, we perform a simulation experiment in **Supplementary Note 4-9** in detail, and the results show that MUTE improves the registration quality also in objective aspects.

In general, MUTE enhanced the contrast of the image by suppressing the haze effect of cataractous retinal images. As more retinal structures are revealed during the dehazing process, more paired features can be detected by GFMR, increasing the success rate of cataractous retinal image registration.

5.5 Potential clinical applications: Retinopathy diagnosis

Finally, we focus on the enhancement results in the areas that indicate retinopathy like hard exudates and hemorrhages. The enhancement results should increase the visual quality of the retinopathy area while do not create unexpected artifacts to guarantee structure fidelity.

We first test MUTE on the DiaRetDB1 dataset, since the retinal images in this dataset are not as severely degraded as that of cataractous retinal images, and retinopathies are already marked by human specialists shown in Fig. 4-15 (b) and Fig. 4-15 (c). As shown in Fig. 4-15 (d1) to Fig. 4-15 (d4), MUTE increases the visual quality of the hard exudates. Some hard exudates which can hardly be observed in the raw image shown in Fig. 4-15 (d3) can be clearly found in the enhanced images in Fig. 4-15 (d4) due to the increase of image contrast. MUTE also increases the visual quality of hemorrhages as shown in Fig. 4-15 (e1) to Fig. 4-15 (e4), as the contrast between hemorrhage areas and the background increases in enhanced images [see Fig. 4-15 (e2) and (e4)].

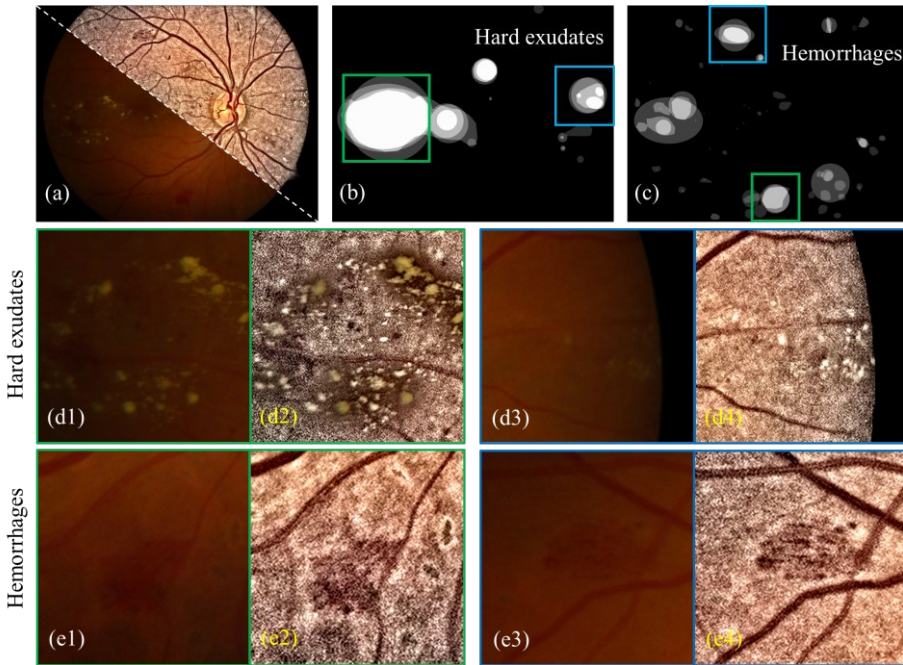


Fig. 4-15. Enhancement of retinopathy areas of MUTE. (a) Montage of raw and enhanced images. (b) labels of hard exudates areas. (c) labels of hemorrhage areas. (d1) to (d4) are enlarged parts of raw and enhanced images corresponding to green and blue boxes in (b). (e1) to (e4) are enlarged parts for (c).

Then we tested MUTE on cataractous retinal images before and after cataract surgery. The retinopathy is manually marked based on the post-surgery retinal images as shown in Fig. 4-16 (b) and (f) for Groups Drusen and RPE Alteration, respectively. After image registrations, we enlarged the area in the yellow box where there is retinopathy (Drusen, and RPE alteration in these cases) for raw images and enhanced images. As shown in Fig. 4-16 (d1), the drusen in the raw images have visual quality due to cataracts, while it can be clearly observed in the post-surgery image shown in Fig. 4-16 (d2), and so does the RPE alteration shown in Fig. 4-16 (e) and Fig. 4-16 (f). After image enhancement by different methods, the visual quality of the retinopathy areas increases, and the MUTE gains a better performance than the other three methods as it

significantly increases the image contrast which has also been shown in Tab 4-2 and Tab 4-3.

MUTE increases the contrast of cataractous retinal images which may potentially facilitate the detection of retinopathy. Meanwhile, by comparing MUTE against the raw image, post-surgery image, and enhanced image by the other three methods, we found that MUTE won't lead to additional artifact structure on enhanced images and guarantees the information fidelity.

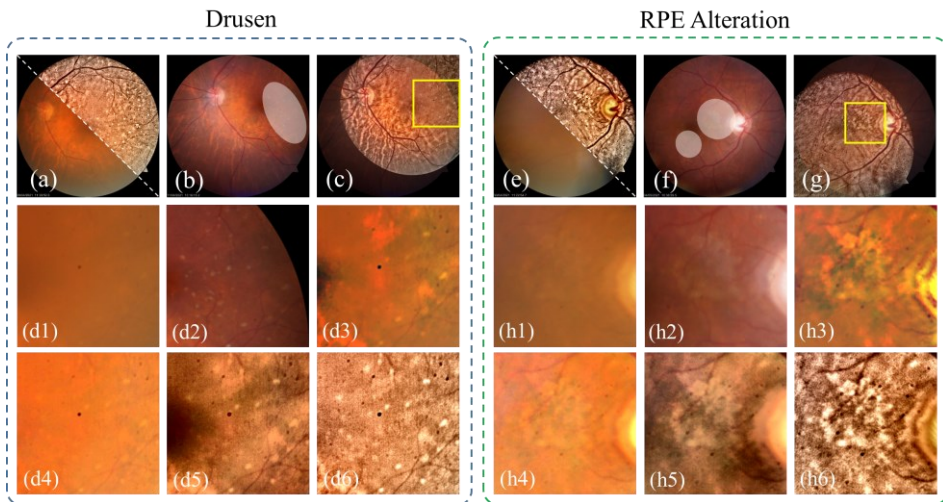


Fig. 4-16. Demonstration of enhancement for cataractous retinal images, with Drusen and RPE alteration. (a) and (e) are Montages of raw and enhanced images by MUTE. (b) and (f) are post-surgery retinal images marked by specialists. (c) and (g) denotes the image registration between post-surgery images and enhanced images. (d1) to (d6) are enlarged parts in the yellow box in (c) for raw image, and post-surgery image, enhanced by DVA, DRC, LPAR and MUTE respectively. (h1) to (h6) are zoomed-in for yellow box in (g).

6. Ablation Study

In this subsection, we perform ablation studies on the number of denoising layers, the type of denoising layers, the selection of I_0 values, the function of illumination correction, and the treatment in Pb/Pr channels.

6.1 On the denoising layers

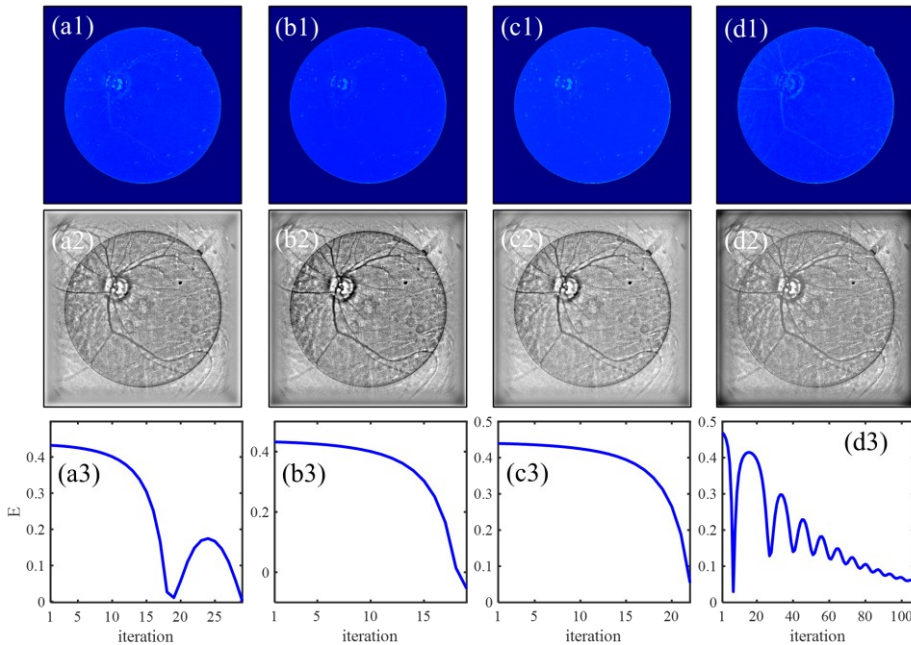


Fig. 4-17. Dehazing of Fig. 3 (d3) with different levels of Gaussian denoising. (a1) to (d1) are T_{sc} for $\mathcal{L}=2$, $\mathcal{L}=4$, $\mathcal{L}=6$, and $\mathcal{L}=8$. (a2) to (d2) are corresponding dehazed Y channel. (a3) to (d3) are evolution of cost function.

Fig. 4-17 shows dehazing results of the image in Fig. 4-3 (d3) with different total numbers of Gaussian denoising levels. More levels give a more refined estimation of T_{sc} but increase the computational complexity. The visual quality of the dehazed image in Fig. 4-17 (b2) is better than that in Fig. 4-17 (d2), which still has a slight haze effect. The cost function for Fig. 4-17 (d2) is oscillating and decreasing with respect to iteration

numbers. This means that the optimization process for Fig. 4-17 (d2) is unstable due to large numbers of undermined parameters. According to the experimental trial, we found that $\mathcal{L} \in [4, 6]$ was appropriate for most of the cataractous retinal images.

Our proposed dehazing method is not sensitive to the choice of \mathcal{S}_i . One can use total variance regularization [53] and its variants, or image gradient L_0 -norm regularization [54] for alternative methods when calculating \mathcal{S}_i . **Supplementary Note 4-11** compares the results of MUTE with different denoisers including Domain transform recursive filter and total-variations. The intermedia output of the parameter \mathcal{A} and $\varphi_i(\mathbf{r})$ are also shown in detail.

In an extreme case, we can even simply let $\mathcal{L} = 1$ and $\mathcal{S}_0(\mathbf{r}) = \mathcal{S}_1(\mathbf{r}) = 1$ since the uniform pattern can be regarded as an extremely denoised image. In this case, as shown in Fig. 4-18, the adaptive pixel-wise stimulated algorithm still returns an appropriate estimation of T_{sc} together with dehazed O_Y .

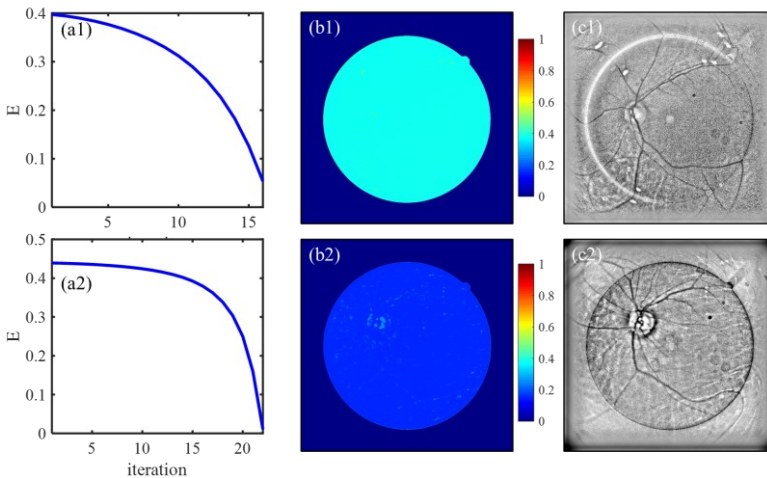


Fig. 4-18. Results for Y channel dehazing when $\mathcal{L} = 2$ and $\mathcal{S}_1(\mathbf{r}) = \mathcal{S}_2(\mathbf{r}) = 1$ for the image in Figs 3(c3) and (d3). (a1) and (a2) are evolution of cost functions. (b1) and (b2) are T_{sc} . (c1) and (c2) are O_Y .

6.2 On the selection of I_0

To show the reason why $I_0 = 0.52$ is used in our experiments, we perform a statistical analysis of the Y channel average pixel value for raw images, images after illumination correction, and final outputs of MUTE. The analysis is applied to 194 cataractous retinal images (100 images in the CATARACT dataset, and 95 images in our proprietary dataset), and 1200 retinal images from the MESSIDOR dataset.

As shown in Fig. 4-19 (a1) and Fig. 4-19 (a2), the intensity of raw images from both the cataract dataset and the MESSIDOR dataset varies from 0.2 to 0.6. The normalized probability density functions (PDF) are fitted with Gaussian distribution (plotted in red curves) which has large standard variations. The distribution pattern implies the fact that not all retinal images are collected under good conditions, some of them suffer from bad and insufficient illumination problems.

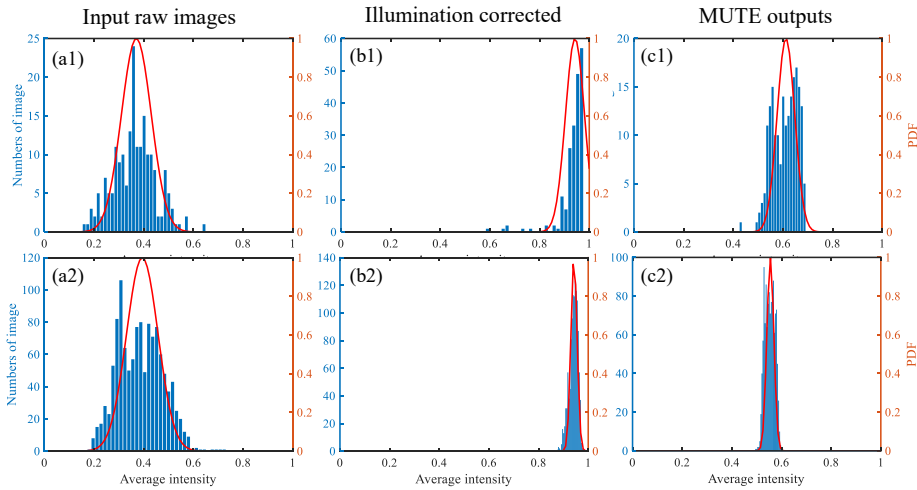


Fig. 4-19. Distribution of Y channel average value for (a1), (a2) raw images; (b1), (b2) illumination corrected images, and (c1), (c2) MUTE output images. (a1) to (c1) are results for cataractous retinal images (194 images included). (a2) to (c2) are results for MESSIDOR dataset (1200 images included).

The distribution patterns after the illumination correction process are shown in Fig. 4-19 (b1) and (b2), where the distributions are concentrated near 0.95. Since the

illumination correction is performed in an adaptive manner, it is suitable for retinal images of different illumination conditions. Moreover, the adaptive illumination correction process has a similar function as intensity normalization - that is adjusting the intensity of raw images to similar levels. This feature will benefit the subsequent dehazing process.

As the dehazing process is performed based on the intensity cost function and the dehazing always decreases the intensity level, we consider the target intensity level being a constant ratio I_0 between the corrected intensity level and the final output intensity level. As shown in Fig. 4-19 (a1) and (a2) the average Y channel value for the raw image approaches 0.4, and the average Y channel value for the illumination corrected images approach 0.95, we, therefore, can let $I_0 \approx 0.4 / 0.95 = 0.4211$.

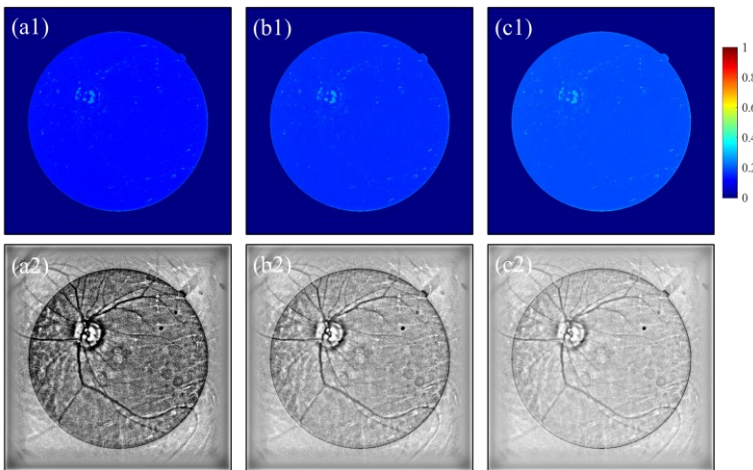


Fig. 4-20. Results for Y channel dehazing when $\mathcal{L} = 4$ with different values of I_0 (a1) to (c1) are T_{sc} with $I_0 = 0.35$, $I_0 = 0.55$, and $I_0 = 0.65$, respectively. (a2) to (c2) are dehazed Y channels.

In our experiment, we choose $I_0 = 0.52$ to let the intensity level for the output images becomes larger than 0.4211. The distribution for the Y channel average value of MUTE outputs are shown in Fig. 4-19 (c1) and (c2). Moreover, Fig. 4-20 shows the

dehazed results with different values of I_0 . In general, a small value of I_0 increases the degree of dehazing as the hazed effect is significantly suppressed. However, the image intensity also decreases. According to the experimental results in Fig. 4-20 and statistical analysis, we believe that $I_0 = 0.52$ is an appropriate setting.

Finally, according to Eq. (4-3), the Pb/Pr channel can be restored through an illumination correction process with known of $L(\mathbf{r}) \circ T_{lens}^2(\mathbf{r})$ from illumination correction in Section 3.2, and $T_{sc}^2(\mathbf{r})$ from dehazing in Section 4. According to the experiments (**Supplementary Note 4-5**), the Pb/Pc treatment can be skipped since this process won't influence the final dehazing quality of MUTE. Please refer to **Supplementary Note 4-5** for detailed experiment results.

6.3 On the necessity of illumination correction

As shown in Section 6.2, the illumination correction will normalize the intensity of various input retinal images into similar intensity levels approach 0.95. This high-intensity level gives sufficient intensity adjusting tolerance for the dehazing process since image dehazing always decreases the intensity level of the image. As such, the illumination correction benefits MUTE dehazing. If we skip the illumination correction process, the MUTE may not output good dehazing retinal images.

For example, if the input retinal images have low-intensity levels, the dehazing effect is not significant and may cause bad performance since the intensity level of input raw images is already very low. Further dehazing will decrease the image intensity and will distort image colors.

For retinal images that have relatively good illumination conditions, if we skip the illumination correction, we need to manually adjust parameter I_0 according to individual input images to optimize the dehazing results since the average intensity of the Y channel varies from 0.2 to 0.6. For example, Fig. 4-21 shows the dehazing results without illumination correction. The input image is shown in Fig. 4-21 (a) which has good and uniform illumination conditions. Its Y channel is shown in Fig. 4-21 (b) where

the mean value is 0.5357. Fig. 4-21 (c1) and Fig. 4-21 (c2) show the dehazed Y channel and T_{sc} for $I_0 = 0.1$ and $I_0 = 0.2$, respectively. The MUTE outputs are shown in Fig. 4-21 (d1) for $I_0 = 0.1$ and 21 (d2) for $I_0 = 0.2$, where dehazing effects for both images are not significant, as the MUTE overestimates the transmission matrix T_{sc} as shown in Fig. 4-21 (c3) and Fig. 4-21 (c4).

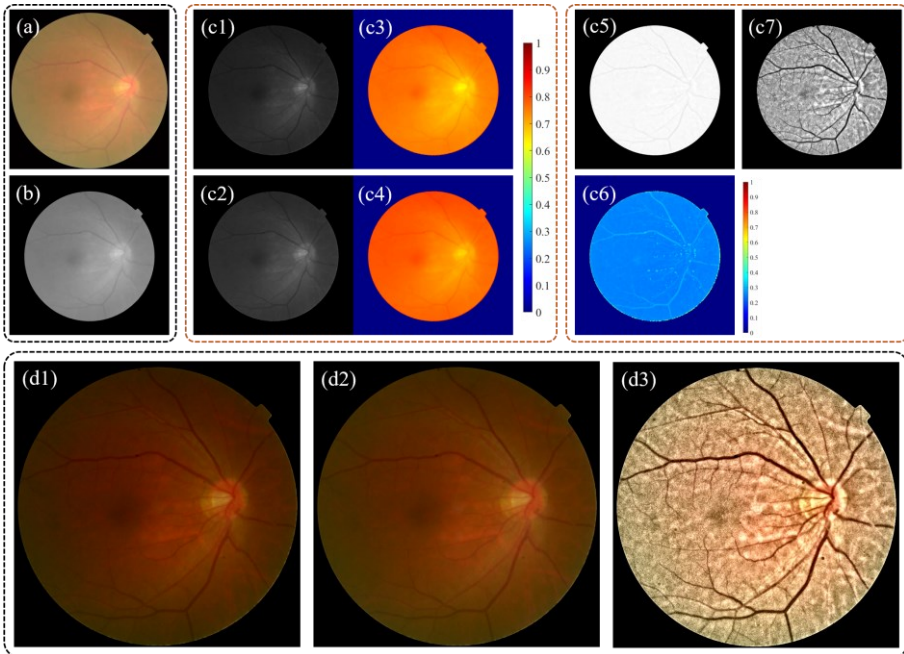


Fig. 4-21. MUTE dehazing output without and with illumination correction process. (a) raw input image. (b) Y channel of (a). Without the illumination correction process, (c1) and (c3) are dehazed Y channel and T_{sc} with $I_0 = 0.1$. (c1) and (c3) are dehazed Y channel and T_{sc} with $I_0 = 0.2$. The MUTE outputs are in (d1) and (d2). With illumination correction, the illumination-corrected Y channel, T_{sc} and dehazed Y channel are shown in (c5), (c6), and (c7), respectively. (d3) is the corresponding MUTE output.

With illumination correction as shown in Fig. 4-21 (c5), the Y channel is at a high-intensity level where the mean value is 0.9721 and has a sufficient intensity range for the dehazing process to darken the image. The transmission map can be correctly

estimated as shown in Fig. 4-21 (c6), $I_0 = 0.52$. The dehazed Y channel is shown in Fig. 4-21 (c7). The final output is shown in Fig. 4-21 (d3), where the contrast of the image is improved.

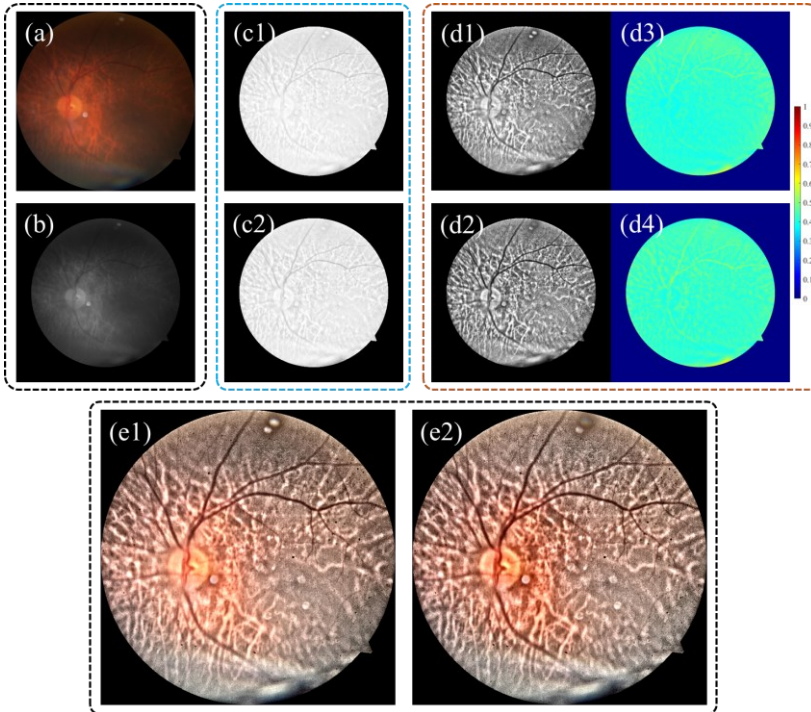


Fig. 4-22. MUTE dehazing output without and with coarse illumination correction process. (a) raw input image. (b) Y channel of (a). (c1) and (c2) are outputs of illumination correction without and with coarse illumination correction. (d1) to (d4) are MUTE dehazing for (c1) and (c2). (e1) and (e2) are final outputs.

As proposed in **Supplementary Note 4-2**, our adaptive illumination correction contains two steps: coarse illumination correction and fine illumination correction. The first step uses Retinex to correct the uneven illumination pattern so that the illumination pattern over the entire image is flat. Without the first step, the dehazing quality of MUTE will be degraded for retinal images that have uneven illumination conditions.

For example, as shown in Fig. 4-22 (a) the raw image has an uneven illumination pattern where the intensity near the optical disk is higher than the areas far from the optical disk. Fig. 4-22 (c1) and (c2) show the output of intensity correction without and with coarse illumination correction. Although the intensity level can still be corrected based on the fine illumination correction step, we can still find that the intensity level near the optical disk in Fig. 4-22 (c1) is higher than the intensity in other places, and so does Fig. 4-22 (d1) for dehazed Y channel. The final outputs are shown in Fig. 4-22 (e1) and (e2) without and with coarse illumination correction, where the dehazing quality near the optical disk in Fig. 4-22 (e2) is better than that of Fig. 4-22 (e1). In general, the illumination correction process is essential for MUTE dehazing, while each step of illumination correction benefits the dehazing quality of MUTE.

7. Suppressing noise effect of MUTE using Hessian regularization

For retinal images that have poor illumination and heavy haze effects, the enhancement algorithm increases the contrast of the image while also enlarging the impact of noises simultaneously. A denoising process can be involved after Eq. (4-8) is finished. For biologic images, we would recommend using L_1 -norm Hessian (second-order gradient) regularization for image denoising, which is given by minimizing the following equation

$$\mathbf{O}_{dn} = \arg \min \left\| \mathbf{O} - \mathbf{O}_{dn} \right\|_2^2 + \lambda \left(\left\| \nabla_{xx} \mathbf{O}_{dn} \right\|_1 + \left\| \nabla_{yy} \mathbf{O}_{dn} \right\|_1 + 2 \left\| \nabla_{xy} \mathbf{O}_{dn} \right\|_1 \right), \quad (4-19)$$

where \mathbf{O} is the image calculated by Eq. (4-8), and \mathbf{O}_{dn} is the denoised image. The TV regularization is not recommended as it is calculated from the first-order partial derivatives and might over-sharpen the boundary between different regions while reducing the differences within a single region, which would result in staircase-like and aberrant reconstructed images

In this section, we quantitatively analyze the noise level of MUTE output and propose using Hessian regularization to suppress the noise signals of MUTE. To

quantitatively measure the noise level without a reference image, we use the global noise estimation [27, 55]:

$$\sigma = \sqrt{\frac{\pi}{2}} \frac{1}{6(M-2)(N-2)} \sum_{x,y} |N \otimes \mathbf{O}_y|, \quad N = \begin{bmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{bmatrix}, \quad (4-20)$$

where \mathbf{O}_y denotes the input Y channel. A larger value of σ denotes stronger noise levels.

Tab. 4-4. Noise level for 4 methods on cataract retinal image dehazing

Databases mean ± std.	Noise level	Raw	Methods			
			DVA	DRC	LPAR	MUTE
Cataract (100 images)	σ	0.5127 ± 0.0724	0.8148 ± 0.3496	1.6710 ± 0.2535	8.5224 ± 2.2421	11.7200 ± 3.0493
Proprietary (94 images)	σ	0.6979 ± 0.3680	1.1143 ± 1.0150	2.1283 ± 1.3075	6.0175 ± 3.4333	16.0702 ± 4.7823
MESSIDOR (1200 images)	σ	0.5445 ± 0.1710	1.4447 ± 0.4177	1.8126 ± 0.6255	5.1842 ± 1.8195	8.5259 ± 2.4833

We use Eq. (4-20) to measure the noise level of all four methods (DVA, DRC, LPAR, and MUTE) among cataract retinal images (100 images from the CATARACT dataset and 94 images of our proprietary) and the MESSIDOR retinal images (1200 images), the results are listed in Tab. 4-4.

As shown in Tab. 4-4, the MUTE can also enlarge the noise signals for retinal images corrupted by poor illumination and heavy hazy effects. Therefore a denoising process can be applied to the MUTE output. Here we choose L_1 -norm Hessian regularization for image denoising. The denoising process is only applied to the Y channel since Pb and Pr channels remain unchanged.

As the λ is the key parameter that controls the denoising level and data fidelity, it should be chosen carefully based on the noise level. We consider Eq. (4-20) as a reference value, and test $\lambda = \sigma / 10$, $\lambda = \sigma / 20$, and $\lambda = \sigma / 30$ for the 194 cataract retinal images. Again, we use IE, C_{ramm} , UISM, UIConM, and FADE to measure the dehazing performance. Results are shown in Tab. 4-5.

As listed in Tab. 4-5, the Hessian regularization significantly suppresses the noise level of MUTE outputs. When $\lambda = \sigma / 30$ the noise level of denoised MUTE is less than the original MUTE by 3-fold, and is less than LAPR (Tab. 4) by 2-fold in the CATARACT dataset. The dehazing performance is also maintained, for example, the FADE score is less than 0.2 for both CATARACT and our proprietary datasets.

Tab. 4-5. MUTE with Hessian regularization denoising on cataractous retinal images.

Databases mean \pm std.	Matrixes	Raw	MUTE, λ value			
			No denoise	$\sigma / 10$	$\sigma / 20$	$\sigma / 30$
Cataract (100 images)	UISM	0.9002 \pm 0.1432	6.7884 \pm 0.7869	5.3386 \pm 0.4284	5.5770 \pm 0.3952	5.7044 \pm 0.4106
	UIConM	0.0573 \pm 0.0123	0.3890 \pm 0.0093	0.3298 \pm 0.0303	0.3419 \pm 0.0212	0.3467 \pm 0.0189
	IE	5.8443 \pm 0.4449	7.4151 \pm 0.1338	7.1076 \pm 0.2255	7.1581 \pm 0.1787	7.1840 \pm 0.1755
	C _{RAMM}	1.2013 \pm 0.1741	8.1483 \pm 0.6035	6.0133 \pm 0.8880	6.3177 \pm 0.8098	6.4882 \pm 0.7938
	FADE	0.4715 \pm 0.1155	0.1469 \pm 0.0228	0.2109 \pm 0.0272	0.1903 \pm 0.0234	0.1790 \pm 0.0215
	σ	0.5127 \pm 0.0724	11.7200 \pm 3.0493	2.4091 \pm 0.1913	2.7044 \pm 0.5011	3.1508 \pm 0.8263
Proprietary (94 images)	UISM	1.2065 \pm 0.3028	6.8936 \pm 1.1603	5.8507 \pm 0.5926	6.1099 \pm 0.6901	6.2199 \pm 0.7329
	UIConM	0.0446 \pm 0.0133	0.3811 \pm 0.0646	0.3482 \pm 0.0492	0.3548 \pm 0.0497	0.3571 \pm 0.0498
	IE	6.8377 \pm 0.5775	7.4889 \pm 0.4049	7.3438 \pm 0.3763	7.3909 \pm 0.3762	7.4144 \pm 0.3760
	C _{RAMM}	1.0378 \pm 0.1741	7.9087 \pm 1.5078	6.1696 \pm 1.2147	6.5879 \pm 1.2283	6.7936 \pm 1.2309
	FADE	1.0455 \pm 0.3905	0.1857 \pm 0.1177	0.2062 \pm 0.0474	0.1865 \pm 0.0473	0.1790 \pm 0.0469
	σ	0.6979 \pm 0.3680	16.0702 \pm 4.7823	3.9869 \pm 1.5204	5.8252 \pm 2.4833	7.2014 \pm 2.9582
Average (194 images)	UISM	1.0486 \pm 0.2798	6.8393 \pm 0.9844	5.5869 \pm 0.5737	5.8352 \pm 0.6172	5.9542 \pm 0.6419
	UIConM	0.0512 \pm 0.0143	0.3852 \pm 0.0455	0.3390 \pm 0.0418	0.3483 \pm 0.0386	0.3519 \pm 0.0379
	IE	6.3164 \pm 0.7181	7.4568 \pm 0.2996	7.2258 \pm 0.3313	7.2745 \pm 0.3161	7.2992 \pm 0.3146
	C _{RAMM}	1.1221 \pm 0.1920	8.0322 \pm 1.1388	6.0915 \pm 1.0642	6.4528 \pm 1.0465	6.6409 \pm 1.0443
	FADE	0.7496 \pm 0.4038	0.1657 \pm 0.0856	0.2086 \pm 0.0386	0.1884 \pm 0.0373	0.1790 \pm 0.0364
	σ	0.5986 \pm 0.2733	12.7902 \pm 3.8432	3.1980 \pm 1.3393	4.2648 \pm 2.3749	5.1761 \pm 2.9691

Tab. 4-6. MUTE with Hessian regularization denoising $\lambda = \sigma / 30$ on other datasets.

Databases mean \pm std.	Matrixes	Raw	Methods				
			DVA	DRC	LPAR	MUTE	MUTE $\sigma / 30$
DiaRetDB0+DB1 (219 images)	UISM	3.2857 \pm 0.2107	3.5914 \pm 0.5870	4.2797 \pm 0.3261	5.3378 \pm 0.8601	6.7373 \pm 0.3610	6.0116 \pm 0.2993
	UIConM	0.0396 \pm 0.0140	0.1793 \pm 0.0393	0.0989 \pm 0.0389	0.2419 \pm 0.0471	0.3977 \pm 0.0068	0.3862 \pm 0.0092
	IE	5.6481 \pm 0.3389	6.2941 \pm 0.2314	5.7972 \pm 0.3434	5.8912 \pm 1.8016	6.6910 \pm 0.1443	6.6076 \pm 0.1612
	C _{RAMM}	0.7169 \pm 0.1332	2.4855 \pm 0.7205	1.4165 \pm 0.4100	3.7898 \pm 1.2053	8.0577 \pm 0.5667	6.6918 \pm 0.4953
	FADE	0.4752 \pm 0.0395	0.4188 \pm 0.0648	0.3094 \pm 0.0461	0.2495 \pm 0.1344	0.1278 \pm 0.0156	0.1468 \pm 0.0122
	σ	0.6522 \pm 0.0507	0.9402 \pm 0.2096	1.4415 \pm 0.2781	4.6964 \pm 1.5661	12.3944 \pm 1.6480	6.6762 \pm 1.2523
STARE (397 images)	UISM	2.6542 \pm 0.6460	4.1758 \pm 1.0861	4.6212 \pm 0.5306	5.5971 \pm 1.0953	6.5124 \pm 0.8229	6.0333 \pm 0.7143
	UIConM	0.0879 \pm 0.0195	0.2440 \pm 0.0458	0.1930 \pm 0.0444	0.2863 \pm 0.0499	0.3693 \pm 0.0180	0.3619 \pm 0.0189
	IE	5.4185 \pm 0.3468	6.3837 \pm 0.2481	5.5470 \pm 0.3198	5.5085 \pm 1.2873	5.8185 \pm 0.2760	5.7667 \pm 0.2869
	C _{RAMM}	1.9010 \pm 0.2998	5.1454 \pm 1.4516	3.3505 \pm 0.7190	5.4806 \pm 1.2965	7.8314 \pm 0.7522	7.2886 \pm 0.7203
	FADE	0.4744 \pm 0.1274	0.3947 \pm 0.1296	0.2367 \pm 0.0617	0.2739 \pm 0.1453	0.1331 \pm 0.0236	0.1363 \pm 0.0159
	σ	0.8479 \pm 0.1725	1.7854 \pm 0.6727	2.5389 \pm 0.6197	5.1547 \pm 2.0209	8.5333 \pm 2.3205	4.1078 \pm 0.7908
MESSIDOR (1200 images)	UISM	1.8419 \pm 0.3224	4.0235 \pm 0.7075	5.0035 \pm 0.4089	6.2763 \pm 0.6584	7.0831 \pm 0.4933	6.3791 \pm 0.4139
	UIConM	0.0746 \pm 0.0185	0.2383 \pm 0.0318	0.1608 \pm 0.0418	0.2777 \pm 0.0280	0.3660 \pm 0.0073	0.3608 \pm 0.0083
	IE	5.1908 \pm 0.3454	6.2726 \pm 0.1881	5.6377 \pm 0.3846	5.9916 \pm 0.7165	6.0725 \pm 0.2111	6.0158 \pm 0.2074
	C _{RAMM}	1.1810 \pm 0.1639	3.9134 \pm 0.7737	2.2784 \pm 0.4610	5.7179 \pm 0.7411	7.6652 \pm 0.5981	6.9967 \pm 0.5910
	FADE	0.4652 \pm 0.0673	0.4361 \pm 0.0842	0.2320 \pm 0.0410	0.1812 \pm 0.0726	0.1252 \pm 0.0136	0.1303 \pm 0.0098
	σ	0.5445 \pm 0.1710	1.4447 \pm 0.4177	1.8126 \pm 0.6255	5.1842 \pm 1.8195	8.5259 \pm 2.4833	4.5249 \pm 1.2562
NORMAL (300 images)	UISM	1.5512 \pm 0.2604	3.4903 \pm 0.7928	4.7259 \pm 0.3344	5.7581 \pm 0.6043	6.9285 \pm 0.4592	6.3241 \pm 0.4392
	UIConM	0.0768 \pm 0.0152	0.2389 \pm 0.0342	0.1700 \pm 0.0336	0.3092 \pm 0.0267	0.3733 \pm 0.0047	0.3697 \pm 0.0065
	IE	5.5313 \pm 0.2497	6.4120 \pm 0.1878	5.8949 \pm 0.2886	6.3338 \pm 0.7880	6.3062 \pm 0.1898	6.2581 \pm 0.1809
	C _{RAMM}	1.2468 \pm 0.1495	3.9861 \pm 0.8219	2.5810 \pm 0.4295	6.3165 \pm 0.8719	8.3668 \pm 0.4882	7.6748 \pm 0.4834
	FADE	0.6671 \pm 0.1226	0.5629 \pm 0.1132	0.3143 \pm 0.0622	0.2096 \pm 0.0876	0.1497 \pm 0.0186	0.1551 \pm 0.0131
	σ	0.4923 \pm 0.0778	1.1698 \pm 0.3010	1.7765 \pm 0.2532	5.4016 \pm 0.9604	8.8852 \pm 1.5514	3.9687 \pm 0.9355

We compared the dehazing results for MUTE without and with the denoising process against other methods on 2116 non-cataract retinal images. $\lambda = \sigma / 30$ for the Hessian regularization. As listed in Tab. 4-6, the Hessian denoising decreases the noise level of MUTE. The noise level is smaller than other methods while the dehazing qualities including UIISM, UIConM, C_{RAMM} , and FADE scores for Hessian denoised MUTE are maintained as it outperforms other state-of-the-art methods and approaches to un-denoised MUTE.

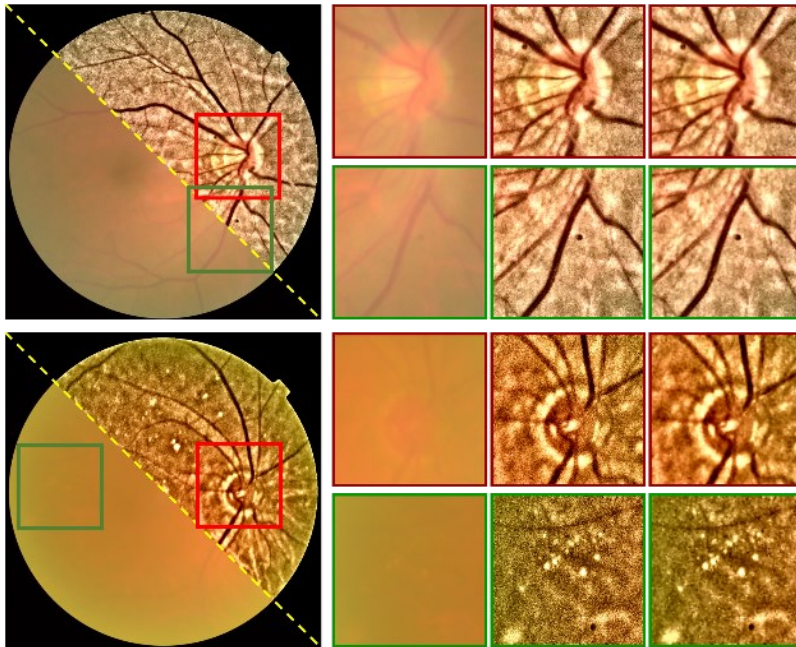


Fig. 4-23. Demonstration of MUTE with and without Hessian denoising process. $\lambda = \sigma / 30$. Images in the first column are montages of raw and dehazed images. The third column and fourth column are original output and denoised output of MUTE for the zoomed-in area of red and green boxes.

Fig. 4-23 shows visual assessments for two cataractous retinal images. As shown in the fourth column of Fig. 4-23, the Hessian regularization suppresses the noise signals, and the dehazing quality is still similar to those without the denoising process shown in the third column of Fig. 4-23.

8. Concluding remarks

8.1 Discussion

Due to the existence of the ocular lens and absorption of biological tissues, retinal images have much higher intensity in red/green channels than that in blue channels. Algorithms for cataractous retinal image dehazing should be designed under consideration of the spectral feature of retinal images.

In this research, we applied the YPbPr color space transformation on the double-pass fundus reflection (DPFR) model and obtain the YPbPr-DPFR model, which shows the haze effect of cataracts only impacts the Y channels of the image. For the YPbPr transformation matrix, it assigned the most weight to the green channel and the least weight to the blue channels, which is just corresponding to the prior feature of retinal images that is the green channel contains important retinal structures and the blue channel has less information due to the absorption.

Then we were inspired by the duality between image dehazing and image denoising and proposed a multilevel-stimulated denoising strategy, termed MUTE, that is dedicated to cataractous retinal image dehazing. Through our experiments, we found that preprocessing is important as it eliminates the boundary effect and normalized the intensity levels for retinal images of different problematic illumination conditions.

In general, MUTE has six hyperparameters that need to be manually determined which are I_0 for dehazing degree, and η , α , β , τ , ξ for Adam-gradient descent. However, according to our massive experiment throughout 2318 images, we have found that the default values ($I_0 = 0.52$, $\eta = 0.01$, $\alpha = \tau = 0.9$, $\beta = \xi = 0.99$) given in the main text are appropriate in order to obtain promising enhancement results. Therefore, there is only one parameter I_0 that is needed to be manually adjusted to optimize the performance for MUTE.

The algorithm for MUTE is composed using MATLAB 2018a and executed using a personal laptop with Intel Core i7-8750H CPU and 16 RAM. No GPU acceleration is used. Average executing times for images of different sizes are listed in Tab. 4-7. As the image

size increases, MUTE takes a longer time to proceed with each image. For image size of 2048 by 2048, it takes about 98.44s on average for each image. To speed up the MUTE, we propose a down-up-sampling strategy in **Supplementary Note 4-10**, which significantly decreases the executing times of MUTE while maintaining the enhancement quality.

Tab. 4-7. Average executing time for DVA, DRC, LPAR and MUTE.

Image Size	DVA (s)	DRC (s)	LPAR (s)	MUTE (s), $\mathcal{L}=5$
512 ²	7.21	1.71	2.73	6.23
1024 ²	16.22	3.11	8.22	23.05
2048 ²	48.13	15.76	38.96	98.44

8.2 Contribution of MUTE

In this research, we propose the multilevel stimulated denoising strategy (MUTE) to achieve retinal image dehazing through the process of image denoising. The MUTE has the following novelty and contributions.

First, we adapted our recently developed image formation model, the double-pass fundus reflection model, to formulate the optical process of retinal image formation. This model is physically self-consistent and can better model the retinal image formation than traditional models. We further revealed the mathematical insight that the haze effect in RGB color space only appears in its Y channel in the YPbPr color space. This allows us to separate the haze component from RGB-colored images and thus better handle the dehazing problem for retinal images since they have different statistical features from natural scenes/objects.

Secondly, we mathematically unify the image dehazing task into the framework of image denoising which was historically thought to be unrelated (Supplementary Note 3). We discover that an accurate estimation of the haze transmission matrix plays a crucial role in successful image dehazing, and this estimation can be achieved through

proper image denoising techniques. This finding has the potential to advance the combination of research in image denoising and dehazing, fostering new insights and techniques in both traditional and deep learning-based approaches. We employ an intensity-based cost function that exploits the observation that the dehazing process consistently decreases the image's intensity. This idea facilitates the development of adaptive and unsupervised tuning of dehazing parameters.

4 Thirdly, MUTE outperforms SOTA traditional and learning-based methods, particularly in increasing the contrast of cataractous retinal images as presented in Section 5. The simplicity of the MUTE model structure allows for full monitoring and understanding of every intermediate step by human specialists such as mathematicians and ophthalmologists. This interpretability aspect ensures reliable and safe clinical trials and future applications of MUTE.

8.3 Conclusion

MUTE significantly improves the visual quality and the contrast of cataractous retinal images and reveals retinal structures that might not be detected by clinical doctors due to insufficient contrast of the raw images. It enables and facilitates the detection of features like drusen, hard exudates, bleeding hemorrhages, or microaneurysms, as these tasks may use more color or local detail information. Early detection of these findings through cataractous retinal images may improve the outcome of treatment of retinal diseases, and influence the decision-making on surgery that might be actually not necessary. For example, patients with severe cataracts might also have macular degeneration. In these cases, if the macular degeneration can be observed through enhanced cataractous retinal image, then ophthalmologists are able to choose a better treatment plan to minimize risks and stress for patients, for example, the patients can be counseled on the guarded visual prognosis and a less expensive lens can be used.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding. This research is supported by China Scholarship Council (CSC) (201908340078)

Reference

1. J. Zhang, B. Dashtbozorg, F. Huang, T. T. J. M. Berendschot, and B. M. ter Haar Romeny, "Analysis of Retinal Vascular Biomarkers for Early Detection of Diabetes," in *VipIMAGE 2017*, J. M. R. S. Tavares, and R. M. Natal Jorge, eds. (Springer International Publishing, 2018), pp. 811-817.
2. M. Siegrist, H. Hanssen, M. Neidig, M. Fuchs, F. Lechner, M. Stetten, K. Blume, C. Lammel, B. Haller, M. Vogeser, K. G. Parhofer, and M. Halle, "Association of leptin and insulin with childhood obesity and retinal vessel diameters," *Int. J. Obes.* **38**, 1241-1247 (2014).
3. F. N. Doubal, T. J. MacGillivray, N. Patton, B. Dhillon, M. S. Dennis, and J. M. Wardlaw, "Fractal analysis of retinal vessels suggests that a distinct vasculopathy causes lacunar stroke," *Neurology* **74**, 1102-1107 (2010).
4. M. Ao, X. Li, W. Qiu, Z. Hou, J. Su, and W. Wang, "The impact of age-related cataracts on colour perception, postoperative recovery and related spectra derived from test of hue perception," *BMC Ophthalmology* **19**, 1-9 (2019).
5. P. Zhu, F. Huang, F. Lin, Q. Li, Y. Yuan, Z. Gao, and F. Chen, "The relationship of retinal vessel diameters and fractal dimensions with blood pressure and cardiovascular risk factors," *PloS one* **9**, e106551 (2014).
6. M. C. V. S. Mary, E. B. Rajasingh, and G. R. Naik, "Retinal Fundus Image Analysis for Diagnosis of Glaucoma: A Comprehensive Survey," *IEEE Access* **4**, 4327-4354 (2016).
7. S. Zafar, J. McCormick, L. Giancardo, S. Saidha, A. Abraham, and R. Channa, "Retinal Imaging for Neurological Diseases: "A Window into the Brain", " **59**, 137-154 (2019).
8. E. Peli, and T. Peli, "Restoration of retinal images obtained through cataracts," *IEEE Transactions on Medical Imaging* **8**, 401-406 (1989).
9. L. Xiong, H. Li, and L. Xu, "An enhancement method for color retinal images based on image formation model," *Computer Methods and Programs in Biomedicine* **143**, 137-150 (2017).
10. A. Mitra, S. Roy, S. Roy, and S. K. Setua, "Enhancement and restoration of non-uniform illuminated Fundus Image of Retina obtained through thin layer of cataract," *Computer Methods and Programs in Biomedicine* **156**, 169-178 (2018).
11. L. Cao, and H. Li, "Enhancement of blurry retinal image based on non-uniform contrast stretching and intensity transfer," *Medical & Biological Engineering & Computing* **59** (2020).
12. Y. Luo, K. Chen, L. Liu, J. Liu, J. Mao, G. Ke, and M. Sun, "Dehaze of Cataractous Retinal Images using an Unpaired Generative Adversarial Network," *IEEE Journal of Biomedical and Health Informatics* (2020).
13. L. Cao, and H. Li, "Detail-richest-channel based enhancement for retinal image and beyond," *Biomedical Signal Processing and Control* **69**, 102933 (2021).
14. A. Gaudio, A. Smailagic, and A. e. Campilho, "Enhancement of Retinal Fundus Images via Pixel Color Amplification," *International Conference on Image Analysis and Recognition* (2020).
15. T. T. J. M. Berendschot, P. J. DeLint, and D. v. Norren, "Fundus reflectance—historical and present ideas," *Progress in retinal and eye research* **22**, 171-200 (2003).
16. S. Zhang, C. A. B. Webers, and T. T. J. M. Berendschot, "A double-pass fundus reflection model for efficient single retinal image enhancement," *Signal Processing* **192**, 108400 (2022).
17. P. Artal, I. Iglesias, and N. L'opez-Gil, "Double-pass measurements of the retinal-image quality with unequal entrance and exit pupil sizes and the reversibility of the eye's optical system," *J. Opt. Soc. Am. A* **12**, 2358-2366 (1995).
18. P. Artal, S. Marcos, R. Navarro, and D. R. Williams, "Odd aberrations and double-pass measurements of retinal image quality," *J. Opt. Soc. Am. A* **12**, 195-201 (1995).
19. H. Koschmieder, "Luftlicht und sichtweite," *Naturwissenschaften*, **26**, 521-528 (1938).

20. E. J. McCartney, "Optics of the Atmosphere: Scattering by molecules and particles," New York, 408 (1976).
21. D. Singh, and V. Kumar, "A Comprehensive Review of Computational Dehazing Techniques," Archives of Computational Methods in Engineering (2018).
22. K. P. Senthilkumar, and P. Sivakumar, "A Review on Haze Removal Techniques," Computer Aided Intervention and Diagnostics in Clinical and Medical Images, 113-123 (2019).
23. D. Mujbaile, and D. Rojtkar, "Model based Dehazing Algorithms for Hazy Image Restoration – A Review," (2020).
24. G. Saxena, and S. S. Bhadauria, "An efficient single image haze removal algorithm for computer vision applications," Multimedia Tools and Applications **79**, 28239–28263 (2020).
25. P. Dai, H. Sheng, J. Zhang, L. Li, J. Wu, and M. Fan, "Retinal Fundus Image Enhancement Using the Normalized Convolution and Noise Removing," International Journal of Biomedical Imaging **2016** 1-12 (2016).
26. L. Cao, H. Li, and Y. Zhang, "Retinal image enhancement using low-pass filtering and α -rooting," Signal Processing **170**, 107445 (2020).
27. J. Wang, Y.-J. Li, and K.-F. Yang, "Retinal fundus image enhancement with image decomposition and visual adaptation," Computers in Biology and Medicine **128**, 104116 (2021).
28. M. Zhou, K. Jin, S. Wang, J. Ye, and D. Qian, "Color Retinal Image Enhancement Based on Luminosity and Contrast Adjustment," IEEE Transactions on Biomedical Engineering (2018).
29. B. Gupta, and M. Tiwari, "Color retinal image enhancement using luminosity and quantile based contrast enhancement," Multidimensional Systems and Signal Processing **30**, 1829-1837 (2019).
30. B. Savelli, A. Bria, A. Galdran, C. Marrocco, M. Molinara, A. e. Campilho, and F. Tortorella, "Illumination correction by dehazing for retinal vessel segmentation," IEEE 30th International Symposium on Computer-Based Medical Systems (2017).
31. K. He, J. Sun, and X. Tang, "Single Image Haze Removal Using Dark Channel Prior," (2009).
32. Q. Zhu, J. Mai, and L. Shao, "A Fast Single Image Haze Removal Algorithm Using Color Attenuation Prior," IEEE Transactions on Image Processing **24**, 3522-3533 (2015).
33. M. Ju, C. Ding, Y. J. Guo, and D. Zhang, "IDGCP: Image Dehazing Based on Gamma Correction Prior," IEEE Trans. Image Process. **29**, 3104 - 3118 (2021).
34. A. Arias, and P. Artal, "Wavefront-shaping-based correction of optically simulated cataracts," Optica **7**, 22-27 (2020).
35. A. Galdran, A. Alvarez-Gila, A. Bria, J. Vazquez-Corral, and M. Bertalm'io, "On the Duality Between Retinex and Image Dehazing," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 8212-8221 (2018).
36. D. P. Kingma, and L. J. Ba, "Adam: A Method for Stochastic Optimization," presented at the International Conference on Learning Representations 2015.
37. K. Panetta, C. Gao, and S. Agaian, "Human-Visual-System-Inspired Underwater Image Quality Measures," IEEE Journal of Oceanic Engineering **41**, 1-11 (2016).
38. J. Wang, K. Lu, J. Xue, N. He, and L. Shao, "Single Image Dehazing Based on the Physical Model and MSRCR Algorithm," IEEE Transactions on Circuits and Systems for Video Technology **29**, 2190-2199 (2018).
39. S. Hong, M. Kim, and M. Kang, "Single image dehazing via atmospheric scattering model-based image fusion," Signal Processing **178**, 107798 (2021).
40. A. Rizzi, T. Algeri, G. Medeghini, and D. Marini, "A proposal for Contrast Measure in Digital Images," presented at the Conference on colour in graphics, imaging, and vision 2004.
41. L. K. Choi, J. You, and A. C. Bovik, "Referenceless prediction of perceptual fog density and perceptual image defogging," IEEE Trans. Image Process. **24**, 3888-3901 (2015).

42. L. K. Choi, J. You, and A. C. Bovik, "LIVE Image Defogging Database," https://live.ece.utexas.edu/research/fog/fade_defade.html.
43. "DIARETDB0 - Standard Diabetic Retinopathy Database," <https://www.it.lut.fi/project/imageret/diaretdb0/>.
44. "DIARETDB1 - Standard Diabetic Retinopathy Database Calibration level 1," <http://www2.it.lut.fi/project/imageret/diaretdb1/>.
45. E. Decencière, X. Zhang, G. Cazuguel, B. Lay, B. Cochener, C. Trone, P. Gain, R. Ordonez, P. Massin, A. Erginay, B. Charton, and J.-C. Klein, "Feedback on a publicly distributed database: the Messidor database," 2014 **33**, 4 %J Image Analysis & Stereology (2014).
46. "cataract dataset," <https://www.kaggle.com/jr2ngb/cataractdataset>.
47. C. Y. Cheung, S. Ong, M. K. Ikram, Y. T. Ong, C. P. Chen, N. Venketasubramanian, and T. Y. Wong, "Retinal vascular fractal dimension is associated with cognitive dysfunction," J Stroke Cerebrovasc Dis **23**, 43-50 (2014).
48. F. Huang, B. Dashtbozorg, J. Zhang, E. Bekkers, S. Abbasi-Sureshjani, T. T. Berendschot, and B. M. Ter Haar Romeny, "Reliability of Using Retinal Vascular Fractal Dimension as a Biomarker in the Diabetic Retinopathy Detection," J. Ophthalmol. **2016**, 6259047 (2016).
49. M. Ciancaglini, G. Guerra, L. Agnifili, R. Mastropasqua, V. Fasanella, M. Cinelli, C. Costagliola, and L. Ambrosone, "Fractal dimension as a new tool to analyze optic nerve head vasculature in primary open angle glaucoma," In Vivo **29**, 273-279 (2015).
50. A. F. Frangi, W. J. Niessen, K. L. Vincken, and M. A. Viergever, "Multiscale vessel enhancement filtering," in *International conference on medical image computing and computer-assisted intervention*(Springer1998), pp. 130-137.
51. T. Chakraborti, D. K. Jha, A. S. Chowdhury, X. J. M. V. Jiang, and Applications, "A self-adaptive matched filter for retinal blood vessel detection," **26**, 55-68 (2015).
52. J. Wang, J. Chen, H. Xu, S. Zhang, X. Mei, J. Huang, and J. J. S. P. Ma, "Gaussian field estimator with manifold regularization for retinal image registration," Signal Processing **157**, 225-235 (2019).
53. L. I. Rudin, S. Osher, and E. J. P. D. n. p. Fatemi, "Nonlinear total variation based noise removal algorithms," Physica D **60**, 259-268 (1992).
54. L. Xu, C. Lu, Y. Xu, and J. Jia, "Image smoothing via L 0 gradient minimization," in *Proceedings of the 2011 SIGGRAPH Asia conference*(2011), pp. 1-12.
55. J. Immerkær, "Fast Noise Variance Estimation," Comput. Vis. Image Understand **64**, 300-302 (1996).

Supplementary Notes for Chapter 4

MUTE: a multilevel-stimulated denoising strategy for single cataractous retinal image dehazing

Supplementary Note 4-1: Failure of the natural scene prior knowledge on retinal image dehazing.

4-1.1 Color retinal image spectral property and failure of dark channel prior

The dark channel prior (DCP) [1] and its improved version have been widely used for natural scene dehazing including underwater image enhancement and haze removal even for thick fog situations. The results of DCP dehazing are promising. Applying the DCP to retinal images for dehazing seems to be natural, since (1) the visual effect of haze effect in natural scenes and retinal images are similar to each other. (2) the physical model for haze-scattering and intraocular scattering are consistent with each other [2-4]. Both can be modeled as random phase perturbations. However, it is found that the performance of DCP as its variations on retinal image dehazing is limited especially for thick cataracts.

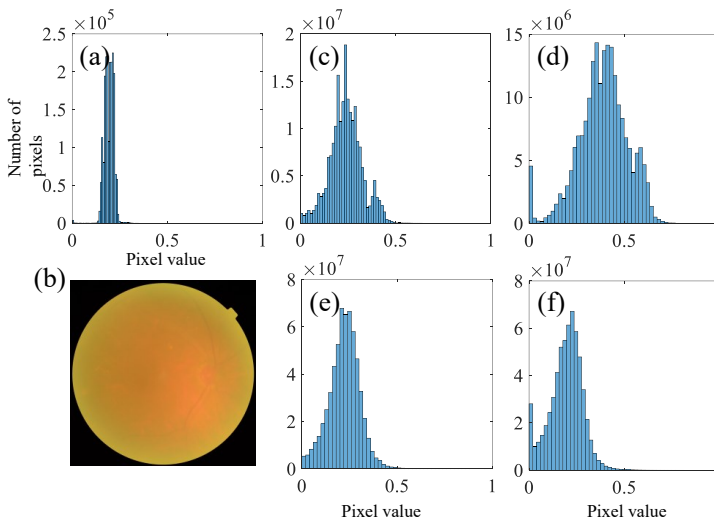


Fig. s4-1. Histograms of (a) blue channel for a retinal image shown in (b); (c) and (d) are blue channel and dark channel for all images in the cataract retinal image database (100 images included); (e) and (f) are the blue and the dark channels for all images in the normal retinal image database (300 images included).

The principle of DCP tells that in any haze-free image (in RGB color space), at least one pixel has zero intensity in at least one channel. This assumption does not hold for normal retinal images and retinal images with severe cataracts, due to the unique spectral property of the ocular lens. In order to protect the retina from harmful blue-wavelength light, the human lens has high absorption for light rays of short wavelengths, which results in low intensities of the blue channel [5-7]. That is why most of the retinal images show reddish or yellowish in their photographs. When the cataract is present, the absorption of short wavelengths will be even higher than the normal condition due to the presence of protein precipitation in the crystalline lens, and lens aging.

The image formation model ignores the condition of light absorption. When DCP is applied to cataractous retinal images (searching for the minimum pixel value in red, green, and blue channels in a given small patch), the DCP will mostly extract the pixel value in the blue channel since it always tends to have small intensity whether there is a cataract or not due to the absorption of short wavelength. As a result, the transmission maps will be over-estimated (under-estimate the haze thickness).

For example, Fig. s4-1(a) shows the histogram of the blue channel of a cataractous retinal image in Fig. s4-1 (b). Most of the pixels in the blue channel have small values that are close to 0.2. Fig. s4-1 (c) shows the histogram of the blue channel for all 100 images in a public cataract retinal image of different degrees from <https://www.kaggle.com/jr2ngb/cataractdataset>. The pixel values in the blue channel range from 0 to 0.5, whose pixel values are also small. By applying DCP to those cataract retinal images, the pixel values in the dark channel are shown in Fig. s4-1 (d) with a range of [0, 0.5]. The distribution of the dark channel [Fig. s4-1 (d)] is similar to that of the blue channel in Fig. s4-1 (c), and most of the pixel values are located at 0.25 to 0.5.

We have also investigated the pixel value distributions for retinal images of normal eyes. A total of 300 images are taken from the public normal retinal images database at <https://www.kaggle.com/jr2ngb/cataractdataset>. Their blue channel and dark channel

distribution are shown in Fig. s4-1 (e) and Fig. s4-1 (f), respectively. Apparently, this distribution does not agree with the DCP assumption for haze-free images, in which the pixel values should approach zero. Meanwhile, there is no distinct difference in the pixel values' distribution between cataractous and normal retinal images. As a result, the DCP and its variations will treat the cataractous retinal images as normal retinal images and over-estimate the transmission map.

4-1.2 Demonstrations of failure of natural scene dehazing methods on cataractous retinal images.

Apart from DCP, we have also tested other dehazing algorithms including boundary constraint dehazing [8], non-local dehazing [9], and color attenuation prior dehazing [10]. Demonstration results are shown in Fig. s4-2, where the above-mentioned dehazing methods fail to suppress the haze effect of cataractous retinal images.

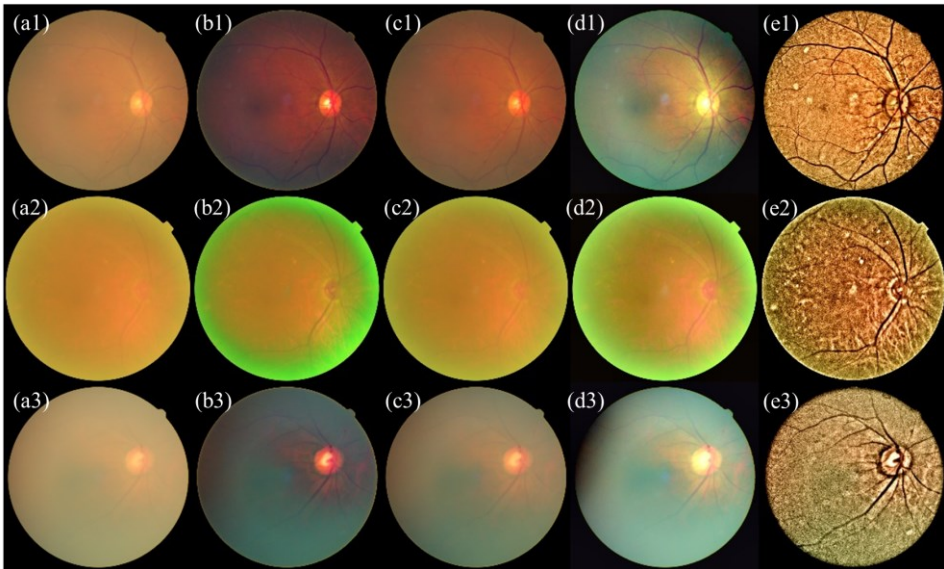


Fig. s4-2. Cataractous retinal image dehazing using different methods. (a1) to (a3) are raw cataractous retinal images. (b1) to (b3) are dehazed results for boundary constraint dehazing. (c1) to (c3) are dehazed results for nonlocal dehazing. (d1) to (d3) are results for color attenuation prior dehazing. (e1) to (e3) are results of MUTE.

The boundary constraint dehazing can be categorized as a variation of DCP, therefore, it fails to eliminate the haze effect for cataractous retinal images. The non-local dehazing utilizes the haze-line to estimate the transmission matrix in Eq. (4-1), however, it also fails to estimate the haze density as the small value of blue channel in retinal images leads to under overestimating the transmission matrix. The color attenuation prior dehazing is based on the observation that haze changes the brightness and saturation of the natural scene images. This observation does not hold for retinal images due to its spectral property, and the model parameters are trained based on natural images rather than retinal images.

Supplementary Note 4-2: Adaptive illumination correction for retinal images

4-2.1 Image background filling

A retinal image is composed of two major regions - a circular retina and a black background. Since the image will be filtered using a low-pass filter in the following section, unexpected results often occur when the convolution kernel slides through the retinal due to the sudden change of pixel values. Padding the black areas before an enhancement is an effective way to reduce an over enhancement of the retinal boundary. We fill the black background in a reflective manner [11].

4-2.2 Adaptive illumination correction for Y channel.

The double-pass fundus reflection model for Y channel is given by

$$\mathbf{S}_Y = \mathbf{L} \cdot \mathbf{T}_{lens}^2 \cdot \mathbf{S}_Y^{IC} . \tag{s4-1}$$

We rewrite the illumination term, $\mathbf{L} \cdot \mathbf{T}_{lens}^2$ as $\mathbf{L} \cdot \mathbf{T}_{lens}^2 = \mathbf{L}_{coarse} \cdot \mathbf{L}_{fine}$, where \mathbf{L}_{coarse} and \mathbf{L}_{fine} denote the coarse illumination pattern (slowly varying in the spatial domain) and fine illumination pattern (fast varying in the spatial domain) projected on the retina, respectively.

$$\mathbf{S}_Y = \mathbf{L}_{coarse} \cdot \mathbf{S}_0, \quad \mathbf{S}_0 = \mathbf{L}_{fine} \cdot \mathbf{S}_Y^{IC}. \quad (s4-2)$$

Practically, the coarse illumination pattern denotes the overall illumination provided by the light source, while the fine illumination pattern can be due to the intensity fluctuations during the camera capturing, and some noise.

Since retinal images might have good illumination conditions or suffer from different levels of uneven illumination. We first need to correct the coarse illumination pattern, making the retinal images have even illumination conditions.

Here we take advantage of the Retinex theory on illumination correction. Taking logarithm to both sides of Eq. (s4-2), yielding

$$\log(\mathbf{S}_Y) = \log(\mathbf{L}_{coarse}) + \log(\mathbf{S}_0), \quad (s4-3)$$

Since \mathbf{L}_{coarse} is slowly varying in the spatial domain, a good estimation of \mathbf{L}_{coarse} can be generated by filtering \mathbf{S} with a low-pass filter which is

$$\mathbf{L}_{coarse} = \mathbf{S}_Y \otimes \left[\frac{1}{2\pi\omega^2} \exp\left(-\frac{\mathbf{r}^2}{2\omega^2}\right) \right]. \quad (s4-4)$$

Here, \otimes denotes convolution. ω controls the full width at half maximum (FWHM) of the Gaussian kernel. The size of ω is determined according to the size of the input image, where $\omega = \lfloor M/20 \rfloor$, M is the width (pixels) of the input image, and $\lfloor \cdot \rfloor$ denotes round-toward-zero. Substituting Eq. (s4-4) to Eq. (s4-3), we have

$$\mathbf{S}_0(\mathbf{r}) = \exp\left(\log(\mathbf{S}_Y + \varepsilon) - \log\left\{\mathbf{S}_Y \otimes \left[\frac{1}{2\pi\omega^2} \exp\left(-\frac{\mathbf{r}^2}{2\omega^2}\right)\right] + \varepsilon\right\}\right) - \varepsilon. \quad (s4-5)$$

Here ε is a small value to avoid the situation of $\log(0)$, $\varepsilon = 0.001$ in the following studies.

The formular for \mathbf{S}_0 is rewritten as

$$\mathbf{I} - \mathbf{S}_0(\mathbf{r}) = \mathbf{L}_{fine}(\mathbf{r}) \cdot [\mathbf{I} - \mathbf{S}_Y^{IC}(\mathbf{r})] + \mathbf{I} - \mathbf{L}_{fine}(\mathbf{r}), \quad (s4-6)$$

We therefore can apply the imaging de-hazing to Eq. (s4-6) and obtain the illumination corrected reflectance object \mathbf{S}_Y^{IC} . Here we use the gray-scale dark-channel prior refined by the guided filter for image illumination correction [12]. The illumination correction for gray-scale images using dark channel prior does not rely on the spectral properties, therefore it can correctly estimate the illumination pattern and correct the uneven illuminations.

The estimation of \mathbf{S}_Y^{IC} is given by

$$\mathbf{S}_Y^{IC}(\mathbf{r}) = \mathbf{I} - \frac{[\mathbf{I} - \mathbf{S}_0(\mathbf{r})] + \mathbf{L}_{fine}(\mathbf{r}) - \mathbf{I}}{\mathbf{L}_{fine}(\mathbf{r})}, \quad (\text{s4-7})$$

with

$$\mathbf{L}_{fine} \approx \mathbf{I} - \alpha \cdot (\mathbf{I} - \mathbf{S}_0)_{dark}. \quad (\text{s4-8})$$

$(\mathbf{I} - \mathbf{S}_0)_{dark}$ is the dark channel of $\mathbf{I} - \mathbf{S}_0$. The dark channel of a grey-scale image \mathbf{X} estimated in a local neighborhood $\Omega(\mathbf{r})$ with the size of w pixels is equal to filtering the image using the local minimum filter [1]:

$$\mathbf{X}_{dark} = \min_{\rho \in \Omega(\mathbf{r})} \mathbf{X}(\rho). \quad (\text{s4-9})$$

w is the patch size, which is given by $w = \lfloor M / 150 \rfloor$, according to the size of the input image. The parameter $\alpha \in [0, 1]$ is used to control the degree of illumination boosting. A large value of α will lighten the gray-scale image, and may also cause an overexposure effect where the pixel value is larger than 1. An appropriate value of α is important for ensuring good enhancement results, we, therefore, use an adaptive manner to determine the value of α .

To determine the value of α with low complexity but high accuracy, a global wise optimization function is designed as

$$\alpha = \arg \min \left(\left\{ \frac{\sum_{r \in \Gamma} [\mathbf{S}_Y^{IC}(\mathbf{r})]}{\# \text{pixels}(\Gamma)} - 0.02 \right\}^2 \right), \quad (\text{s4-10})$$

where the function $f[\mathbf{S}_Y^{IC}(\mathbf{r})] = 1$ for $\mathbf{S}_Y^{IC}(\mathbf{r}) > 1$, $f[\mathbf{S}_Y^{IC}(\mathbf{r})] = 0$ for $\mathbf{S}_Y^{IC}(\mathbf{r}) \leq 1$. Γ is the retinal area apart from the black background. $\# \text{pixels}(\Gamma)$ is the number of the total pixels in Γ . $\sum_r f[\mathbf{S}_Y^{IC}(\mathbf{r})] / \# \text{pixels}(\Gamma)$ implies the percentage of overexposed pixels. Equation (s4-10) ensures that the percentage of pixels that are overexposed during the illumination boosting should approach to 2%.

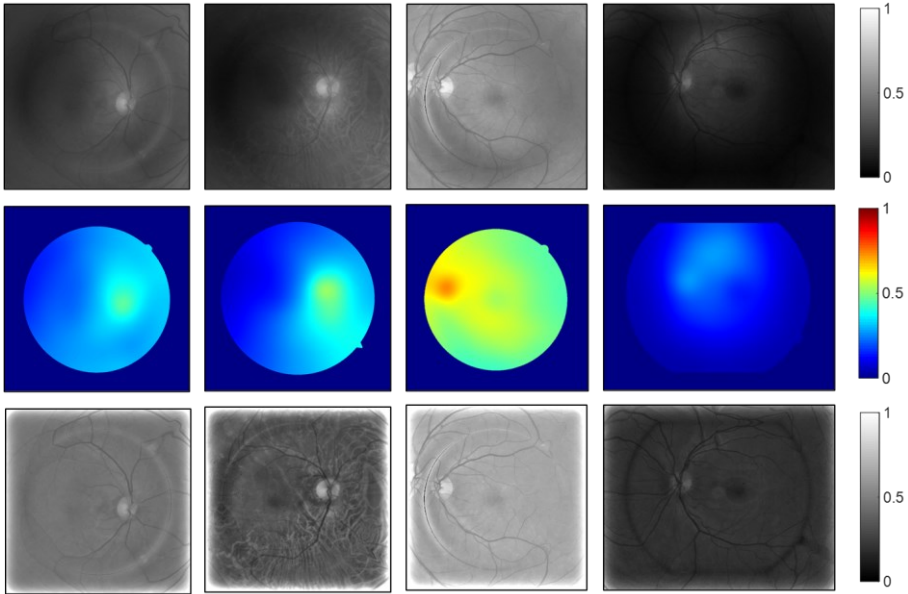


Fig. s4-3. Illumination coarse correction in the Y channel. First row: raw Y channel images. Second row: coarse illumination estimation. Third row: Y channel after coarse illumination correction.

Considering that Eq. (s4-10) is a one-dimensional optimization function. In this work, the Fibonacci method (FM) is adopted to solve Eq. (s4-10) since it is able to gradually narrow the search interval for a one-dimensional optimization problem until

the convergence condition is satisfied [13]. Fig. s4-3 and Fig. s4-4 demonstrate the illumination corrections for retinal images under different illumination conditions. The coarse illumination correction can correct the uneven illumination pattern, while the fine illumination correction increases the intensity levels.

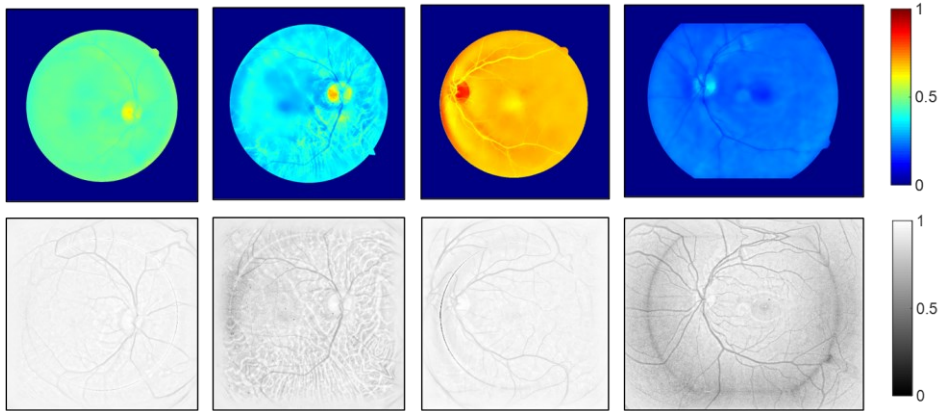


Fig. s4-4. Illumination fine correction in the Y channel. First row: Fine illumination estimation based on the coarse-correction results. Second row: Final outputs of the illumination correction process.

Supplementary Note 4-3: On the relationship of image dehazing, Retinex theory, and image denoising

4-3.1 Image denoising as the retinex solvers

Retinex was introduced by Edwin H. Land, and was originally defined as a color vision model of human perception which aims to explain the human ability to perceive color as stable regardless of changes in global illumination. Today, the Retinex theory has been applied to different image processing tasks including color constancy, shadow removal, illumination correction, and contrast enhancement. Especially, the Retinex theory was widely used in illumination correction tasks where the image formation model is given by the product of illumination component B , and reflectance R :

$$I(x, y) = B(x, y)R(x, y), \quad (\text{s4-11})$$

Taking logarithm to both sides of Eq. (s4-11) yielding

$$i(x, y) = b(x, y) + r(x, y), \quad (\text{s4-12})$$

where $i = \log(I)$. **The core idea of Retinex is that illumination b is supposed to be varying smoothly, and can be estimated by image smoothing according to practical algorithms.**

Based on the idea, the Retinex has various models and can be categorized into the threshold-based Retinex, reset-based Retinex, center-surround Retinex, and variational Retinex. Later, as pointed out in Ref. [14], all of the Retinex models can be unified in the framework of variational retinex with different local or non-local filters. For example, the center-surround retinex with a Gaussian filter can be regarded as a non-local filter by assigning Gaussian distributed weights to adjacent pixels. The unified Retinex model is written as

$$r = \arg \min \|\nabla_w r - \nabla_w i\|_2^2 + \lambda \varphi(r), \quad (\text{s4-13})$$

where ∇_w is the local/non-local gradient, and $\varphi(r)$ is the regularization term which can be total variation, non-local total variation, gradient L_0 -norm, or the gradient L_2 -norm [14].

As the illumination b is supposed to be varying smoothly, any image smoothing methods can be used for estimating b and thus benefit retinex in different degrees. It is also true that the (non-local) total variation and gradient L_0 -norm used in the unified retinex are originally designed for image denoising or image smoothing. The Gaussian filter is often used for center-surround retinex with single and multi-scales. If we consider image smoothing being an extreme case of image denoising, we can write the following notation:

$$\begin{aligned} b &= \text{Denoise}(i) \\ r &= \text{Retinex}_{linear}(i) = i - b = i - \text{Denoise}(i), \end{aligned} \quad (\text{s4-14})$$

The exponential form is

$$R = \text{Retinex}_{\log}(I) = \exp\left[\text{Retinex}_{\text{linear}}(\log(I))\right], \quad (\text{s4-15})$$

where the image denoising is a solver for retinex.

4-3.2 The duality between nature scene image dehazing and Retinex theory

The nature scene image dehazing seems to be unrelated to Retinex theory since they are dealing with different problems. Later, as pointed out in Ref. [15] they are connected by a simple algebra modification of the haze formation model by assuming the input image is globally white-balanced that is

$$S(x, y) = T(x, y)O(x, y) + (1 - T), \quad (\text{s4-16})$$

$$\begin{aligned} \Leftrightarrow 1 - S &= 1 - TO - 1 + T \\ \Leftrightarrow (1 - S) &= T(1 - O) \end{aligned} \quad (\text{s4-17})$$

where S is the hazed image, O is the haze-free scene, and T is the transmission matrix of haze. If we consider $(1 - S) = I$ and $(1 - O) = R$, Eq. (s4-17) is identical to Eq. (s4-11). By assuming T is spatially slow varying, according to Eq. (s4-17) and Eq. (s4-15), we have

$$O = 1 - \text{Retinex}_{\log}(1 - S) = \text{Dehazing}(S). \quad (\text{s4-18})$$

This formula is also proofed in Ref [15], and shows that the image dehazing task can be finished under the retinex theory.

4-3.3 Dehazing of cataractous retinal images can be regarded as a special image denoising task

Our image formation model for the cataractous retinal image is given by Eq. (4-7) in the main text, and we also assume that T_{sc} is spatially slow varying. However, Eq. (4-7) can not be reformulated as a multiplication form since the orders of T_{sc} in Eq. (4-7) are different. Considering that T_{sc} is slowly varying, $1 - T_{sc}$ should also be slowly varying.

We, therefore, find the similarity between Eq. (4-7) and Eq. (4-4) in the main text. Since Eq. (4-4) is identical to Eq. (s4-12), we are able to regard $1 - \mathbf{T}_{sc} = b$, and can be obtained by image denoising according to Eq. (s14):

$$\mathbf{T}_{sc} = 1 - \text{Denoise}(\mathbf{S}) \quad (\text{s4-19})$$

As long as \mathbf{T}_{sc} is determined, the haze-free image can be obtained. Therefore, our task is to find a denoise algorithm to solve Eq. (s4-19).

Current denoising algorithms are not designed for cataractous retinal image dehazing tasks, and the denoised image needs to be further refined to form a good estimation of \mathbf{T}_{sc} . Inspired by multi-scale treatment, we consider the \mathbf{T}_{sc} as a linear combination of multiple denoised images of different denoised strengths. For strong denoised images, they are regarded as coarse estimation of \mathbf{T}_{sc} since they contain large-scale features of \mathbf{T}_{sc} . While weak denoised images can be regarded as fine estimation of \mathbf{T}_{sc} but corrupted by the reflected components. We are further inspired by stimulating functions, and we assign stimulated functions to different layers of denoised image, and design a cost function that can tell the model which pixel should contain more details while which pixel shouldn't. The resultant formula is given by Eq. (4-10) in the main text.

In our model, we have a total of $(\mathcal{L}-1)MN+1$ unknown parameters ($\mathcal{L}-1$ layers with M -by- N pixels, plus the parameter \mathcal{A}) to be determined. If the stimulation function is assigned to S_0 , then we have a total of $\mathcal{L}MN+1$ unknown parameters. It is possible to solve the unknown parameters since we have $(\mathcal{L}+1)MN$ data points (\mathcal{L} denoised image plus one single input image), but we found according to our pioneering experiment trials that the robustness is not very good. One solution is to let the fine (coarse) denoised image (blurred by the smallest Gaussian kernel) as a basement and use other denoised images (blurred by different sizes of Gaussian kernels) to modify it.

In this way, we have enough data redundancy as there are $(\mathcal{L}+1)MN$ data points for solving $(\mathcal{L}-1)MN+1$ unknown parameters. The optimization can be more efficient and stable for different input cataractous retinal images.

Supplementary Note 4-4: Learning for model parameters

In this section, we perform the mathematical calculation for the closed-form gradient of Eq. (4-12) with respect to parameters $\boldsymbol{\varphi}_l(\mathbf{r})$ and \mathcal{A} , respectively.

4-4.1 Learning for $\boldsymbol{\varphi}_l(\mathbf{r})$

Since we assume that there is no correlation between each pixel, when learning $\boldsymbol{\varphi}_l(\mathbf{r})$ Eq. (4-12) can be converted into a pixel-wise optimization problem which is

$$E = (\mathbf{O}_Y - I_0)^2 - \lambda |\nabla \mathbf{O}_Y|^2, \quad (\text{s4-20})$$

The derivatives of Eq. (4-13) with respect to $\boldsymbol{\varphi}_l(\mathbf{r})$ are calculated as

$$\frac{\partial E}{\partial \boldsymbol{\varphi}_l} = \left[2(\mathbf{O}_Y - I_0) - 2\lambda (\nabla^T \nabla \mathbf{O}_Y) \right] \cdot \frac{\partial \mathbf{O}_Y}{\partial \boldsymbol{\varphi}_l}, \quad (\text{s4-21})$$

where

$$\frac{\partial \mathbf{O}_Y}{\partial \boldsymbol{\varphi}_l} = \frac{2\mathcal{A} \cdot \exp(-\boldsymbol{\varphi}_l)}{(\mathcal{L}-1) \cdot [1 + \exp(-\boldsymbol{\varphi}_l)]^2} \cdot \frac{(\mathbf{S}_Y^{IC} - 1)}{\mathbf{T}_{sc}^3} \cdot \mathcal{S}_l(\mathbf{S}_Y^{IC}). \quad (\text{s4-22})$$

We update $\boldsymbol{\varphi}_l$ using the gradient descent method with adaptive moment estimation (Adam) [16], since our problem is non-convex. The initial value is given by $\boldsymbol{\varphi}_{l,0}(\mathbf{r}) = 0$.

At the n -th iteration, $\boldsymbol{\varphi}_l$ is updated according to

$$\begin{cases} \mathbf{g}_l(\mathbf{r}) = \frac{\partial E}{\partial \boldsymbol{\varphi}_{l,n}}(\mathbf{r}) \\ \mathbf{m}_{l,n+1}(\mathbf{r}) = \alpha \cdot \mathbf{m}_{l,n}(\mathbf{r}) + (1-\alpha) \cdot \mathbf{g}_l(\mathbf{r}) \\ \mathbf{v}_{l,n+1}(\mathbf{r}) = \beta \cdot \mathbf{v}_{l,n}(\mathbf{r}) + (1-\beta) \cdot \mathbf{g}_l^2(\mathbf{r}) \\ \boldsymbol{\varphi}_{l,n+1}(\mathbf{r}) = \boldsymbol{\varphi}_{l,n}(\mathbf{r}) - \eta \cdot \frac{\mathbf{m}_{l,n+1}(\mathbf{r})}{\sqrt{\mathbf{v}_{l,n+1}(\mathbf{r}) + \varepsilon}} \end{cases}, \quad \mathbf{r} \in \Gamma, \quad (\text{s4-23})$$

η is the learning rate at n -th iteration, \mathbf{m}_n ($\mathbf{m}_0 = 0$) is the first momentum estimation, and \mathbf{v}_n ($\mathbf{v}_0 = 0$) is the second-moment estimation. α and β are decay rates for the moment estimations. ε is a very small value used to avoid the condition of dividing by 0. We choose $\eta = 0.01$, $\alpha = 0.9$ and $\beta = 0.99$ in all following experiments.

4-4.2 Learning for \mathcal{A}

The derivative of E with respect to \mathcal{A} is given by

$$\frac{\partial E}{\partial \mathcal{A}} = \frac{1}{\# \text{pixels}(\Gamma)} \sum_{\mathbf{r} \in \Gamma} \left\{ \left[2(\mathbf{O}_Y - I_0) - 2\lambda(\nabla^T \nabla \mathbf{O}_Y) \right] \cdot \frac{\partial \mathbf{O}_Y}{\partial \mathcal{A}} \right\}, \quad (\text{s4-24})$$

where $\# \text{pixels}(\Gamma)$ denotes the total number of the pixels in Γ , and

$$\frac{\partial \mathbf{O}_Y}{\partial \mathcal{A}} = \frac{2(\mathbf{S}_Y^{IC} - 1)}{\mathbf{T}_{sc}^3} \left\{ \mathcal{S}_0[\mathbf{S}_Y^{IC}(\mathbf{r})] + \sum_{l=1}^{\mathcal{L}} \boldsymbol{\theta}_l(\mathbf{r}) \circ \mathcal{S}_l[\mathbf{S}_Y^{IC}(\mathbf{r})] \right\}. \quad (\text{s4-25})$$

The initial value of \mathcal{A} can be an arbitrary value within $[0.1, 0.2]$. Similar to Eq. (s4-23),

\mathcal{A} is also updated by the Adam-gradient method which is

$$\begin{cases} \mathbf{g} = \frac{\partial E}{\partial \mathcal{A}_n} \\ \boldsymbol{\varpi}_{n+1} = \tau \cdot \boldsymbol{\varpi}_n + (1-\tau) \cdot \mathbf{g} \\ \mathbf{v}_{n+1} = \xi \cdot \mathbf{v}_n + (1-\xi) \cdot \mathbf{g}^2 \\ \mathcal{A}_{n+1} = \mathcal{A}_n - \eta \cdot \frac{\boldsymbol{\varpi}_{n+1}}{\sqrt{\mathbf{v}_{n+1} + \varepsilon}} \end{cases}. \quad (\text{s4-26})$$

We choose $\tau = 0.9$ and $\xi = 0.99$ in all following experiments. Let $R_n = \text{mean}_{r \in I'} (\mathbf{O}_y - \bar{I})^2$ be part of residual of cost function in n -th iteration and we stop the iteration process when $R_n < 0.005$. Figure 4-5 shows the flow chart of the proposed dehazing algorithm.

Supplementary Note 4-5: Treatment for Pb/Pr channels

In this section we analyze the treatment for Pb/Pr channels and show that this process can be skipped since it won't influent the final dehazing quality of MUTE. According to Eq. (4-3) in the main text, the image formation for Pb/Pr channel is given by

$$\begin{bmatrix} \mathbf{S}_{\text{Pb}}(\mathbf{r}) \\ \mathbf{S}_{\text{Pr}}(\mathbf{r}) \end{bmatrix} = \mathbf{L}(\mathbf{r}) \circ \mathbf{T}_{\text{lens}}^2(\mathbf{r}) \circ \mathbf{T}_{\text{sc}}^2(\mathbf{r}) \circ \begin{bmatrix} \mathbf{O}_{\text{Pb}}(\mathbf{r}) \\ \mathbf{O}_{\text{Pr}}(\mathbf{r}) \end{bmatrix}. \quad (\text{s4-27})$$

With known of $\mathbf{L}(\mathbf{r}) \circ \mathbf{T}_{\text{lens}}^2(\mathbf{r})$ from illumination correction in Section 3.2 in the Chapter 4, and $\mathbf{T}_{\text{sc}}^2(\mathbf{r})$ from dehazing in Section 4, the Pb/Pr channel can be restored by

$$\mathbf{O}_{\text{Pb/Pr}}(\mathbf{r}) = \frac{\mathbf{S}_{\text{Pb/Pr}}(\mathbf{r})}{\mathbf{L}(\mathbf{r}) \circ \mathbf{T}_{\text{lens}}^2(\mathbf{r}) \circ \mathbf{T}_{\text{sc}}^2(\mathbf{r})}. \quad (\text{s4-28})$$

However, since \mathbf{T}_{sc} includes many pixels of small value, Eq. (s4-28) can lead to severe color distortion if the image is converted back to the RGB color space. Therefore, a color correction treatment is applied to both \mathbf{O}_{Pb} and \mathbf{O}_{Pr} after Eq. (s4-28). Here we use the min-max approach to correct the color difference, which is given by

$$\mathbf{O}_{P, \text{correction}} = [\max(\mathbf{S}_P) - \min(\mathbf{S}_P)] \cdot \left[\frac{\mathbf{O}_P - \min(\mathbf{O}_P)}{\max(\mathbf{O}_P) - \min(\mathbf{O}_P)} \right] + \min(\mathbf{S}_P), \quad (\text{s4-29})$$

where $P \in \{\text{Pb}, \text{Pr}\}$.

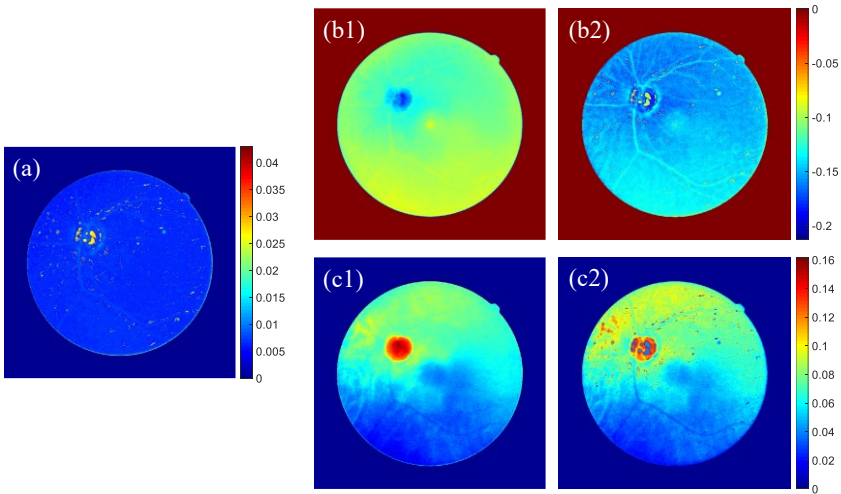


Fig. s4-5. Treatment of Pb and Pr channels for image in Fig. 3 (d1). (a) production of $L(\mathbf{r}) \circ T_{lens}^2(\mathbf{r}) \circ T_{sc}^2(\mathbf{r})$. (b1) and (b2) are the Pb channels before and after treatment. (c1) and (c2) are the Pr channels before and after treatment.

Fig. s4-5(a) shows the product result of $L(\mathbf{r}) \circ T_{lens}^2(\mathbf{r}) \circ T_{sc}^2(\mathbf{r})$ for image in Fig. 4-3 (d1). The Pb and Pr channel before and after treatment are shown in Fig. s4-5 (b1) to (b2) and Fig. s4-5 (c1) to (c2). Then we combine the treated Pb and Pr channel with the dehazed Y channel and transform them back to the RGB color space, yielding dehazed retinal images.

Fig. s4-6 demonstrates 5 groups of raw images (the first row) and corresponding dehazed images (the second row). We also found according to the experiment that the treatment for the Pb/Pr channel can be skipped as it will not lead to large differences in the final enhanced images, as shown in the third row of Fig. s4-6. Therefore, in our experiment, we skip the Pb/Pr channel process.

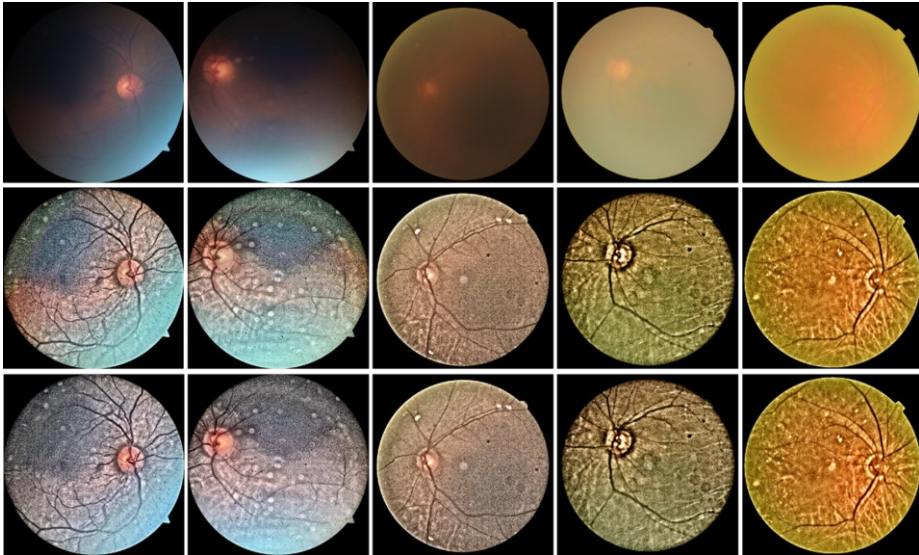


Fig. s4-6. Demonstration of proposed retinal image dehazing. First row: raw images. Second row: enhanced images with Pb and Pr channel treatments. Third row: enhanced images without Pb and Pr channel treatments.

Supplementary Note 4-6: Additional results for cataractous retinal images using MUTE

In this section of the Supplementary Materials, additional three groups of experimental results using MUTE are shown for visual assessment.

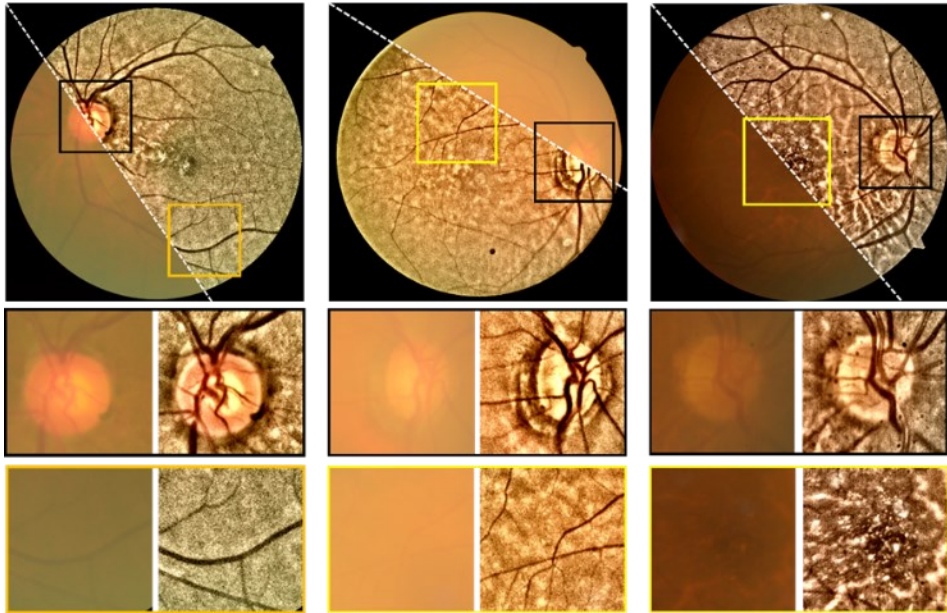


Fig. s4-7. Three additional groups of experimental results for MUTE.

Supplementary Note 4-7: Enhancement results for non-cataractous retinal images using MUTE

In this section of the Supplementary Materials, additional four groups of experimental results are shown for visual assessment. The raw images are from DRIVE and DiaRetDB01 datasets.

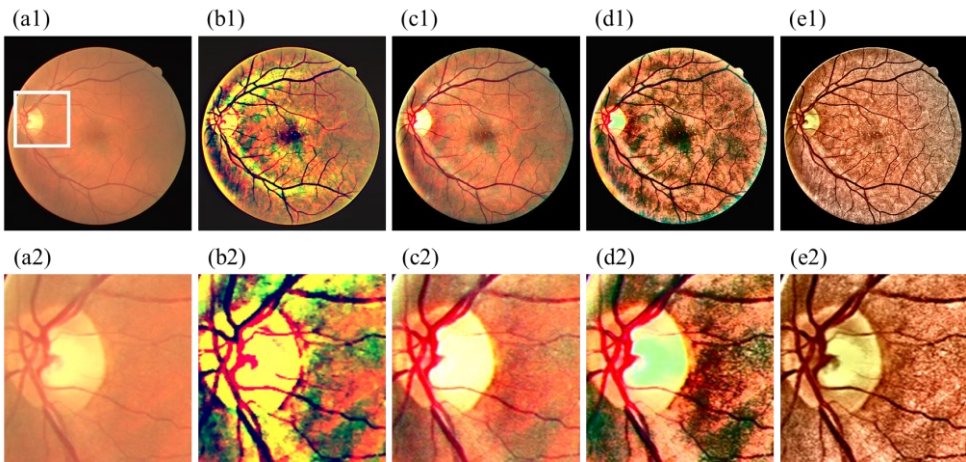


Fig. s4-8. Group 1 for visual assessment. The raw image is from DRIVE dataset. Left to right are raw images, results of DVA, DRC, LPAR, and MUTE methods.

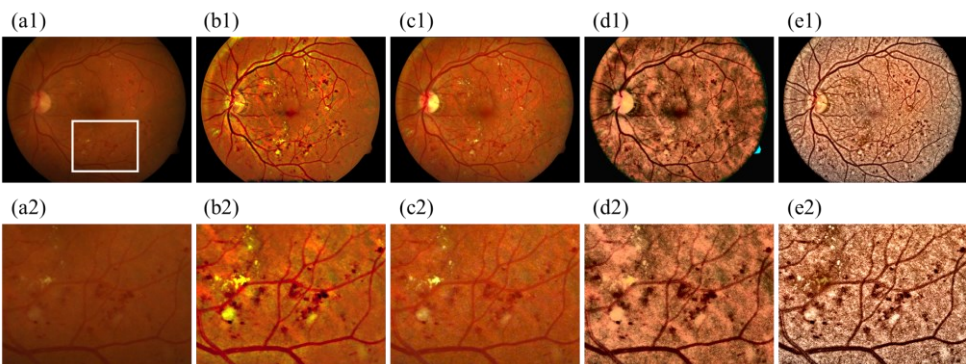


Fig. s4-9. Group 2 for visual assessment. The raw image is from DiaRetDB1 dataset. Left to right are raw image, results of DVA, DRC, LPAR, and MUTE methods.

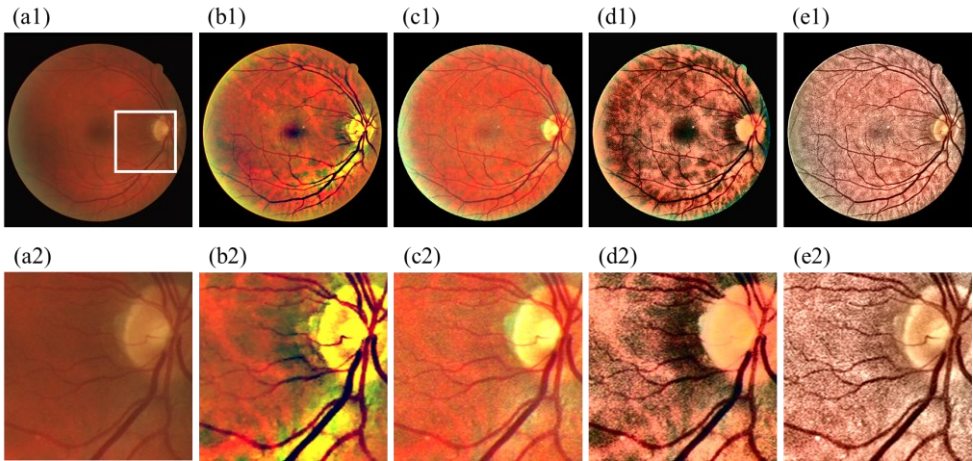


Fig. s4-10. Group 2 for visual assessment. The raw image is from DRIVE dataset. Left to right are raw images, results of DVA, DRC, LPAR, and MUTE methods.

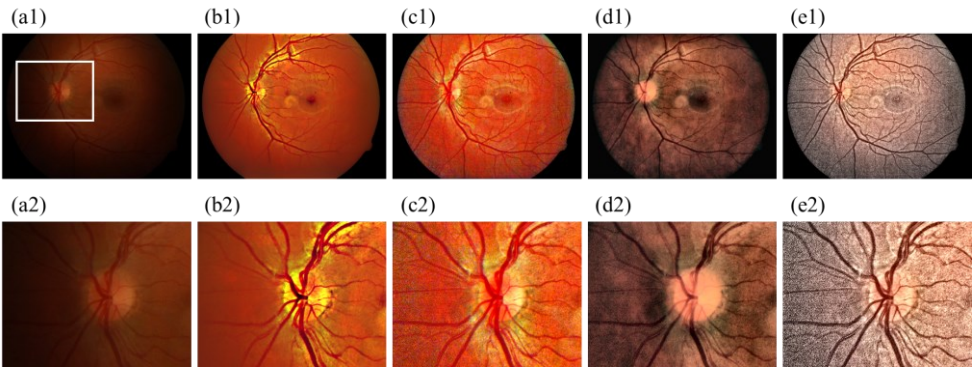


Fig. s4-11. Group 2 for visual assessment. The raw image is from DiaRetDB1 dataset. Left to right are raw image, results of DVA, DRC, LPAR, and MUTE methods.

Supplementary Note 4-8: Quantitative evaluation of blood vessel segmentation using simulation data

In section 5.3 in the main text, we showed that MUTE benefits the blood vessel segmentation for cataractous retinal images. However, as there is no ground-truth blood vessel map for cataractous retinal images due to the difficulty of manual segmentation for cataractous images, we cannot perform a quantitative evaluation to show the performance of MUTE.

In order to quantitatively evaluate the segmentation results, in this section, we perform simulations and generate the cataractous retinal image using the DPFR model. The raw images are taken from the DRIVE dataset, with paired ground-truth blood vessel map as shown in Fig. s4-12. As shown in Fig. s4-12 (d1) and (d2) the images are severely hazed due to the small value of the transmission map in Fig. s4-12 (c).

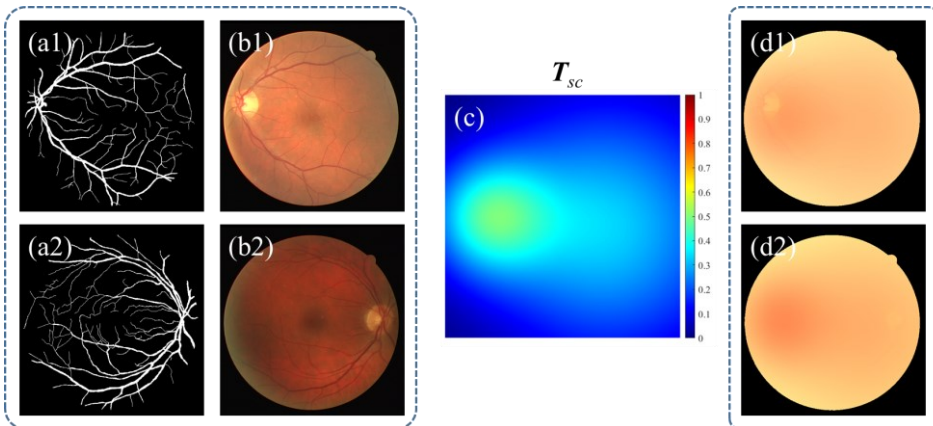


Fig. s4-12. Simulation generation of cataractous retinal images (d1) and (d2) using raw images in (b1) and (b2) and transmission map (c). (a1) and (a2) are ground-truth blood vessel maps.

Based on the simulated cataractous images in Fig. s4-12 (d1), we perform retinal image enhancement using DVA, DRC, LPAR, and MUTE. Enhancement results are shown in Fig. s4-13 (a1) to Fig. s4-13 (a5). Where both LPAR and MUTE show their ability on retinal image dehazing, which is corresponding to the results on real data. Then we

perform blood vessel segmentation using the Frangi filter as shown in Fig. s4-13 (b1) to (b5), and Matched filter as shown in Fig. s4-13 (c1) to (c5). The segmentation results are also consistent with those of real data shown in the main text.

Segmentation results for Fig. s4-12(d2) are shown in Fig. s4-14, where the filter responses on MUTE are also better than on enhanced images of other SOTA methods.

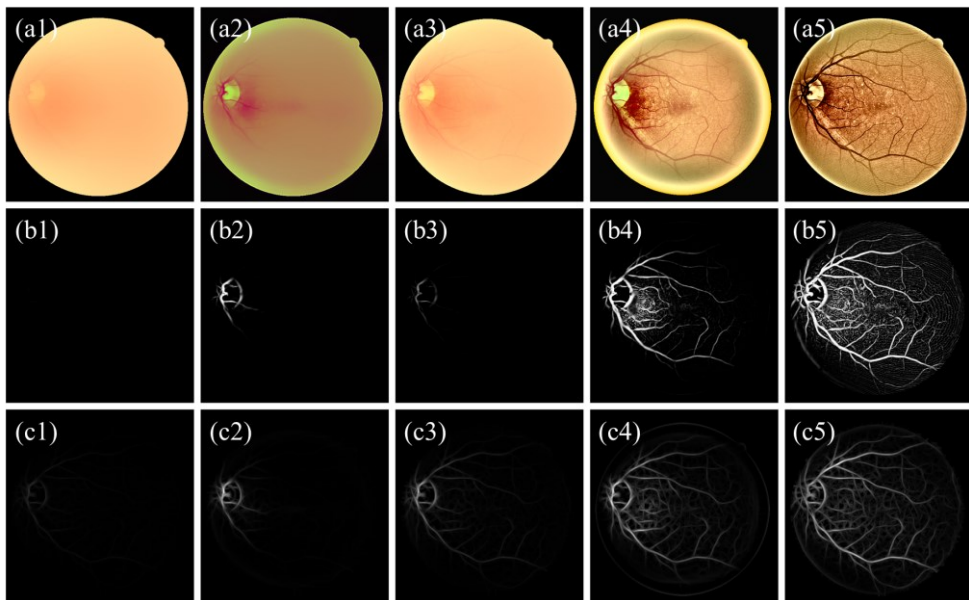


Fig. s4-13. Demonstration of retinal blood vessels segmentation for the cataractous retinal image. (a1) to (a5) are raw, and enhanced images by DVA, DRC, LAPR, and MUTE, respectively. (b1) to (b5) are segmentation results of the Frangi filter. (c1) to (c5) are results for the Match filter.

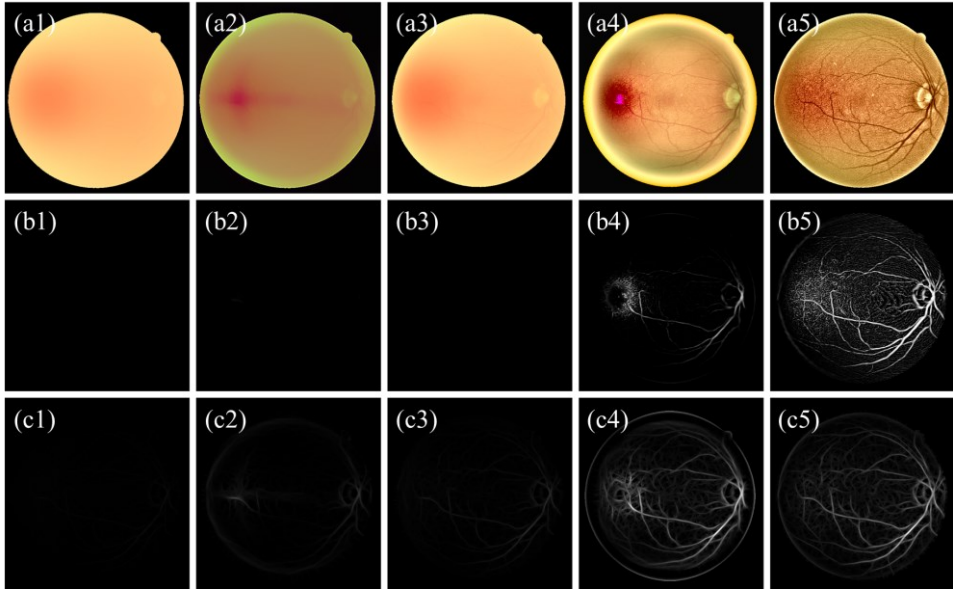


Fig. s4-14. Demonstration of retinal blood vessels segmentation for the cataractous retinal image. (a1) to (a5) are raw, and enhanced images by DVA, DRC, LAPR, and MUTE, respectively. (b1) to (b5) are segmentation results of the Frangi filter. (c1) to (c5) are results for the Match filter.

We use the five matrixes, including Accuracy (ACC), Precision (PSS), Recall, F_1 -value, and Dice score, to evaluate the segmentation results for different enhancement methods. The matrixes are calculated according to:

Matrix	Formular
Accuracy	$ACC = \frac{TP + TN}{TP + TN + FP + FN}$
Precision	$PSS = \frac{TP}{TP + FP}$
Recall	$Recall = \frac{TP}{TP + FN}$
F_1 -value	$F_1 = 2 / \left(\frac{1}{PSS} + \frac{1}{Recall} \right)$
Dice score	$Dice = 2 \frac{V_{seg} \cap V_{gt}}{V_{seg} + V_{gt}}$

where TP is True Positives denoting any pixels marked as the vessel in both ground truth and segmented image; FP is False positive denoting any pixels marked as a vessel pixel in segmented image which is marked as background pixel in ground truth; FN is False Negative denoting any pixels marked as a background pixel in segmented image which is marked as vessel pixel in ground truth; TN is True Negative denoting any pixels marked as background in both ground truth and segmented image. V_{seg} is the binary set of pixels denoting the vessels in the segmentation image. V_{gt} is the binary set of pixels denoting the vessels in the ground-truth image. The larger value of ACC denotes better segmentation quality, and so do the other four matrixes.

Those matrixes are calculated based on the binarized blood vessel map, and we binarized each gray-scale blood vessels map in Fig. s4-13 and Fig. s4-14 according to the regressed threshold τ_{re} by solving the following 1-D optimization problem

$$\tau_{re} = \arg \max_{\tau \in N, \tau < 256} Dice(V_{seg}, V_{gt}), \quad V_{seg} = \begin{cases} 1, & Map(x, y) \geq \tau \\ 0, & else \end{cases}$$

where Map denotes the gray-scale blood vessels map given by the Frangi filter or Matched filter. We use the regressed threshold to bypass the effects of weak filter response due to low image contrast. For example, the blood vessel structures in Fig. s4-14 (c1) do exist, but can hardly be observed due to their small gray values (< 10) since the hazed image has very low contrast and the filter responses on the vessel structures are very weak.

Tab. s1 lists the segmentation quality for the hazed image in Fig. s4-12 (d1). We also calculated the segmentation quality for raw non-hazed images (not shown), as listed in the third column of Tab. s4-1, the segmentation quality based on the raw image is good. While the quality directly based on the simulated hazy image is problematic as listed in the fourth column of Tab. s4-1, where the F_1 -value, Accuracy, and Dice scores are very small. While the segmentation based on MUTE enhanced image gains a larger score than other SOTA methods, especially for Dice score. This implies that the two filters can

recognize more vessel structure on MUTE enhanced images than on other images, which shows that MUTE benefits the segmentation of cataractous retinal images. A similar analysis is applied to Tab. s4-2.

Tab. s4-1. Segmentation quality for Fig. s4-15 (d1) with Frangi and Matched filters

Fig. s4-15(d1)	Matrix	Segmentation based on					
		Non-hazed raw image	Hazed image	DVA	DRC	LPAR	MUTE
Frangi Filter	ACC	0.1192	0.0016	0.0360	0.0275	0.1346	0.1123
	PSS	0.7306	0.9694	0.5398	0.7616	0.4378	0.5673
	Recall	0.7778	0.0143	0.1738	0.1873	0.5263	0.5692
	F ₁ -value	0.7534	0.0281	0.2629	0.3006	0.4780	0.5682
	Dice score	0.7493	0.0278	0.2607	0.2979	0.4755	0.5650
Matched Filter	ACC	0.1187	0.1281	0.3046	0.1200	0.1301	0.1163
	PSS	0.6761	0.4946	0.2381	0.5059	0.4443	0.5820
	Recall	0.7169	0.5663	0.6479	0.5423	0.5163	0.6046
	F ₁ -value	0.6959	0.5281	0.3482	0.5234	0.4776	0.5931
	Dice score	0.6921	0.5252	0.3471	0.5206	0.4751	0.5898

Tab. s4-2. Segmentation quality for Fig. s4-15 (d2) with Frangi and Matched filters

Fig. s4-15(d1)	Matrix	Segmentation based on					
		Non-hazed raw image	Hazed image	DVA	DRC	LPAR	MUTE
Frangi Filter	ACC	0.1042	0.0000	0.0058	0.0019	0.1481	0.1108
	PSS	0.7442	NaN	0.2451	0.9951	0.3783	0.4922
	Recall	0.6618	0.0000	0.0122	0.0160	0.4782	0.4655
	F ₁ -value	0.7006	NaN	0.0233	0.0316	0.4224	0.4785
	Dice score	0.6934	0.0000	0.0228	0.0310	0.4188	0.4738
Matched Filter	ACC	0.1253	0.1400	0.2497	0.1786	0.1687	0.1199
	PSS	0.6459	0.4601	0.2331	0.4160	0.3731	0.5450
	Recall	0.6910	0.5499	0.4967	0.6343	0.5372	0.5578
	F ₁ -value	0.6677	0.5010	0.3173	0.5024	0.4404	0.5513
	Dice score	0.6614	0.4966	0.3154	0.4986	0.4369	0.5461

In general, the MUTE removes the hazed effect of cataractous retinal images and improves the image contrast, which benefits the blood vessels segmentation tasks on cataractous retinal images.

Supplementary Note 4-9: Quantitative evaluation of retinal image registration using simulation data

In section 5.4 we showed that MUTE benefits the retinal image registration of cataractous retinal images by suppressing the haze effect. However, as there is no ground-truth for a cataractous retinal image due to the difficulty of manual feature segmentation for cataractous images, we cannot perform a quantitative evaluation to show the performance of MUTE.

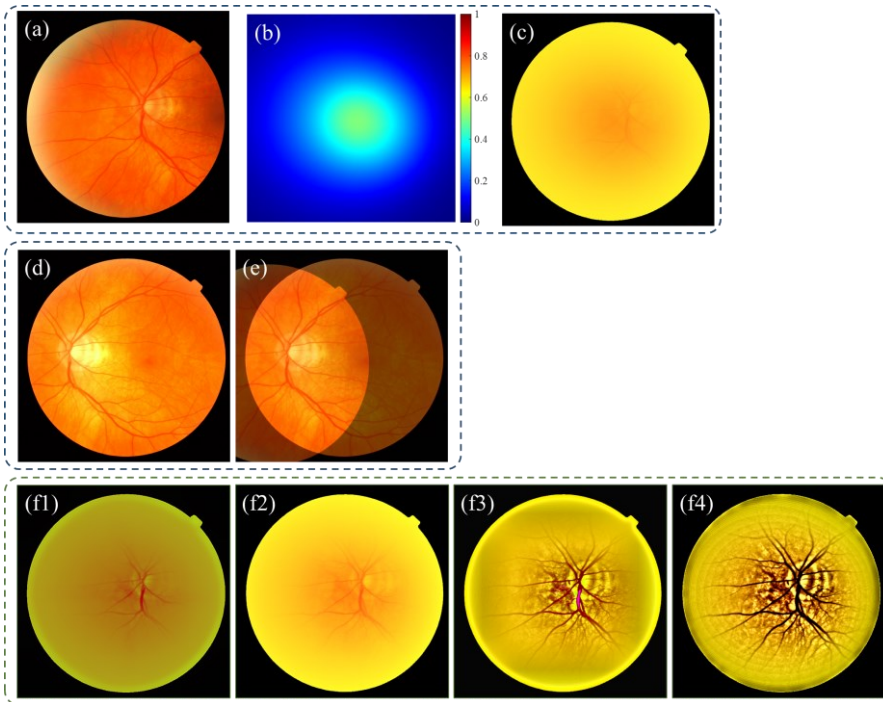


Fig. s4-15. Simulation generation of cataractous retinal images for image registration. (a) raw image. (b) transmission map. (c) simulated cataractous retinal image. (d) paired image for registration. (e) registration results. (f1) to (f4) are enhancement results with DVA, DRC, LPAR, and MUTE methods, respectively.

In order to quantitatively evaluate the registration results, in this section we perform a simulation experiment based on paired retinal images as shown in Fig. s4-15

(a) and Fig. s4-15 (d) and generate the cataractous retinal image using the DPFR model as shown in Fig. s4-15 (c). The ground-truth registration result is shown in Fig. s4-15 (e). Enhancement results for four different methods are shown in Fig. s4-15 (f1) to (f4).

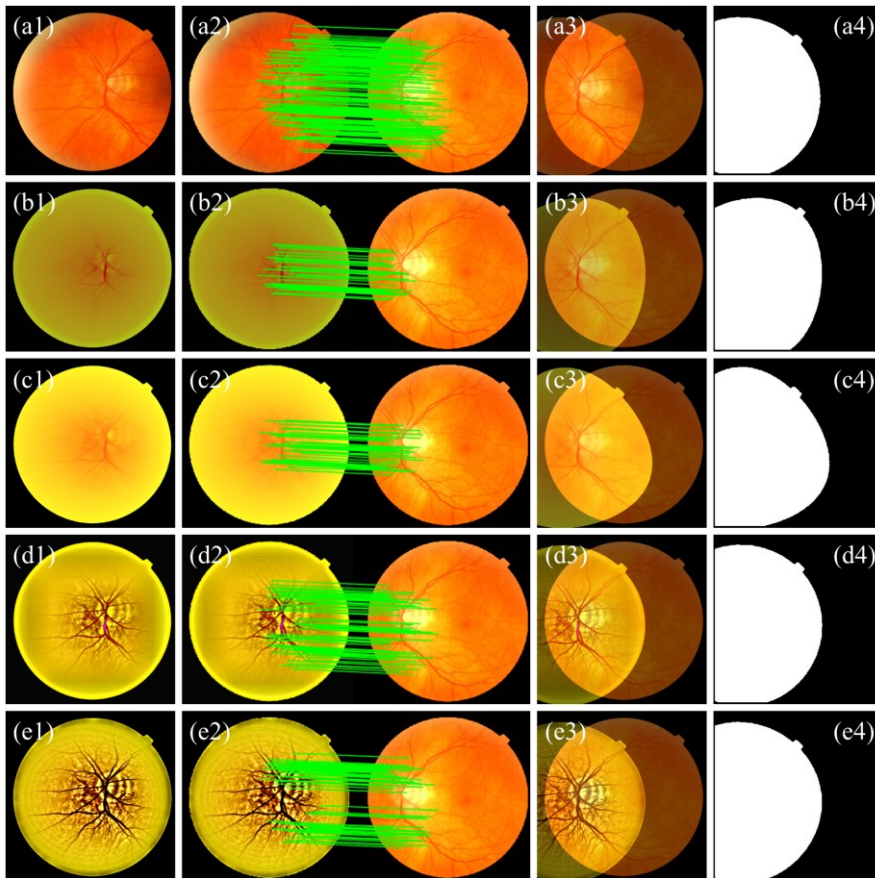


Fig. s4-16. Demonstration of retinal image registration on the simulated cataractous retinal image. (a1) to (e1) are images to be registered to Fig. s4-15 (d). (a2) to (e2) are features for registration. (a3) to (e3) are registration results. (a4) to (e4) are masks for the region of interest.

Image registration results are shown in Fig. s4-16, where the registration qualities for LPAR and MUTE are good according to the visual assessment as shown in Fig. s4-16 (d3) and (e3). The binary masks for the region of interest are shown in Fig. s4-16 (d4)

and (e4). The mask for ground truth is shown in Fig. s4-16 (a4). The registration quality precision can be directly calculated from the root mean square (RMS) of the difference between the ground-truth mask, and masks for other methods. The smaller the RMS, the better the registration quality.

As listed in Tab. s4-3, the MUTE has the lowest RMS score denoting that registration using MUTE enhanced image has better quality than other methods.

Tab. s4-3. Registration quality for enhanced image with different methods.

Matrix	Registration based on the enhanced image			
	DVA	DRC	LPAR	MUTE
RMS	0.1888	0.2006	0.1100	0.0798

Supplementary Note 4-10: Further increasing the speed of MUTE using down-up-sampling strategy

As listed in Tab. 4-4 in the main text, the executing time for MUTE increases significantly when the size of the image becomes large. In this section of the Supplementary Materials, we propose the down-up-sampling strategy to speed up the MUTE and maintain the enhancement quality.

Let Y_H be the Y channel of the raw retinal images, and \tilde{Y}_H be the Y channel of the final output (enhanced image) of MUTE. We consider a pixel-wise transfer matrix \mathcal{T}_H that converts the Y_H to \tilde{Y}_H in an end-to-end manner, and we write the following equation

$$\tilde{Y}_H = \mathcal{T}_H \circ Y_H. \quad (\text{s4-30})$$

To determine the \mathcal{T}_H , we first down-sample the raw high-resolution image to a small size image. The Y channel of the small size image is noted by Y_L . Then, the Y_L is

processed by MUTE which returns us the processed (enhanced) Y channel \tilde{Y}_L . Then we calculate the transfer matrix \mathcal{T}_L through pixel-wise division

$$\mathcal{T}_L = \frac{\tilde{Y}_L}{Y_L + \delta}, \quad (\text{s4-31})$$

where δ is a small value to avoid the divided by 0 conditions. Then we up-sample the \mathcal{T}_L to the same size of the raw high-resolution image to obtain an estimated version of \mathcal{T}_H , and finally, we use Eq. (s4-30) to obtain the enhanced \tilde{Y}_H . The entire progress can be expressed by

$$\tilde{Y}_H = \text{Us} \left\{ \frac{\text{Mute}[\text{Ds}(Y_H)]}{\text{Ds}(Y_H) + \delta} \right\} \circ Y_H, \quad (\text{s4-32})$$

where Ds and Us denote down- and up-sampling operations. Mute denotes our proposed retinal image enhancement methods.

Fig. s4-17 and Fig. s4-18 show the enhanced results where the sizes of raw images are 2448 by 2448. We down-sample the raw images into 875 by 875 and calculate the \mathcal{T}_L so that MUTE takes only **17 sec.** to finish the retinal image enhancement. After then, \mathcal{T}_L is up-sampled into 2448 by 2448 which is regarded as an estimation of \mathcal{T}_H . As shown in Fig. s4-17 and Fig. s4-18, MUTE significantly increase the contrast of the raw image for both non-cataracts and cataract retinal images, and the executing time is reduced.

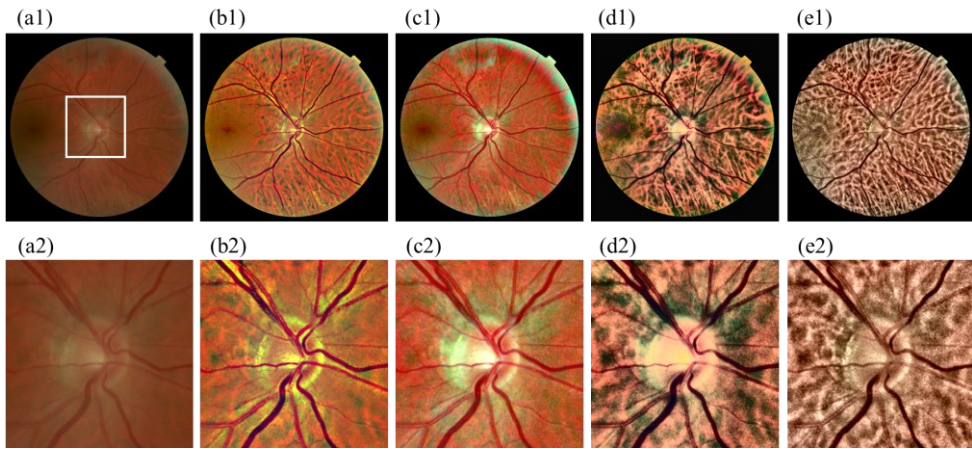


Fig. s4-17. Enhanced results for high-resolution non-cataractous retinal images. Left to right are raw images, results of DVA, DRC, LPAR, and MUTE methods.

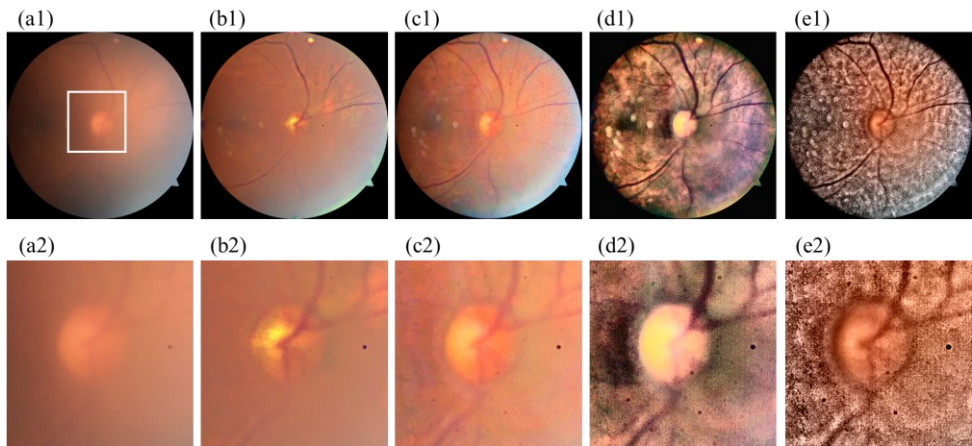


Fig. s4-18. Enhanced results for high-resolution cataractous retinal images. Left to right are raw images, results of DVA, DRC, LPAR, and MUTE methods.

Supplementary Note 4-11: MUTE with different denoisers and the reason of choose Gaussian filter as our denoiser

4-11.1 Visual assessment

In this section, we consider using the other two denoisers, which are domain transform recursive filter (DRF) [17] and total variation (TV) filter [18], to estimate the denoised layers. We demonstrate the MUTE on non-cataractous and cataractous retinal images as shown in Fig. s4-19 (a) and Fig. s4-19 (b), respectively.

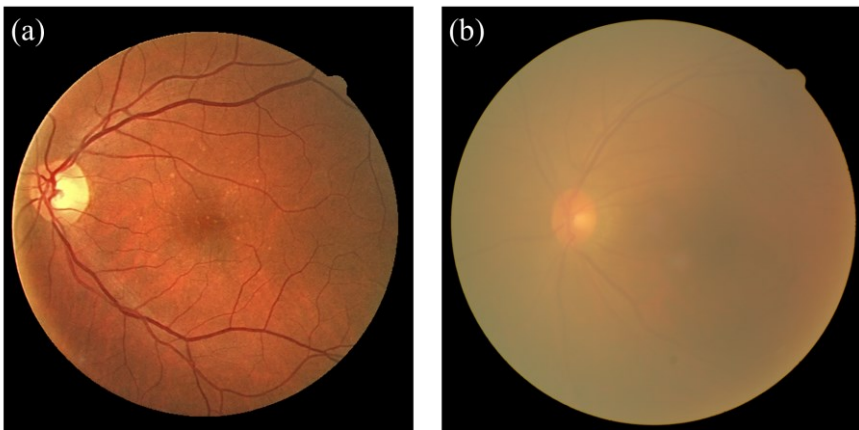


Fig. s4-19. Retinal images that were used to show the results of different denoisers.

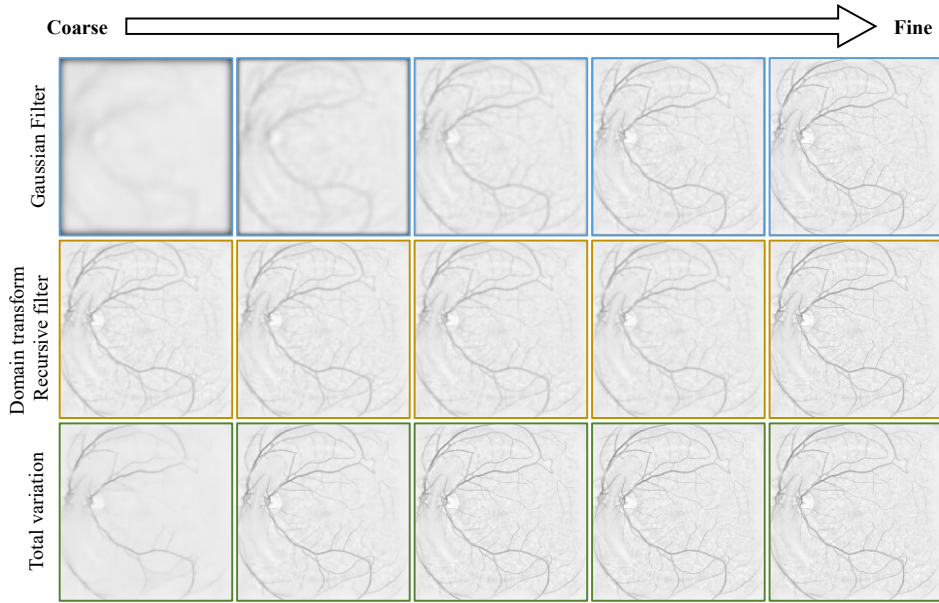


Fig. s4-20. Demonstration of S_l , ($l=0,1,2,3,4$) generated by different denoisers.

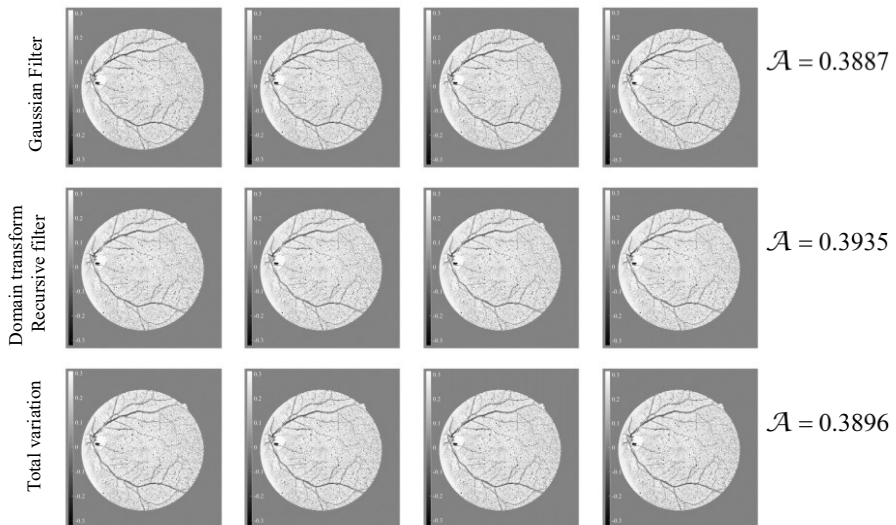


Fig. s4-21. Latent parameters φ_l and \mathcal{A} after optimization for different denoisers.

The domain transforms recursive filter has three parameters indication filter spatial standard deviation, filter range standard deviation, and noise standard deviation. The TV has one parameter corresponding to the penalty strength on the L_1 -norm of the image gradient. These parameters are varying from strong denoise to weak denoise according to the denoising level in MUTE. For example, as shown in Fig. s4-20, different denoisers generate different estimations of denoising layers.

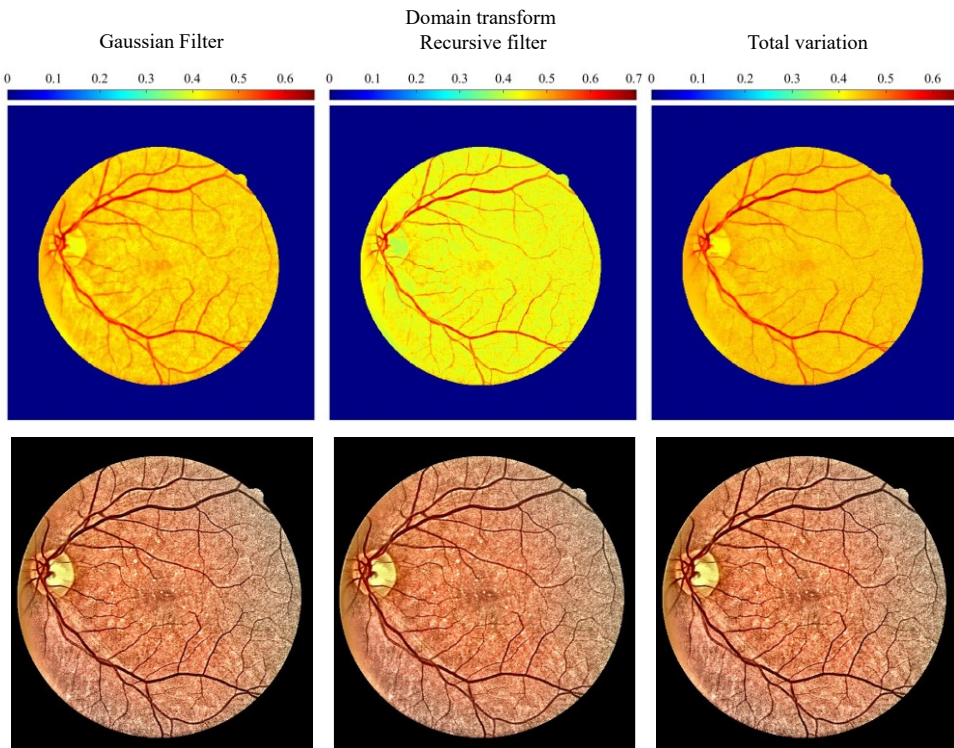


Fig. s4-22. MUTE output for Fig. s4-19 (a) with different denoiser. First row: pattern of T_{sc} . Second row: MUTE enhancements.

The optimization outputs including φ_l and \mathcal{A} for three denoisers are shown in Fig. s4-21. The latent parameters φ_l for different denoisers and different layers are similar

from the visual aspect as they converge to similar values, however, they are different in details. The parameter \mathcal{A} also converges to different values but approaches 0.39.

As T_{sc} is the aggregate result of φ_l , \mathcal{A} , and \mathcal{S}_l , and is directly related to the dehazing results, we show T_{sc} together with the dehazed image in Fig. s4-22. As shown in the first row of Fig. s4-22, although the T_{sc} are different according to given denoisers, the final outputs are similar as the visual performances are the same. A similar analysis was applied to the retinal image in Fig. s4-19 (b), and the latent products are shown in Fig. s4-23 to Fig. s4-25.

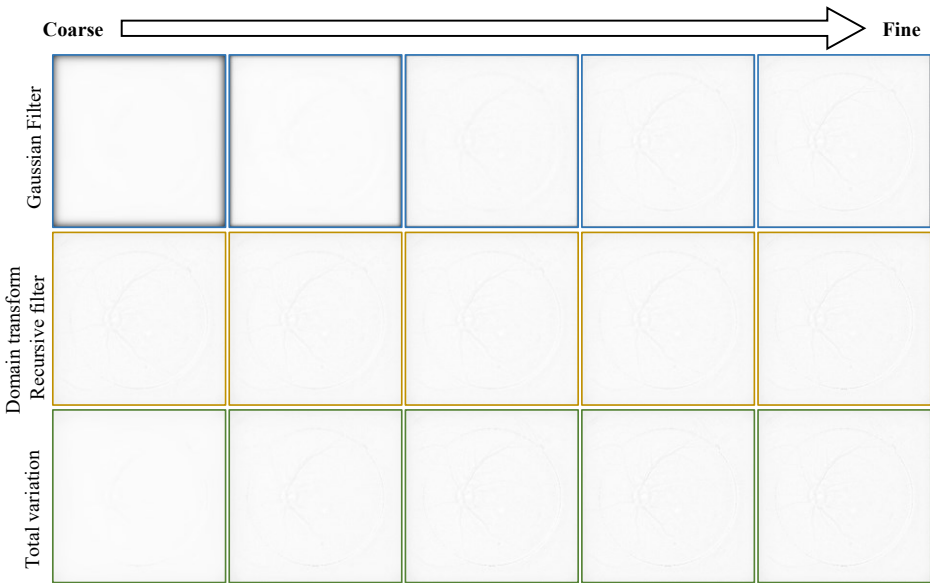


Fig. s4-23. Demonstration of \mathcal{S}_l , ($l=0,1,2,3,4$) generated by different denoisers for image in Fig. s4-19 (b).

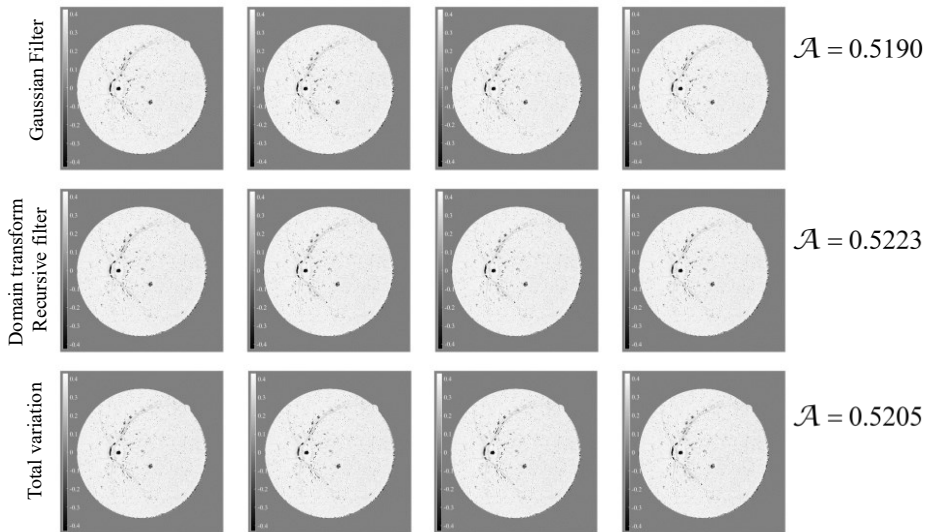


Fig. s4-24. Latent parameters φ_i and \mathcal{A} after optimization for different denoisers.

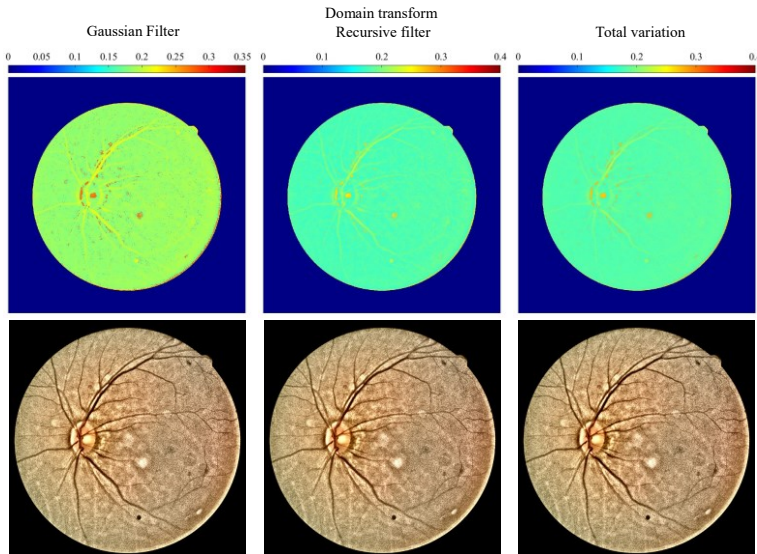


Fig. s4-25. MUTE output for Fig. s4-19 (b) with different denoiser. First row: pattern of T_{sc} . Second row: MUTE enhancements.

4-11.2 Objective assessment

Based on the domain transform recursive filter and TV regularization filter, we perform MUTE on the CATARACT dataset and compare their performance. The MUTE outputs are denoised with Hessian regularization $\lambda = \sigma / 30$. As listed in Tab. 4-6 in the maintext, there is no significant difference in the dehazing performance among three different denoisers. This shows the feature that MUTE is not sensitive to the type of denoiser, as mentioned in the main text.

Tab. s4-4. Quantitative assessment for MUTE with different denoisers.

Databases mean \pm std.	Matrixes	Raw	MUTE, Denoiser		
			Gaussian filter	DRF	TV
Cataract (100 images)	UISM	0.9002 \pm 0.1432	5.4597 \pm 0.4898	5.4325 \pm 0.4686	5.4417 \pm 0.5393
		UIConM		0.0573 \pm 0.0123	0.3632 \pm 0.0222
	IE		5.8443 \pm 0.4449	7.1898 \pm 0.1756	
		C _{RAMM}	1.2013 \pm 0.1741		6.7353 \pm 0.7251
	FADE		0.4715 \pm 0.1155	0.1790 \pm 0.0215	
		σ	0.5127 \pm 0.0724		3.1508 \pm 0.8263

Since MUTE is not sensitive to the denoiser, we can choose the simplest Gaussian filter to achieve the estimation of \mathcal{S}_f , as the Gaussian filter needs only one parameter (kernel size), and requires only two calculations of Fourier transform and one pixel-wise matrix multiplication.

Supplementary Note 4-12: Comparing against State-of-the-art deep-learning methods.

In this section, we compare our proposed MUTE against three SOTA deep-learning methods including ArcNet [19], ScrNet [20], and ArSrNet [21]. All three methods are published in 2022. We test the performance on cataractous retinal image dataset. For ArcNet and ScrNet, we used the published pre-trained model in <https://github.com/liamheng/Annotation-free-Fundus-Image-Enhancement>. For ArSrNet, we asked the author to process the images in their local machine and sent them back to us.

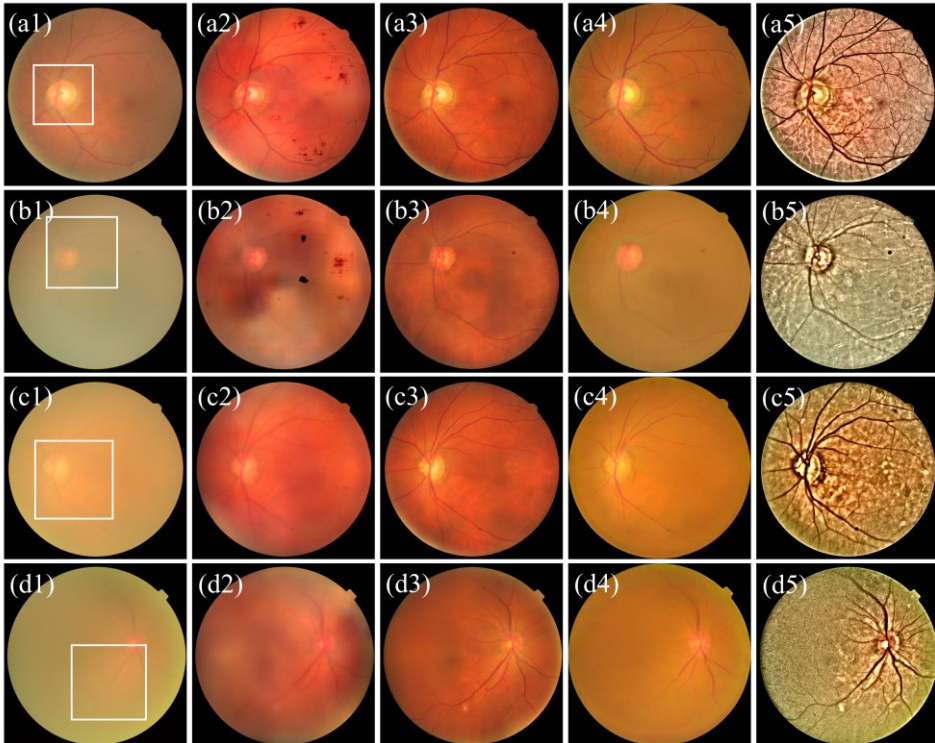


Fig. s4-26. Comparison against SOTA deep learning methods. (a1)-(d1) raw images. (a2)-(d2) outputs of ArcNet. (a3)-(d3) outputs of ScrNet. (a4)-(d4) outputs of ArSrNet. (a5)-(d5) output of MUTE.

The raw images are resized into 512 by 512 pixels in order to produce good results since these three networks are trained based on simulated cataract retinal images of the same size. Where the ArcNet is trained based on images of 256 by 256 pixels. For the larger size of input images, the network should be trained again due to their lack of generalization. While our proposed MUTE can be applied to any size of input images.

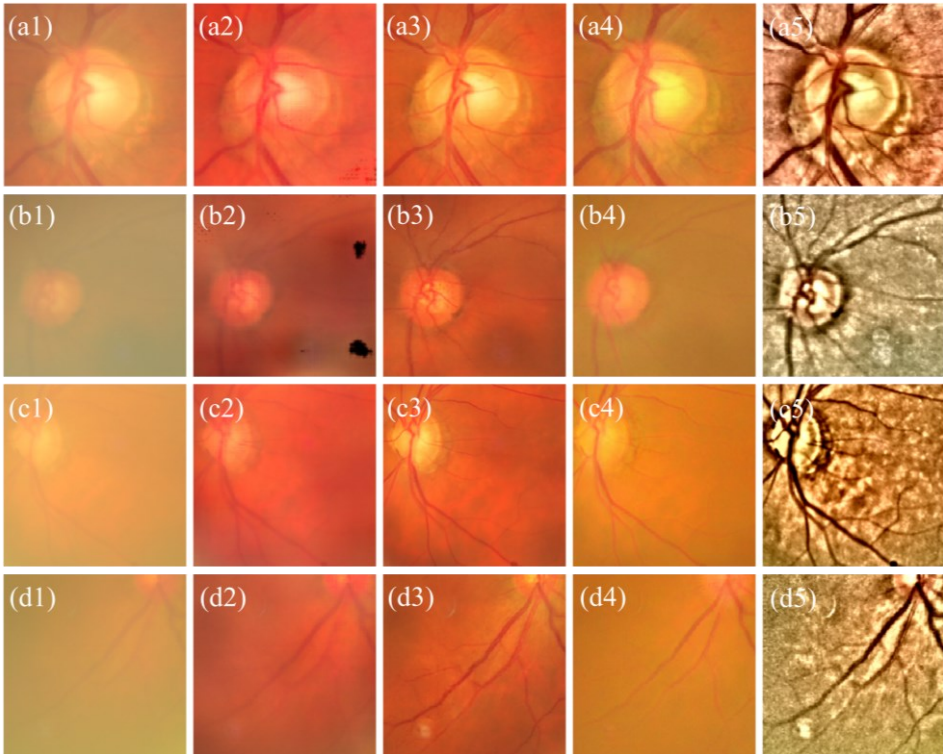


Fig. s4-27. Zoomed-in images for the region in the white boxes in Fig. s26. (a1)-(d1) raw images. (a2)-(d2) outputs of ArcNet. (a3)-(d3) outputs of ScrNet. (a4)-(d4) outputs of ArSrNet. (a5)-(d5) output of MUTE.

Fig. s4-26 and Fig. s4-27 show the output images of four methods. Except for Fig. s4-26 (a1), where the cataract is mild, the cataracts in Fig. s4-26 (b1) to (d1) are rather severe. The enhancements for all three SOAT Networks can merely improve the contrast of the image to a limited degree. The cataract effect as shown in Fig. s4-26 (b2)

for ArcNet, and Fig. s4-26 (b4) for ArSrNet are still pronounced. While the MUTE significantly improves the image contrast as shown in Fig. s4-26 (b5). Similar results can be also found in the zoomed-in images in Fig. s4-27.

Tab. s4-5. Quantitative assessment for retinal image enhancement results in Fig. s4-26.

Raw image	Matrixes	Raw	Methods			
			ArcNet	ScrNet	ArSrNet	MUTE
Fig. s4-26 (a1)	UISM	1.1599	1.7425	1.8711	1.9078	3.8646
	UIConM	0.0797	0.0826	0.1063	0.0727	0.3476
	IE	5.7880	5.9789	6.3074	5.8403	7.3208
	CRAMM	2.0155	2.0973	2.2553	1.9387	7.7648
	FADE	0.3924	0.2695	0.2837	0.2841	0.1352
	σ	0.5862	0.4426	0.9217	1.0030	3.5575
Fig. s4-26 (b1)	UISM	1.1188	1.5567	1.3384	1.4918	4.0252
	UIConM	0.0946	0.0794	0.0736	0.0757	0.3265
	IE	5.2076	5.8317	6.0489	5.5855	7.2288
	CRAMM	2.1192	1.9930	1.8268	1.8556	7.2337
	FADE	0.5572	0.3092	0.3629	0.3070	0.1489
	σ	0.5725	0.3590	0.6832	1.1588	1.3042
Fig. s4-26 (c1)	UISM	1.3660	1.6098	1.3196	1.5484	4.2620
	UIConM	0.0928	0.0760	0.0721	0.0773	0.2999
	IE	5.5182	5.7829	6.2343	5.4319	7.0442
	CRAMM	2.0319	1.8701	1.7859	1.7791	5.8601
	FADE	0.7310	0.3168	0.3788	0.3362	0.1505
	σ	0.6837	0.4354	0.5644	1.0740	2.9904
Fig. s4-26 (d1)	UISM	1.1144	1.5886	1.9574	1.2968	3.6856
	UIConM	0.0819	0.0658	0.1019	0.0676	0.2991
	IE	5.2409	5.5612	6.4547	4.8506	7.0558
	CRAMM	1.9275	1.8016	2.1814	1.6819	5.9793
	FADE	0.6085	0.3310	0.3409	0.3881	0.1842
	σ	0.5604	0.3427	0.7253	0.8832	1.5557

Note that the ArcNet is trained on images of 256 by 256 pixels, when it is applied to larger images, it will introduce unexpected problematic pixels as can be seen in the black spot in Fig. s4-26 (b2), and Fig. s4-27 (b2). In practical cases, the size of retinal images can be even larger than 2000 by 2000 pixels and vary between different brands of fundus cameras from different companies. The input image size does not influent the

output quality of MUTE thanks to the down-up-sampling strategy proposed in Supplementary Note 4-10.

Tab. s4-6. Quantitative assessment for retinal image enhancement results in cataract dataset. The images are resized to 512 by 512 pixels

Databases mean \pm std.	Matrixes	Raw	Methods			
			DVA	DRC	LPAR	MUTE
Cataract (100 images)	UI SM.	1.2050	1.7907	1.7668	1.7658	3.7250
		± 0.1054	± 0.1609	± 0.2522	± 0.2260	± 0.3782
	UI ConM	0.0691	0.0772	0.0876	0.0678	0.3315
		± 0.0163	± 0.0066	± 0.0091	± 0.0082	± 0.0224
	IE	5.8285	5.8318	6.2697	5.8190	7.2347
		± 0.4496	± 0.1535	± 0.1655	± 0.3155	± 0.1682
	C _{RAMM}	1.8163	2.0137	2.1168	1.8042	7.1012
		± 0.2616	± 0.1099	± 0.1669	± 0.1519	± 0.8085
	FADE	0.4481	0.2827	0.3023	0.3054	0.1460
		± 0.1050	± 0.0230	± 0.0253	± 0.0295	± 0.0143
σ	0.5825	0.4384	0.8012	0.9835	3.6337	
	± 0.0577	± 0.0762	± 0.1473	± 0.0922	± 1.3543	

The score for objective assessments is listed in Tab. s4-5 for images in Fig. s4-26, and Tab s4-6, for statistic results for 100 images in the cataractous retinal images dataset. Accordingly, our MUTE increases the image edge information (UI SM) larger than SOTA deep-learning method by 2-folds, and increases the image contrast (UI ConM, and C_{RAMM}) by 3-folds. The MUTE also decreases the haze effect of cataract retinal image (FADE) by 2-folds less than SOTA deep-learning methods.

References

1. K. He, J. Sun, and X. Tang, "Single Image Haze Removal Using Dark Channel Prior," (2009).
2. D. Christaras, H. Ginis, A. Pennos, and P. Artal, "Intraocular scattering compensation in retinal imaging," *Biomed. Opt. Express* **7**, 3996-4006 (2016).
3. A. Arias, H. Ginis, and P. Artal, "Light scattering in the human eye modelled as random phase perturbations," *Biomed. Opt. Express* **9**, 2664-2670 (2018).
4. A. Arias, and P. Artal, "Wavefront-shaping-based correction of optically simulated cataracts," *Optica* **7**, 22-27 (2020).
5. T. T. J. M. Berendschot, P. J. DeLint, and D. v. Norren, "Fundus reflectance—historical and present ideas," *Progress in retinal and eye research* **22**, 171-200 (2003).

6. R. E. Hogg, M. B. Zlatkova, U. Chakravarthy, and R. S. Anderson, "Investigation of the effect of simulated lens yellowing, transparency loss and refractive error on in vivo resonance Raman spectroscopy," *Ophthalmic and Physiological Optics* **27**, 225-231 (2007).
7. T. T. J. M. Berendschot, J. v. d. Kraats, M. J. Kanis, and D. v. Norren, "Directional model analysis of the spectral reflection from the fovea and para-fovea," *Journal of Biomedical Optics* **15**, 065005 (2010).
8. G. Meng, Y. Wang, J. Duan, S. Xiang, and C. Pan, "Single Image Dehazing with Boundary Constraint and Contextual Regularization," presented at the ICCV 20132013.
9. D. Berman, T. Treibitz, and S. Avidan, "Non-Local Image Dehazing," presented at the CVPR 20162016.
10. Q. Zhu, J. Mai, and L. Shao, "A Fast Single Image Haze Removal Algorithm Using Color Attenuation Prior," *IEEE Transactions on Image Processing* **24**, 3522-3533 (2015).
11. S. Zhang, "A reflective background filling algorithm for fundus images preprocessing," <https://github.com/ShuheZhang-MUMC/Fundus-image-background-filling>.
12. B. Savelli, A. Bria, A. Galdran, C. Marrocco, M. Molinara, A. e. Campilho, and F. Tortorella, "Illumination correction by dehazing for retinal vessel segmentation," *IEEE 30th International Symposium on Computer-Based Medical Systems* (2017).
13. M. Ju, C. Ding, Y. J. Guo, and D. Zhang, "IDGCP: Image Dehazing Based on Gamma Correction Prior," *IEEE Trans. Image Process.* **29**, 3104 - 3118 (2021).
14. D. Zosso, G. Tran, and S. Osher, "A unifying retinex model based on non-local differential operators," in *Computational Imaging XI(SPIE2013)*, p. 865702.
15. A. Galdran, A. Alvarez-Gila, A. Bria, J. Vazquez-Corral, and M. Bertalmío, "On the Duality Between Retinex and Image Dehazing," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 8212-8221 (2018).
16. D. P. Kingma, and L. J. Ba, "Adam: A Method for Stochastic Optimization," presented at the International Conference on Learning Representations2015.
17. E. S. Gastal, and M. M. Oliveira, "Domain transform for edge-aware image and video processing," in *ACM SIGGRAPH 2011 papers*(2011), pp. 1-12.
18. L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D: nonlinear phenomena* **60**, 259-268 (1992).
19. H. Li, H. Liu, Y. Hu, H. Fu, Y. Zhao, H. Miao, and J. Liu, "An Annotation-Free Restoration Network for Cataractous Fundus Images," *IEEE Transactions on Medical Imaging* **41**, 1699-1710 (2022).
20. H. Li, H. Liu, H. Fu, H. Shu, Y. Zhao, X. Luo, Y. Hu, and J. Liu, "Structure-Consistent Restoration Network for Cataract Fundus Image Enhancement," in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2022*, L. Wang, Q. Dou, P. T. Fletcher, S. Speidel, and S. Li, eds. (Springer Nature Switzerland, Cham, 2022), pp. 487-496.
21. B. Yang, H. Zhao, L. Cao, H. Liu, N. Wang, and H. Li, "Retinal image enhancement with artifact reduction and structure retention," *Pattern Recognition* **133**, 108968 (2023).

Chapter 5

InQue:

An intensity-quenching scheme for large-scale fully unsupervised retinal image enhancement

Shuhe Zhang, Carroll A. B. Webers, and Tos T. J. M. Berendschot

Abstract

We present a semi-learning-based algorithm called intensity-quenching (InQue) to achieve fully unsupervised single retinal image enhancement. It doesn't need any synthetic images in any form, the only input is a single retinal image. InQue consists of four modules namely color implantation, luma heater, luma cooler and luma rectifier. In the color implantation module, the background color of the input retinal image is replaced by a desired color, sampled from a normal retinal image. The image is then converted into the YCbCr color space, and the luma heater, a one-layer network, increases the intensity of the Y-channel of the color-implanted image to a certain level while correcting uneven or insufficient illumination. Next, the luma cooler, also a one-layer network, decreases the intensity of the Y-channel, increasing the contrast of the image. The cooled-down Y channel is further rectified by the luma rectifier including denoising and unifying the intensity distribution. Finally, the output Y channel, together with the original Cb and Cr channels, is then converted back into the RGB color space, yielding the final enhanced retinal image. We tested InQue on cataract retinal images from public databases and compared its performance with other state-of-the-art deep-learning-based methods using both visual and objective assessments, where it outperformed the others. We also conducted ablation studies on each module and demonstrated potential applications of InQue in enhancing ultra-wide-field retinal images, retinal image vessel tracking and diagnosing.

Keywords

Retinal image enhancement; Unsupervised learning; Cataract retinal image; Large scale image.

1. Introduction

The invention of the ophthalmoscope by Helmholtz in the 19th century enabled ophthalmologists to obtain images of the retina, leading to the emergence of ophthalmology as a separate subfield of medicine. Since then, retinal images have been widely used for early detection, diagnosis, and monitoring of ocular diseases such as diabetic retinopathy [1, 2], glaucoma [3-5] and age-related macular degeneration [6-8] as well as for diagnosing neurological diseases like stroke and cognitive dysfunction and cardiovascular diseases [9-11].

However, sufficient quality of retinal images is crucial for accurate diagnosis and monitoring of ocular and neurological diseases. Unfortunately, a significant percentage of acquired retinal images are of low quality, with unsatisfactory images ranging from 3.7% to 19.7% in studies of diabetic retinopathy [12] and mydriatic patients [13].

While modifying the optical system of a fundus camera can improve retinal image quality, such methods can be expensive and inconvenient for universal use. As a result, image processing methods have become increasingly important. Various algorithms have been investigated for enhancing retinal images, including deep-learning and traditional algorithms.

In this research, we propose a semi-learning method called Intensity-Quenching (InQue) to achieve fully unsupervised single retinal image enhancement. In Section 2, we briefly review related works and their limitations. In Section 3, we introduce the framework of InQue and explain the quenching process in detail. Section 4 presents the experimental results of InQue and compares them against state-of-the-art methods. Section 6 shows some clinical applications of InQue including diagnosis using an enhanced-image, enhancement on ultra-wide-field images, and blood segmentations. Section 6 is an ablation study on the main modules of InQue. Finally, we provide concluding remarks in Section 7.

2. Related works

2.1 Non-learning-based methods

Retinal images can be enhanced by traditional image enhancement methods, such as gamma correction for illumination adjustment, histogram equalization (HE), and its improved version including but not limited to the contrast limited adaptive histogram equalization (CLAHE) for contrast improvement. For example, Zhou et al. proposed in [14] to achieve luminosity and contrast adjustment based on the combination of gamma correction in HSV color space and CLAHE in CIE-Lab color space. Gupta et al. further improves this idea in [15] using *q*-quantile-based HE. Although these methods are simple and quick, they cannot be applied to retinal images affected by complex issues such as uneven illumination and cataractous effects.

Filtering-based methods have been proposed to achieve better illumination correction and contrast enhancement. These methods are mainly based on the Retinex theory, which divides a single retinal image into background (low-frequency) and structure (high-frequency) components and enhances the structures by assigning a larger weight to them. For instance, Cao et al. proposed a low-pass Gaussian filtering and α -rooting method to stretch the contrast of retinal images [16], while Wang et al. proposed to use total variation as the low-pass filter to separate the background and detailed components [17].

Filtering-based methods can be combined with an image formation model to further improve the methods' physical consistency. Peli et al. [18] proposed a cataract retinal image formation model and used a Gaussian filter to estimate the transmission matrix of the cataract layer. Mitra et al. [19] followed a similar idea and incorporated the HE into the contrast enhancement procedure to improve the contrast of retinal images with thin cataracts. Xiong et al. [20] used the natural scene hazing model for retinal image enhancement. The natural scene hazing model can be extended for illumination correction [21] and combined with dark-channel prior and its variation for retinal

image dehazing [22, 23]. Zhang et al. [22] proposed a double-pass fundus reflection model which is more physically consistent compared with former image formation models for retinal imaging.

Non-learning-based methods have high interpretability and do not rely on specific data distributions. They can improve the contrast of retinal images, but sometimes severely distort the color and naturalness of the output images.

2.2 Learning-based methods

With the development of computational power, deep-learning-based retinal image enhancements attracted a lot of research interest. Due to the lack of paired real retinal images for good and degenerated quality, most of learning-based retinal image restoration methods published recently can be categorized as extensions of GAN. These methods convert the retinal image restoration task into a style-transform task that transforms the image style from a bad-quality retinal image to a good-quality one. To mitigate the risk of GANs introducing unexpected artifacts, many works focus on preserving information fidelity.

Since there are no paired real retinal images, researchers use synthetic/simulated degenerated retinal images to train the networks. For instance, based on the image formation model proposed by Peli et al. [18], Luo et al. [24] trained an unpaired GAN to achieve cataract retinal image dehazing for mild cataract cases. Li et al. [25] proposed an annotation-free GAN for cataractous retinal image restoration. Based on the natural scene haze formation model, Yang et al. [26] trained a modified cycle-GAN for artifact reduction and structure retention in retinal image enhancement. Shen et al. [27] proposed a new mathematical model to formulate the image-degrading process of fundus imaging and train a network for retinal image restoration. Other researchers have modified the structures of the network or loss function to improve the performance of the networks [28, 29].

While these learning-based methods produce impressive restoration results in both quality and naturalness preservation, they have limitations. Overfitting on synthetic data and lack of generalization are potential issues as we will show in the experimental sections. Additionally, the performance of trained networks is limited by the input image resolution (typically 512×512), which is too small for clinical applications where image resolution, in general, is larger than 2000×1000 . Furthermore, these methods lack interpretability and may introduce unexpected artifacts or elimination of important retinal structures (Section 5.1), which can be detrimental to clinical applications. Thus, there is still a long way to go in both technical and ethical aspects of learning-based retinal image enhancement methods [30].

Overall, the retinal image enhancement/restoration communities seek contributions that combine the strengths of both non-learning-based and learning-based methods. Such methods should have wide generalization, high interpretability, and efficiently enhance the contrast of retinal images while preserving naturalness.

2.3 Our contributions

Our proposed InQue is a semi-learning-based, fully unsupervised retinal image enhancement algorithm. It doesn't need any synthetic images in any form, the only input is a single retinal image. InQue learns the model parameter based on a single input and returns an enhanced retinal image, maintaining color preservation. Fig. 5-1 shows the output of InQue for retinal images under different degeneration situations including cataracts, low light, and uneven illuminations. InQue has the following contributions:

- 1) A color implantation module that standardizes the color distribution of the input images while retaining detailed information of the retinal images. This stabilizes the output and preserves the naturalness of the images.

- 2) **Unifying the retinal image enhancement process into two steps:** illumination correction and dehazing. We propose an intensity-based cost function based on the observation that illumination correction increases the average intensity

of the image, while dehazing decreases it. This allows InQue to learn model parameters based on single images and tune the output performance adaptively.

3) **Our proposed InQue is not limited by the resolution of input images and can be applied to an image of arbitrary resolutions from 512^2 to 4096^2 or larger.** High computational efficiency is maintained.

4) The most significant contribution of InQue is its improvement of retinal image quality and contrast, benefiting clinical applications. **The algorithm does not introduce any artifacts and is interpretable by both mathematicians and clinical specialists.**

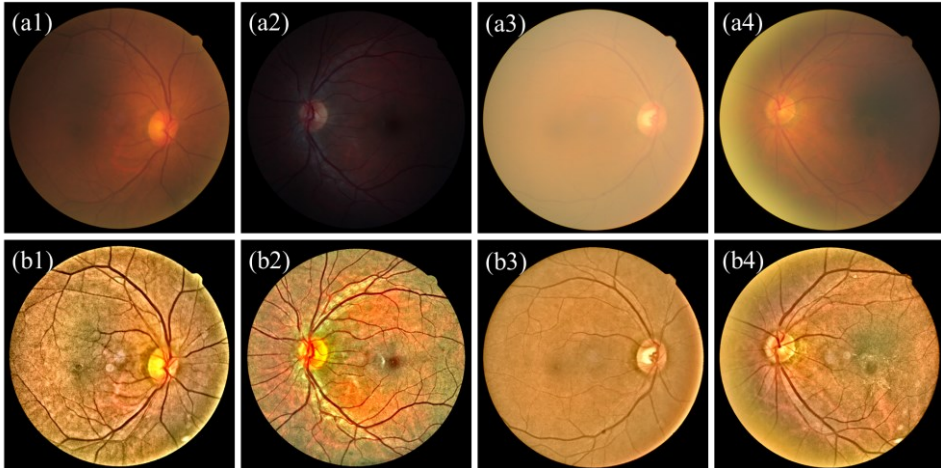


Fig. 5-1. Demonstration of InQue output on different degraded retinal images. First row: input images. Second row: output of InQue. (a1) low-light & cataract. (a2) low-light. (a3) cataract. (a4) uneven-light & cataract.

3. Quenching the intensity of retinal images

The image formation model for fundus photographs can be described using the double-pass fundus reflection model [31], which is given by:

$$S_c(r) = L(r) \circ [T^2(r) \circ O_c(r) + 1 - T(r)], \quad c \in \{R, G, B\}, \quad (5-1)$$

where \mathcal{S}_c is assumed to be the captured image on the camera the size of \mathcal{S}_c is assumed to be of M by N pixels. \mathcal{O}_c is the un-degraded retinal images of which the quality will be decreased by the potential problematic illumination, \mathbf{L} and intraocular scattering such as cataract layers, where $\mathbf{T} \in [0, 1]$ is the transmission matrix of the scattering layer. Both \mathbf{L} and \mathbf{T} remain unchanged between different color channels. Ignoring the absorption of the scattering layer, $1 - \mathbf{T}(\mathbf{r})$ is the backward scattering component that can lead to the haze effect in fundus imaging.

In a more practical situation, the color distribution of captured retinal images varies from image to image due to unexpected aspects such as lens aging, intraocular reflection, and environmental background intensity. For example, as shown in the first row of Fig. 5-1, all four retinal images have distinguished color distributions, where the first one is reddish, and the last one is rather yellowish. To model this, we introduce the background color distortion matrix, \mathcal{C}_c , in Eq. (1) yielding

$$\tilde{\mathcal{S}}_c = \tilde{\mathcal{C}}_c \circ \mathcal{S}_c, \quad (5-2)$$

Here $\tilde{\mathcal{S}}_c$ is the final observation of retinal imaging, and our goal is to restore \mathcal{O} from known only of $\tilde{\mathcal{S}}_c$.

This task is highly ill-posed as $\tilde{\mathcal{C}}_c$, \mathbf{L} , and \mathbf{T} are all unknown. We therefore separated this task into four sequential modules which are (1) Color implant, (2) Luma Heater, (3) Luma Cooler and (4) Luma Rectifier. The Color implant module aims to correct $\tilde{\mathcal{C}}_c$, while the Luma Heater and Luma Cooler modules are for blind estimating \mathbf{L} and \mathbf{T} . With that, we are able to use Eq. (5-1) and (5-2) to restore the \mathcal{O} , and finally, the Luma Rectifier remaps the obtained \mathcal{O} to the noise-reduced pattern.

Fig. 5-2 shows a flow chart of our proposed InQue. As it will be mentioned in Section 3.2, the intensity of Luma will be first increased and then decreased, which is similar to a ‘quenching process’ if the gray-scaled Y channel is regarded as a heated plane, we,

therefore, refer our proposed method to the intensity quenching (InQue) method. In the following subsection, we will present detailed calculations and learning processes for each module.

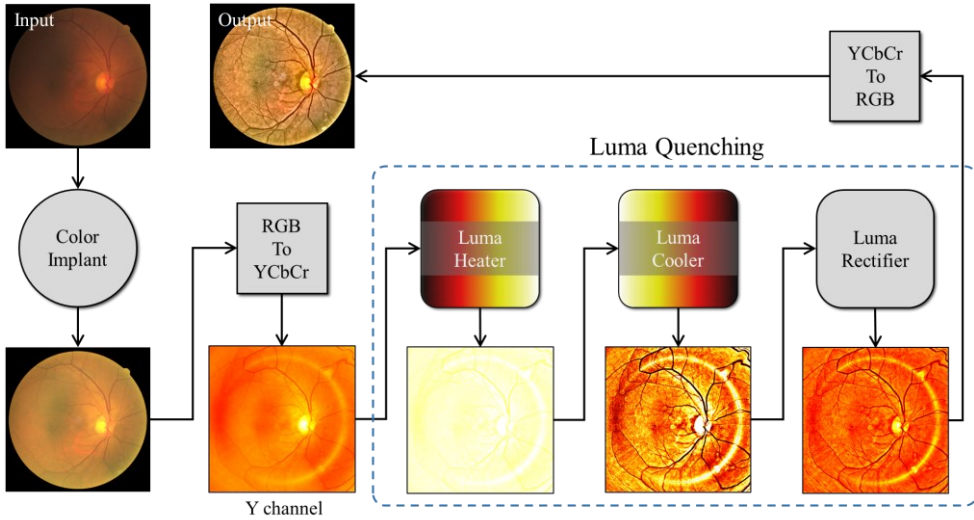


Fig. 5-2. The flowchart of Luma Quenching.

Before entering the retinal image blind deconvolution, an image background reflective padding process was applied in order to remove the boundary effect during the convolution process [32]. Further, throughout the manuscript, pixel values were normalized within the range of [0, 1].

3.1 Color implant

The first “color implant” module aims to correct the impact of the color distortion matrix, C_c . It first removes the original background color matrix \tilde{C}_c and implants a guiding background color C_c from a more natural color distribution to the original input images. Considering the background can be the low-frequency component of the image, we use Eq. (5-3) to achieve the color implant.

$$S_c = \frac{\tilde{S}_c}{\tilde{C}_c} \circ (C_c \otimes g) \circ M, \quad \tilde{C}_d = \tilde{S}_c \otimes g, \quad (5-3)$$

where C_c is the guiding background color to be implanted, which can be sampled from known retinal images of good illumination quality. In this research, we use a uniform color of $(R, G, B) = (0.80, 0.54, 0.24)$ be C_c . Here M is the binary region of interest (ROI), and g is a Gaussian kernel to extract the background component from the input RGB image. The size g is $\min(M, N)/3$ pixels to ensure no high-frequency components in the background. \otimes is the 2D convolution. Zero-padding is used to diminish boundary effects during the convolution.

5

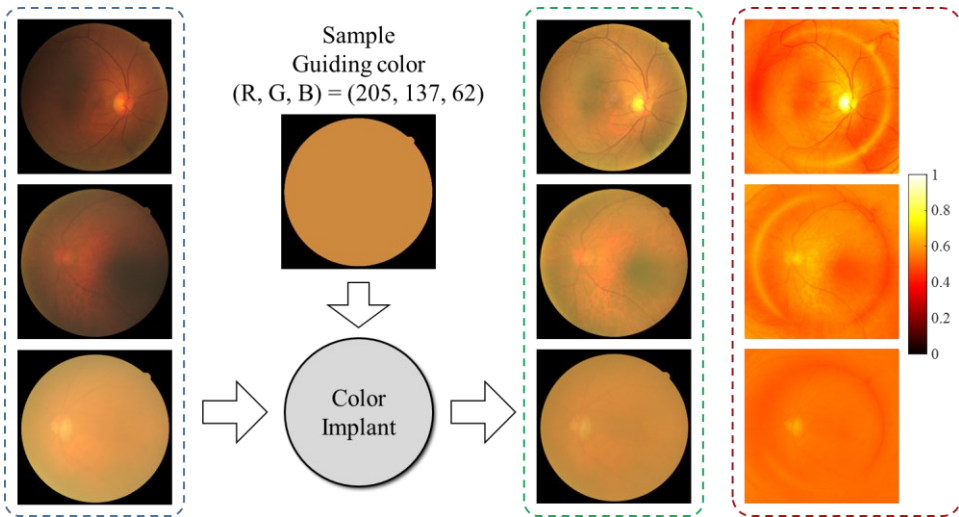


Fig. 5-3. Color implant module that converts the background color of input images (first left column) to a more natural background color distribution. The images in the last column are the Luma components (Y channel in YCbCr color space)

Output images are shown in the third column of Fig. 5-3, where all images now have a similar color distribution. The image formation model of the output images is then given by Eq. (5-1). Let Ψ be the matrix that converts RGB color space to YPbPr color space, then Ψ is written as

$$\Psi = \begin{pmatrix} K_R & K_G & 1 - K_R - K_G \\ -\frac{K_R}{2(K_R + K_G)} & -\frac{K_G}{2(K_R + K_G)} & \frac{1}{2} \\ \frac{1}{2} & -\frac{K_G}{2(1 - K_R)} & -\frac{1 - K_R - K_G}{2(1 - K_R)} \end{pmatrix}, \quad (5-4)$$

and $\Psi(1,1,1)^T = (1,0,0)$. According to ITU-R BT.601 standard, $K_R = 0.299$ and $K_G = 0.587$. Applying Eq. (5-4) to both sides of Eq. (5-1), we obtained the image formation model in the YCbCr color space, which reads

$$\begin{pmatrix} \mathbf{S}_Y \\ \mathbf{S}_{P_b} \\ \mathbf{S}_{P_r} \end{pmatrix} = \mathbf{L} \cdot \mathbf{T}^2 \cdot \begin{pmatrix} \mathbf{O}_Y \\ \mathbf{O}_{P_b} \\ \mathbf{O}_{P_r} \end{pmatrix} + \begin{pmatrix} \mathbf{1} - \mathbf{T} \\ 0 \\ 0 \end{pmatrix}. \quad (5-5)$$

According to Eq. (5-5), the hazing effects only affect the Y channel of the retinal image as the Y channel has the haze term $\mathbf{L} \cdot (\mathbf{1} - \mathbf{T})$ which denotes the backscattering of light from intraocular scatters. For C_b and C_r channels, the existence of \mathbf{T} works similar to \mathbf{L} which only decreases the intensity of corresponding color channels of the haze-free image.

3.2 Heater and cooler

In the quenching process, both the Luma heater and Luma cooler are applied to the Y-channel (\mathbf{S}_Y) of Eq. (5-5). The Luma heater corrects the impact of \mathbf{L} which is an illumination correction task, while the Luma cooler corrects the impact of \mathbf{T} which is a dehazing task. Notably, by employing some simple algebra, the illumination correction task can be converted into a special case of dehazing task [33, 34]. Let $\mathbf{Q}_Y = \mathbf{T}^2 \circ \mathbf{O}_Y + \mathbf{1} - \mathbf{T}$, $\tilde{\mathbf{Q}}_Y = \mathbf{1} - \mathbf{Q}_Y$, and $\tilde{\mathbf{S}}_Y = \mathbf{1} - \mathbf{S}_Y$, we reformulate the Eq. (5-1) for the luma component and we now have two dehazing tasks of different purposes which are

$$\tilde{\mathcal{S}}_Y = \mathbf{L} \circ \tilde{\mathcal{Q}}_Y + 1 - \mathbf{L}, \quad (5-6)$$

and

$$\mathcal{Q}_Y = \mathbf{T}^2 \circ \mathcal{O}_Y + 1 - \mathbf{T}, \quad (5-7)$$

The task for the Luma Heater module is to solve \mathcal{Q}_Y with known of \mathcal{S}_Y based on Eq. (5-6), while the task for the Luma Cooler is to solve \mathcal{O}_Y with known of \mathcal{Q}_Y based on Eq. (5-7). Since both Eq. (5-6) and Eq. (5-7) are of similar mathematical form, we can put them into a similar mathematical framework.

Luma Heater: For Eq. (5-6), we consider that the illumination component \mathbf{L} is spatially slow varying according to the Retinex theory, then $1 - \mathbf{L}$ should also be slow spatially varying. We, therefore, assume that $1 - \mathbf{L}$ can be estimated from the smoothed version of $\tilde{\mathcal{S}}_Y$. Here, we use the linear superposition of multi-level smoothed $\tilde{\mathcal{S}}_Y$ and \mathbf{L} can be expressed as

$$\mathbf{L} = 1 - \sigma^h \cdot \sum_{l=1}^L w_l^h \cdot \text{Smooth}_l(\tilde{\mathcal{S}}_Y), \quad (5-8)$$

Here σ^h and w_l^h are scalar parameters to weigh the global and individual amplitude of smoothed layer. And $\text{Smooth}_l(\mathbf{X})$ is an image smoothing function that smooths the input image \mathbf{X} . Based on Eq. (5-6) and Eq. (5-8), we can design a cost function that guides the updating of these parameters.

Considering that the illumination correction process increases the intensity distribution of \mathcal{Q}_Y , the cost function is designed based on the ratio between the number of over-exposed pixels between the total number of pixels in the ROI which is

$$\mathcal{L}_{\text{Heater}}(\sigma^h, w_l^h) = \frac{1}{R_{\text{Heater}}} \left| \frac{\#\{\mathbf{r} \in \{\mathcal{Q}_Y(\mathbf{r}) > 1\} \cap \mathbf{r} \in \mathbf{M}\}}{\#\{\mathbf{r} \in \mathbf{M}\}} - R_{\text{Heater}} \right|, \quad \mathbb{R}^{M \times N} \rightarrow \mathbb{R}. \quad (5-9)$$

Here, $\mathbf{Q}_Y = \mathbf{S}_Y / \mathbf{L}$ according to Eq. (5-6). # is the counting function that counts the number of all pixels \mathbf{r} that meet the given conditions. $\mathbf{r} \in \{\mathbf{Q}_Y(\mathbf{r}) > 1\}$ denotes the over-exposed pixels where the gray-values are larger than 1. $\#(\mathbf{r} \in \mathbf{M})$ denotes the total number of pixels in the ROI which is a constant and can be calculated once the input images are given. $R_{Heater} \in [0, 1]$ is the target ratio. For example, when $R_{Heater} = 0.5$, minimizing Eq. (5-9) means that there will be 50% percent of pixels in the ROI that are over-exposed.

Since the counting function has the same behavior as the L_0 -norm, Eq. (5-9) is non-differentiable. To make the parameters learnable, we use the sigmoid function to approximate the counting function and Eq. (5-9) can be approximately calculated as

$$\mathcal{L}_{Heater}(\sigma^h, w_l^h) = \frac{1}{R_{Heater}} \left| \frac{\|\text{Sigmoid}[\beta(\mathbf{Q}_Y - 1)]\|_1}{\#(\mathbf{r} \in \mathbf{M})} - R_{Heater} \right|, \quad \mathbb{R}^{M \times N} \rightarrow \mathbb{R}. \quad (5-10)$$

The shifted sigmoid function in Eq. (5-10), approaches to 0 when the pixel value of x is less than 1, while it approaches 1 when the pixel value is larger than 1. Scalar parameter β controls the speed of the stimulation process. When $\beta \rightarrow \infty$, the sigmoid function becomes a unit step function beginning at 1. By giving an appropriate value of β , the sigmoid function in the numerator of Eq. (5-10) can mimic the behavior of the counting function. Eq. (5-10) is now differentiable with respect to σ^{heater} and w_l^{heater} , and with the given target ratio R_{Heater} the model parameters together with the \mathbf{Q}_Y for the Luma Heater can be adaptively determined (learned).

Luma Cooler: Similar to the Luma Heater, we rewrite the \mathbf{T} as

$$\mathbf{T} = 1 - \sigma^c \cdot \sum_{l=1}^L w_l^c \cdot \text{Smooth}_l(\mathbf{Q}_Y), \quad (5-11)$$

given that \mathbf{T} is also spatially slow varying. Since the dehazing process always decreases the intensity of the correction process and increases the intensity distribution of \mathbf{O}_Y ,

we designed a cost function using the ratio between the number of under-exposed pixels between the total number of pixels in the ROI as

$$\mathcal{L}_{Cooler}(\sigma^c, w_l^c) = \frac{1}{R_{Cooler}} \left| \frac{\#\{\mathbf{r} \in \{\mathbf{O}_Y(\mathbf{r}) < 0\} \cap \mathbf{r} \in \mathbf{M}\}}{\#\{\mathbf{r} \in \mathbf{M}\}} - R_{Cooler} \right|, \quad \mathbb{R}^{M \times N} \rightarrow \mathbb{R}. \quad (5-12)$$

Here, $\mathbf{O}_Y = (\mathbf{Q}_Y - 1) / \mathbf{L}^2 + \min(1 / \mathbf{T}, 1)$ according to Eq. (5-7). Again, we used the sigmoid function to approximate the counting function and Eq. (5-12) can be approximately calculated as

$$\mathcal{L}_{Cooler}(\sigma^c, w_l^c) = \frac{1}{R_{Cooler}} \left| \frac{\|\text{Sigmoid}(-\beta \cdot \mathbf{O}_Y)\|_1}{\#\{\mathbf{r} \in \mathbf{M}\}} - R_{Cooler} \right|, \quad \mathbb{R}^{M \times N} \rightarrow \mathbb{R}. \quad (5-13)$$

Model parameters together with the \mathbf{O}_Y for the Luma Cooler can be adaptively learned. The pixels in output \mathbf{O}_Y that are less than -1 are clipped to -1.

In this research, we chose the Gaussian filter to be the smooth function. The Gaussian kernel width is $l \times \sqrt{MN/100}$ pixels for the l -th level. A total of $L = 5$ levels. $\beta = 30$, $R_{Heater} = 0.2$ and $R_{Cooler} = 0.05$. We used auto-differentiation in MATLAB 2022b. Parameters were updated using the Adam optimizer with 0.9, and 0.999 decay rates. The initial learning rate was 0.01 and decreased to 0.001 if the value of the cost function was less than 50%. Following Fig. 5-3 (c), Fig. 5-4 (a) shows the Luma heater output where the intensity of Luma components are at similar levels. Then the Luma cooler decreases the intensity of these Luma components. During this cooling process, the contrast of detailed structure of retinal images like blood vessels gradually becomes prominent as shown in Fig. 5-4 (b).

Luma from the Cooler is not the final output of our InQue since they are not normalized. As shown in Fig. 5-4 (b), the Luma component for the central image is brighter than the other two images, while the intensity in the optics disk is also brighter

than in other regions. To suppress these situations, we introduce the Luma Rectifier to refine the Luma components.

3.3 Luma rectifier

With the O_Y from Luma cooler, the output of Luma rectifier is calculated according to

$$O_{Y,fin} = \frac{I_0}{\frac{1}{\#r \in M} \sum_{\forall r \in M} \bar{O}_Y(r)} \cdot \bar{O}_Y, \quad (5-14)$$

where $O_{Y,fin}$ is the final output of Luma component of InQue. I_0 is a scalar parameter, and \bar{P} is calculated according to

$$\left\{ \begin{array}{l} P = \sum_{l=1}^L O_Y - \text{Smooth}_l(O_Y) \\ \bar{P} = \left[\max_{r \in M}(O_Y) - \min_{r \in M}(O_Y) \right] \cdot \frac{P - \min_{r \in M}(P)}{\max_{r \in M}(P) - \min_{r \in M}(P)} + \min_{r \in M}(O_Y) \\ \bar{O}_Y = \text{Denoising}(\arctan \bar{P}), \end{array} \right. \quad (5-15)$$

The first step in Eq. (5-15) can be regarded as a multi-scaled Retinex process that rectifies the intensity distribution of input O_Y so that the intensity for the optical disk area won't be too much brighter than other areas of in O_Y . Similar to the Cooler/Heater module, the smooth function here is a Gaussian filter but the width of the Gaussian kernel is $3l \times \sqrt{MN/100}$. The second step in Eq. (5-15) shifts the dynamic range of corrected images, P , to the dynamic range of the input images using max-min methods. The arctangent function in the third step of Eq. (5-15) is a stimulated function that normalizes the gray-scale distribution between -1 to 1, so that the output Y-channel \bar{O}_Y will maintain the contrast while won't cause severe color distortion. The 'Denoising' in

Eq. (5-15) means an arbitrary denoising process for given inputs, and in this research, we chose the self-guided filter with a window size of 2 pixels.

After \bar{P} is obtained, Eq. (5-14) remaps the average value of \bar{P} to a target value I_0 , so that the average intensity distribution of $O_{Y,fin}$ becomes similar even for different input images as shown in Fig. 5-4 (c) where $I_0 = 0.52$ for example. Combining the new Luma component $O_{Y,fin}$ with the original S_{Cb} and S_{Cr} calculated from Eq. (5-5), we transfer them back to the RGB color space, and the final RGB retinal images are shown in Fig. 5-4 (d). Ablation studies on each module are detailed in Section 5.

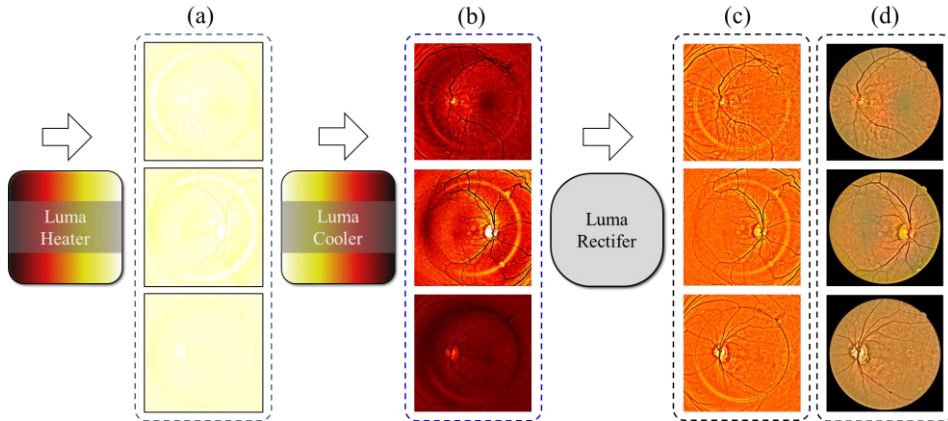


Fig. 5-4. Quenching processing on the intensity pattern of the retinal image. (a) Intensity pattern after heating. (b) after cooling. (c) Intensity pattern after the rectifying. (d) final outputs of InQue.

4. Experimental results

In this research, we chose $L = 5$, $\beta = 30$, $R_{Heater} = 0.05$ and $R_{Cooler} = 0.02$. We used auto-differentiation in MATLAB 2022b. The parameters were updated using the Adam optimizer with 0.9, and 0.999 decay rates. The initial learning rate was 0.01 and decreased to 0.001 if the value of the cost function was less than 50%. $I_0 = 0.52$ for the Luma Rectifier as this value is tested to be appropriate to produce enhanced retinal images with natural visual perspectives [35]. We tested InQue on 6 public datasets

including Messidor- (1200 images) [36], CATARACT - (100 images) [37], Normal- (300 images) [37], DiaRet DB - (219 images) [38, 39], IDRiD- (516 images) [40] and E-ophtMA- (381 images) [1] datasets. In the main text, we show only the experimental results for the CATARACT dataset, more results can be found in **Supplementary Note 5-1**.

InQue is compared against four other SOTA deep-learning-based methods, ArcNet [25], ScrNet [41], DANet [42] and ArSrNet [26]. All images were down-sampled into 512 by 512 pixels in order to fit the input size of deep-learning methods. For ArcNet, ScrNet, and DANet, we used the author-published pre-trained models. While for ArSrNet, we asked the author to process the images on their local machine. Outputs of InQue were also down-sampled into 512 by 512 for fair comparisons. In addition, the InQue is also compared against SOTA non-deep learning methods, namely the Luminosity and Contrast Adjustment (LCA) [14], Decomposition and Visual Adaptation (DVA) [17], the detail-richest-channel method (DRC) [23], the low-pass filtering and α -rooting (LPAR) [16], and the Double-pass Fundus Reflection model (DPFR) [31]. Results can be found in **Supplementary Note 5-2**.

4.1 Visual assessments

Fig. 5-5 shows the results of four groups of enhanced images. All five methods effectively restore color in the images, resulting in more uniform background colors. Additionally, they all improve the visual quality of cataractous retinal images by making retinal structures such as blood vessels and optical disks visible, which were previously obscured by the cataract effect. Examples of these improvements are shown in Fig. 5-5 (b1) to (b6), and Fig. 5-5 (c1) to (c6). Among the methods, InQue output provides the most prominent contrast, as demonstrated in Fig. 5-5 (b6) and Fig. 5-5 (d6) compared to Fig. 5-5 (b2) to (b5) or Fig. 5-5 (d2) to (d5). The outputs of SrcNet as shown in Fig. 5-5 (c3) has black error blocks due to the lack of generalization of the network.

Partially enlarged images for the region in the white boxes are shown in Fig. 5-6 for detailed observation. For the deep-learning-based methods ScrNet, DANet, and ArSrNet, the outputs images have chessboard effects as shown in Fig. 5-6 (b2), (a4), and (c5).

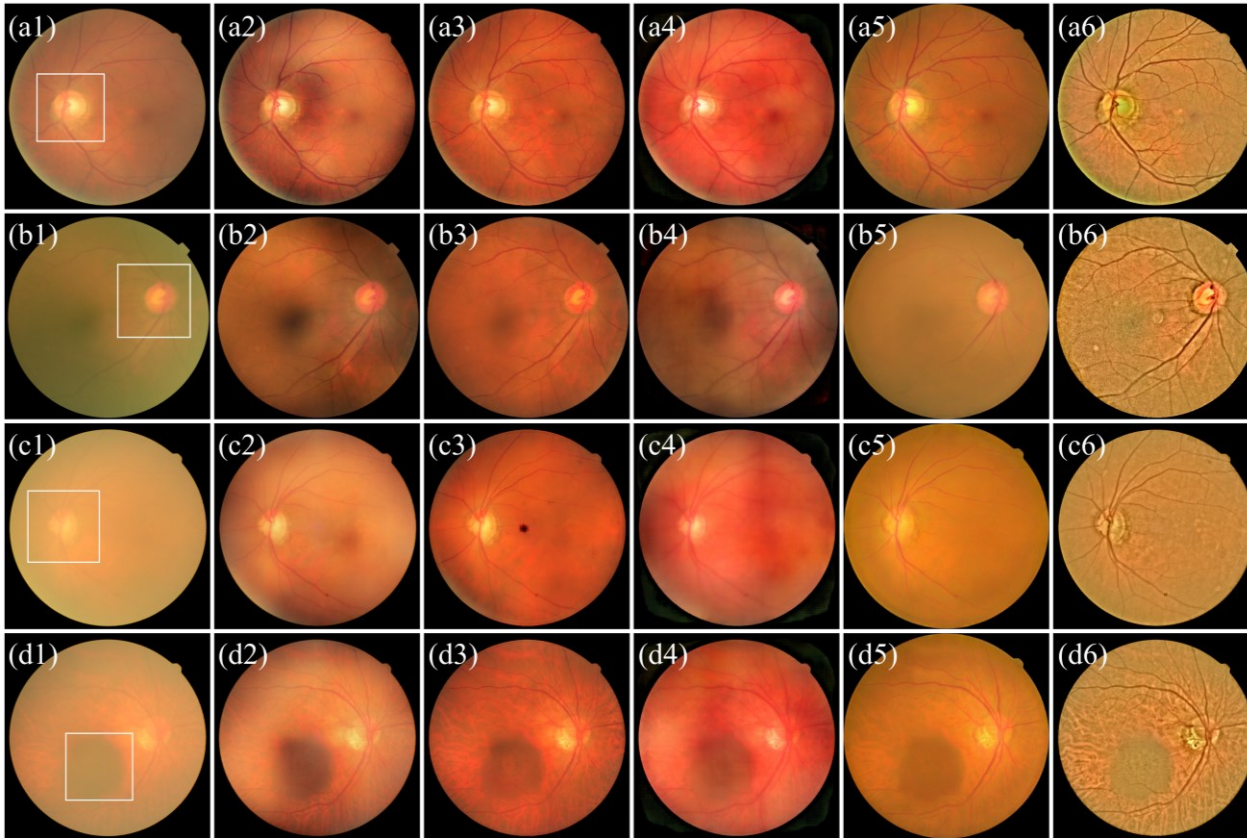


Fig. 5-5. Comparison results on cataractous retinal image enhancement. (a1) to (d1) are raw images. (a2) to (d2) are outputs of ArcNet. (a3) to (d3) are outputs of ScrNet. (a4) to (d4) are outputs of DANet. (a5) to (d5) are outputs of ArSrNet, and the final column is the output of InQue.

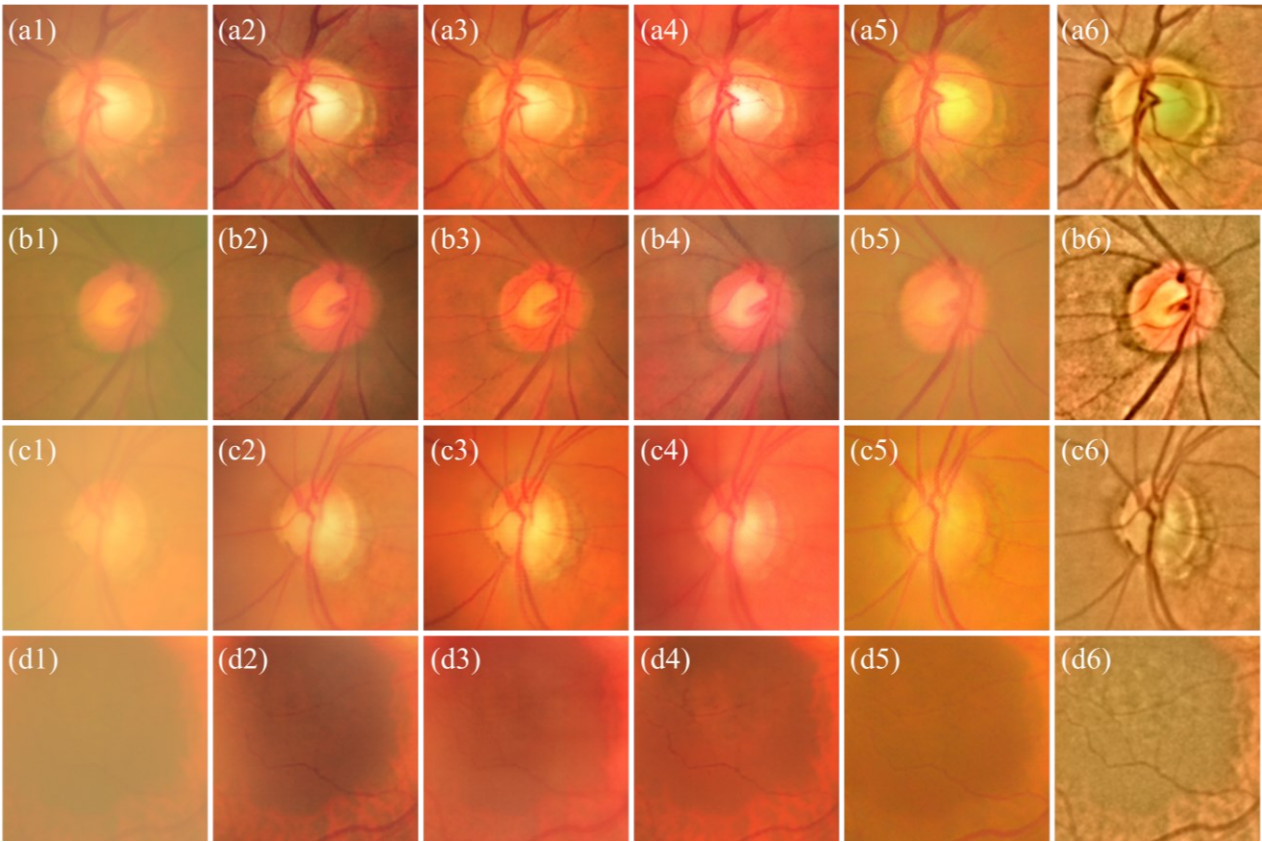


Fig. 5-6. Partially enlarged images for the region in the white boxes in Fig. 5-5. (a1) to (d1) are raw images. (a2) to (d2) are outputs of ArcNet. (a3) to (d3) are outputs of ScrNet. (a4) to (d4) are outputs of DANet. (a5) to (d5) are outputs of ArSrNet, and the final column is the output of InQue

This effect is probably due to the transpose convolution during the generation of images. In Fig. 5-6 (a4) and Fig. 5-6 (c5), there are red spot-like artifacts in the optical disk area, which can be wrongly regarded as hemorrhages leading to misdiagnosis. Such artifacts are common problems in deep-learning-based methods as reported in [25]. Although the ArSrNet reduces the artifacts in the generated images, its dehazing ability as shown in Fig. 5-6 (b5) and Fig. 5-6 (c5) is weaker than that of InQue as shown in the last column of Fig. 5-6.

The dehazing effect of InQue is significant as the branches of small blood vessels can be clearly observed as shown in Fig. 5-6 (b6) to (c6) and Fig. 5-6 (d6). Meanwhile, InQue has no artifacts, preserving the data fidelity of the restored images. Every step and corresponding intermediate products of InQue is explicable, which is important for medical images and their applications.

4.2 Objective assessments

We performed objective assessments on each image in Fig. 5-5 and all images in the cataract dataset. We used non-references-based evaluation matrixes including image entropy (IE), image multi-scale contrast (C_{RAMM}) [43], image sharpness measurement (UISM), image contrast measurement (UIConM) [44] and image fog density estimation (FADE) [45]. The references-based evaluations such PSNR and SSIM are not used because they are not related to human visual prospection [46] and lack fairness [47], since they require real ground-truth images. Such ground truth of medical images is practically not well-defined (some medical images that are regarded as being of good quality when they are collected can be still enhanced for better visual quality).

Results for images in Fig. 5-5 are listed in Tab. 5-1, where InQue obtains better results than the other methods in all evaluation matrixes. Particularly, InQue achieves a lower FADE score, indicating promising dehazing results on cataract retinal images. For example, the FADE score of the image in Fig. 5-5 (b1) is reduced from 0.4485 to 0.1751 after being enhanced by InQue. InQue also significantly improves image contrast, as indicated by its high scores in UISM, UIConM and C_{RAMM} . For the image in

Fig. 5-5 (c1), in UISM the contrast is increased by more than 3 times compared to the original image.

Tab. 5-1. Quantitative assessment of the enhanced images among Fig. 5-5 (a1) to (d1) with five metrics.

Raw image	Matrixes	Raw	Methods				
			ArcNet	ScrNet	DANet	ArSrNet	InQue
Fig. 5-5 (a1)	UIConM ↑	0.0795	0.0967	0.0846	0.0980	0.0727	0.1648
	UISM ↑	1.1856	1.4435	1.5265	2.3394	1.9078	4.4290
	IE ↑	5.8863	6.2905	5.9440	6.7442	6.3616	6.0533
	C _{RAMM} ↑	1.9977	2.4131	2.1578	2.3595	1.9387	3.5313
	FADE ↓	0.3941	0.2693	0.2725	0.2421	0.2841	0.2028
Fig. 5-5 (b1)	UIConM ↑	0.0637	0.0614	0.0659	0.0692	0.0658	0.1858
	UISM ↑	1.3908	1.2140	1.5202	2.4293	1.6773	4.4512
	IE ↑	6.1030	6.3725	6.0189	6.6719	6.4123	6.3375
	C _{RAMM} ↑	1.7081	1.8300	1.8202	1.8958	1.7027	3.8016
	FADE ↓	0.4485	0.3005	0.3151	0.2656	0.3102	0.1751
Fig. 5-5 (c1)	UIConM ↑	0.0944	0.0825	0.0743	0.0860	0.0757	0.1192
	UISM ↑	1.1528	0.9800	1.5461	2.5178	1.4918	3.8217
	IE ↑	5.6353	6.2942	5.8789	6.9150	6.2938	5.6565
	C _{RAMM} ↑	2.0947	2.0407	1.8857	2.1733	1.8556	2.7400
	FADE ↓	0.5622	0.3639	0.3259	0.2891	0.3070	0.2531
Fig. 5-5 (d1)	UIConM ↑	0.0887	0.0864	0.0750	0.0869	0.0752	0.1469
	UISM ↑	1.1570	1.0753	1.4116	2.4253	1.5175	4.0842
	IE ↑	5.7905	6.3311	5.8247	6.8460	6.4521	6.1191
	C _{RAMM} ↑	2.0437	2.1569	2.0449	2.2390	1.9220	3.4423
	FADE ↓	0.4638	0.3190	0.2893	0.2642	0.2916	0.1973

Statistical results are listed in Tab. 5-2 for all 100 images in the CATARACT dataset including both averages and standard deviations. According to Tab. 5-2, the average scores of InQue outperformed the other 3 deep-learning-based methods, while the small standard deviation values imply the robustness of InQue among different retinal

images. More experiment results on the other 5 published dataset, including DiaRet dataset, E-optha_MA dataset, IDRiD dataset, Messidor dataset and NORMAL dataset, can be found in the **Supplementary Note 5-1**.

Tab. 5-2. Averages and standard deviations of enhancement results from both public and proprietary cataractous retinal image datasets.

Databases mean \pm std.	Matrixes	Raw	Methods				
			ArcNet	ScrNet	DANet	ArSrNet	InQue
Cataract (100 images)	UIConM (\uparrow)	0.0689	0.0774	0.0745	0.0856	0.0679	0.1758
		± 0.0163	± 0.0145	± 0.0059	± 0.0073	± 0.0082	± 0.0293
	UISM (\uparrow)	1.2874	1.3390	1.6026	2.4310	1.7656	4.3753
		± 0.1971	± 0.2852	± 0.2792	± 0.1816	± 0.2261	± 0.2686
	IE (\uparrow)	5.8933	6.2380	5.9048	6.7477	6.3738	6.1375
		± 0.2622	± 0.1728	± 0.1197	± 0.1073	± 0.1798	± 0.2310
	CRAMM (\uparrow)	1.8005	2.0926	1.9770	2.2044	1.8050	3.7292
		± 0.2608	± 0.2672	± 0.1364	± 0.1351	± 0.1518	± 0.5667
	FADE (\downarrow)	0.4354	0.2969	0.2983	0.2524	0.3057	0.1911
		± 0.0796	± 0.0363	± 0.0299	± 0.0196	± 0.0296	± 0.0253

Tab. 5-3. Averages execution time of InQue for different size of images.

Image size	512 ²	1024 ²	2048 ²	4096 ²
Average execution duration	2.85	9.35	30.23	107.38

4.3 Execution time

To demonstrate the processing efficiency of InQue, we measured its average execution time on images of varying sizes ranging from 512² to 4096² pixels, using laptop DELL precision 7550 (Intel Xeon W10855M CPU 2.80GHz, 32GB RAM). As listed in Tab. 5-3, for images of 512² pixels, InQue exhibited an average processing time of

fewer than 0.3 seconds, which is comparable to that of learning-based methods. However, for larger images such as those with 2048^2 or 4096^2 pixels, the execution time increased. Nevertheless, for images with sizes less than 2048^2 pixels, InQue's execution speed was less than 30 seconds.

5. Potential clinical applications

5.1 Retinopathy diagnosis on problematic retinal images

As InQue restores image intensity and improves the contrast, it can benefit clinical diagnosis. To illustrate this, we focussed on the restoration results in areas that indicate retinopathy, as shown in Fig. 5-7 (b). The raw image in Fig. 5-7 (a) suffers from insufficient illumination and low contrast, which makes it difficult to diagnose retinopathy. In contrast, InQue-enhanced images provide greater clarity and detail, making it easier to identify abnormalities.

We compared InQue's results with those of deep-learning-based methods in Fig. 5-7 (c2) to (c5) for the purple box area and Fig. 5-7 (d1) to (d5) for the green box area. As seen in Fig. 5-7 (b), the purple box area contains hard exudates and red small dots. InQue improved the contrast of this area, making the white spot of hard exudates more obvious in Fig. 5-7 (c5) than in Fig. 5-7 (c1) (the raw image). On the other hand, deep-learning methods such as ArcNet and ScrNet seem to decrease the intensity of hard exudates as shown in Fig. 5-7 (c2) and (c3), respectively. Additionally, InQue enhanced the area of hemorrhages, as shown in Fig. 5-7 (d5), which is indicated by the green box in Fig. 5-7 (b). In contrast, the raw images in Fig. 5-7 (d1) do not provide clear visualization of the hemorrhages.

Overall, InQue's ability to increase the clarity of retinal images may facilitate the detection of retinopathy without leading to the loss of information or additional artifact structures, as observed in the comparison with deep-learning-based methods.

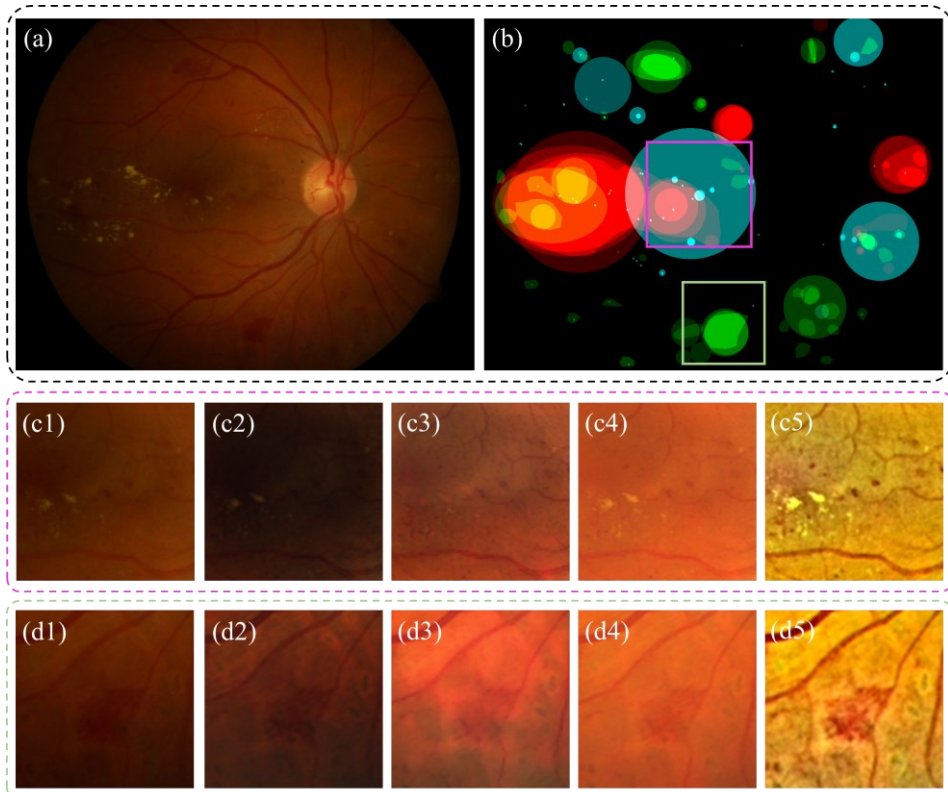


Fig. 5-7, Enhancement of retinopathy. (a) Raw image. (b) Labels of retinopathy areas. Red: Hard exudates; Green: Hemorrhages; Cyan: red small dots. (c1) to (c5) are zoomed-in pictures for regions in the purple box. They are raw image, results of ArcNet, results of ScrNet, results of DANet, and results of InQue, respectively. (d1) to (d5) are zoomed-in pictures for regions in the green box.

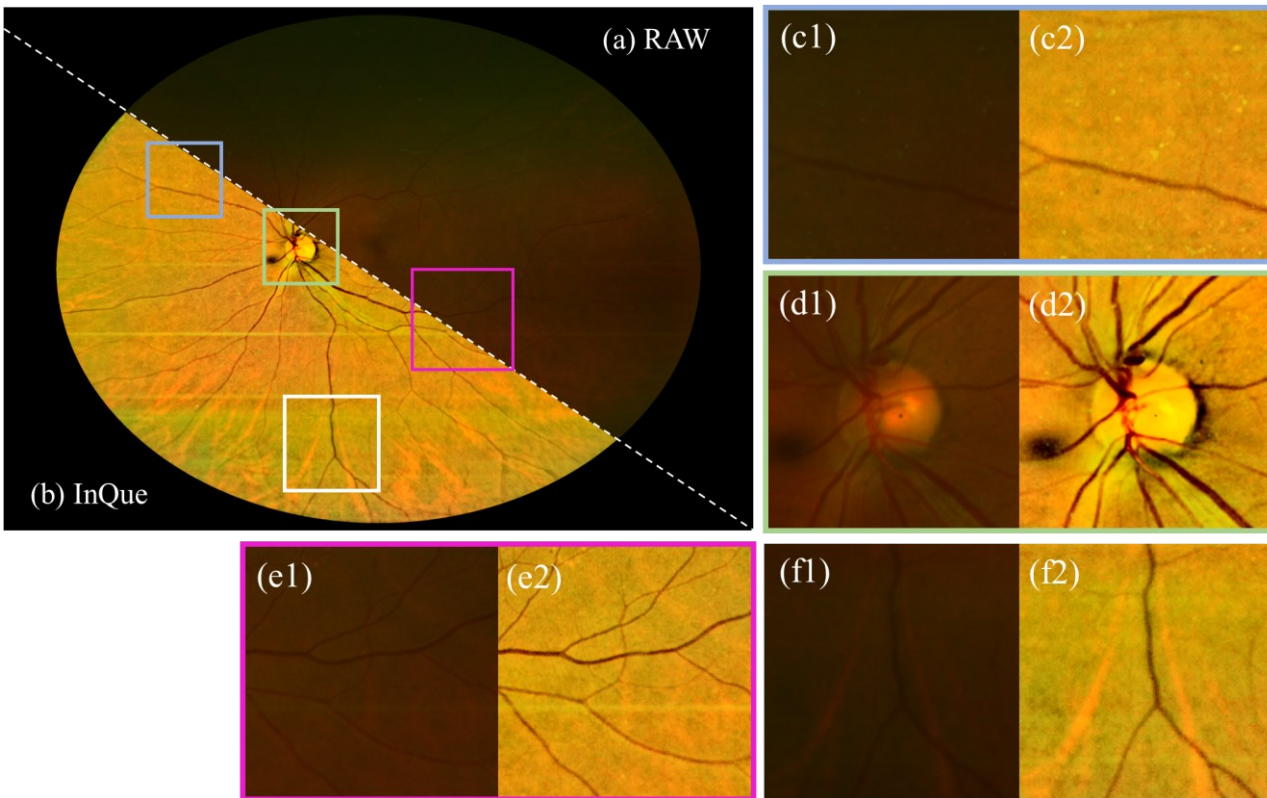


Fig. 5-8. InQue enhancement for UWF retinal image. (a) raw images. (b) enhanced. (c1) and (c2) are zoomed-in images for the blue box. (d1) and (d2) are zoomed-in for the green box. (e1) and (e2) are zoomed in for the purple box, and (f1) and (f2) are zoomed-in of the white box.

5.2 Enhancement of ultra-widefield retinal images

Ultra-wide field (UWF) imaging system allows capture of 200 degrees of the retina (approximately 82% of retinal surface area) in a single shot. It provides non-contact, high-resolution images of retina for clinicians to analyze retinal disorders [48, 49]. Imaging using UWF system can also suffer from illumination problems, and haze effects due to imperfect photographing conditions. Here we demonstrate the application of InQue in enhancing UWF images.

Fig. 5-8 shows the enhancement for UWF retinal images. The raw image is shown in Fig. 5-8 (a), where insufficient illumination and haze effects are prominent. The enhanced one is shown in Fig. 5-8 (b). while the retinal structures including the optical disk and blood vessels are zoomed in Fig. 5-8 (c1) to (f2). Accordingly, the image clarity of the enhanced image is significantly improved, and blood vessels can be clearly observed.

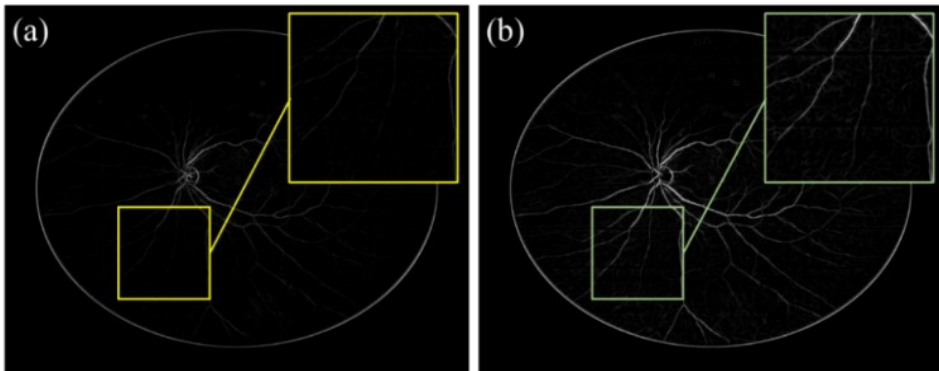


Fig. 5-9. Match-filtering of UWF image for blood vessel detection.

5.3 Retinal vascular tracking

With the enhanced image in Fig. 5-8 (b), we perform the blood vessel tracking using match-filters. Results are shown in Fig. 5-9 (a) for tracking results of the raw image and

in Fig. 5-9 (b) for the enhanced one. By increasing the contrast of the image, InQue benefits the tracking results, as shown in Fig. 5-9 (b) where small blood branches can be correctly tracked by the filters.

6. Ablation study and discussions

6.1 On the Color implant

In this section, we conduct an ablation study on the color implantation module. This module replaces the original background color with a sampled color from normal retinal images, which can achieve color preservation for the enhanced images, making them have more natural color distribution.

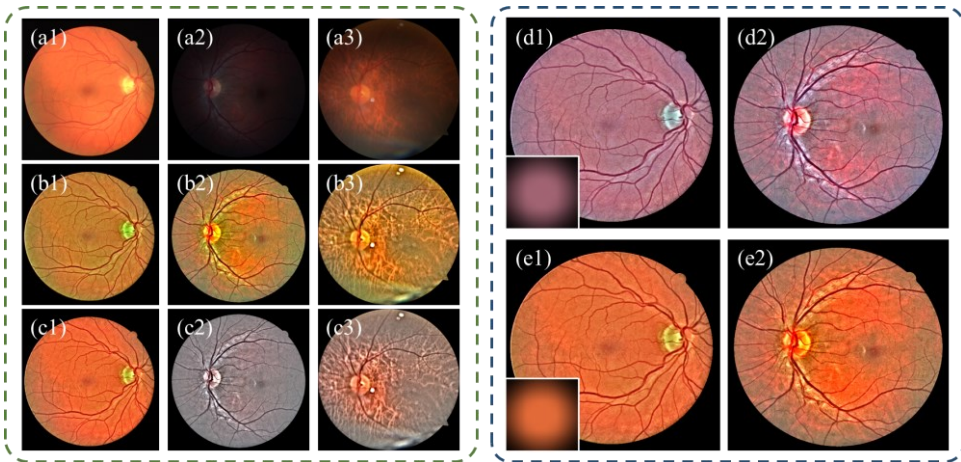


Fig. 5-10. Ablation study on color implant module. (a1) to (a3) are raw images. (b1) to (b3) are enhanced images when the color implant module is turned on. (c1) to (c3) are enhanced images when the color implant module is turned off. (d1) to (e2) are enhanced images with a different sampled background color.

As shown in Fig. 5-10 (a1) to (b3), although the raw input images are different, for example, Fig. 5-10 (a1) are orange while Fig. 5-10 (a2) are black/purple, the output images of the InQue in Fig. 5-10 (b1) to (b3) have similar color distribution. In this way, the color implant module can correct the color distortion during the fundus photograph

process. As the Gaussian kernel for the color implant is sufficiently large, this process won't cause a loss of detailed image information.

Without the color implant module, as shown in Fig. 5-10 (c1) to (c3), the images may lose color saturation due to the insufficient illumination condition. For example, the background of the output image in Fig. 5-10 (c2) is gray, and both the color and naturalness severely degenerate. Note that the color implant won't impact subsequent heating and cooling procedures as the contrast of output images is still improved.

Furthermore, we can use a different sampled background color to achieve the color implant. For example, as shown in Fig. 5-10 (d1) and (d2), we used the purple color sampled from the HRF dataset to achieve the color implant. While we can also use the orange background color sampled from Fig. 5-10 (a1) to achieve the color implant as shown in Fig. 5-10 (e1) and (e2).

6.2 On the Luma heater

In this section, we performed an ablation study on the Luma heater module. As shown in the experimental results section, Luma heater can increase the average intensity of the Luma channel of input images and normalize the intensity distribution to a narrow range. To further analyze this property, we calculated the average intensity distribution for the Y channel of retinal images after each step of InQue. The average intensity distribution of the raw images has a large deviation, as shown by the blue bar in Fig. 5-11. After the color implantation, the average intensity distribution concentrates in a narrower range as shown by the orange bars. The Luma heater further increased the average intensity.

It is worth noting that contrast improvement occurs in the Luma cooler module through dehazing, which generally decreases the intensity level. Insufficient intensity may lead to poor contrast improvement, as the cooling process reaches the requirement of the cost function too early. The results of turning off the Luma heater are presented

in Fig. 5-12 (a1) to (a3), which demonstrate that the contrast of these final output images is not as good as that in Fig. 5-10 (b1) to (b3).

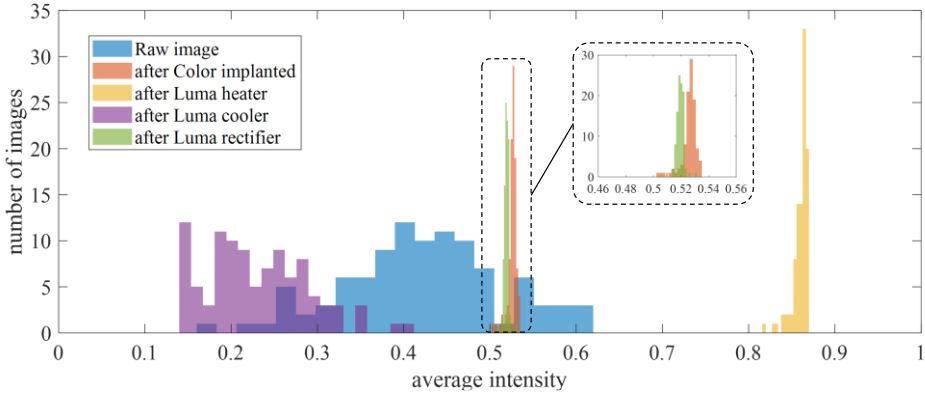


Fig. 5-11. the average intensity distribution of raw images and processed images in each step of InQue.

On the other hand, the Luma cooler significantly decreases the intensity of input images, thereby improving the image's contrast. As shown in the purple bars in Fig. 5-11, the average intensity of the cooled-down images is even lower than that of the raw images, and the deviation is also large. With the Luma rectifier, the intensity distribution becomes concentrated again, as shown by the green bars in Fig. 5-11.

We also studied the impact of parameter R_{Heater} on the final output of InQue. For example, as shown in Fig. 5-12 (b1) to (b3) and (d1) to (d3), the contrast improvement becomes better for larger value of R_{Heater} . For example, the contrast of Fig. 5-12 (d3) where $R_{Heater} = 0.5$ is better than that of Fig. 5-12 (d1) where $R_{Heater} = 0.01$. The intermediate Luma channels are shown in Fig. 5-12 (c1) to (e3). Note that we should not let too many pixels being over-exposed as it may causes problematic pixels and loss of image information. We, therefore, use $R_{Heater} = 0.02$ in our experiments to ensure data fidelity.

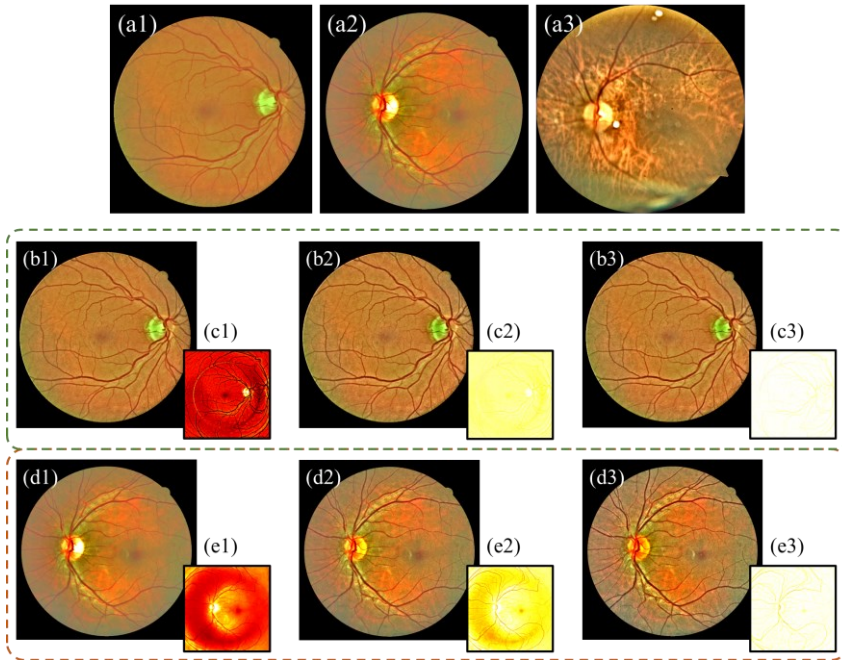


Fig. 5-12. Ablation study on Luma heater module. (a1) to (a3) are enhanced images without Luma heater. (b1) to (b3) are enhanced images on (a1) with Luma heater, for $R_{Heater} = 0.01$, $R_{Heater} = 0.05$, and $R_{Heater} = 0.5$, respectively. (c1) to (c3) are corresponding Luma channel after the heating process. (d1) to (d3) and (e1) to (e3) are similar experiments applied to (a2).

6.3 On the Luma rectifier

In this section, we present results of our ablation study on the Luma rectifier module. As discussed in the experimental results section, the Luma rectifier plays a crucial role in normalizing the output of intensity distribution from the Luma heater. Without the Luma rectifier, the pixel values in the Luma channel may exceed the correct intensity range of 0 to 1, leading to over-exposed or under-exposed images, as shown in Fig. 5-13 (a1) to (a3). Although the contrast of the images is significantly improved, the naturalness is severely degraded, which is considered to be an incorrect enhancement.

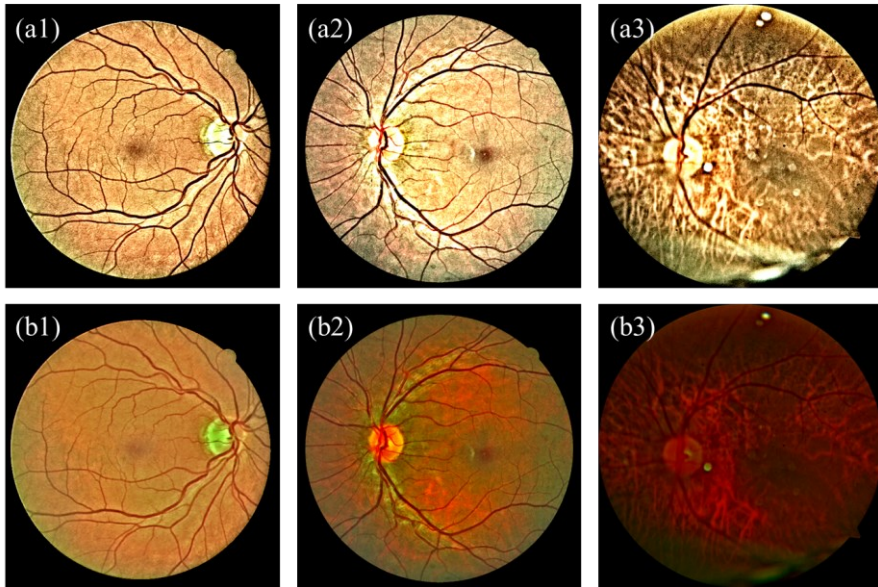


Fig. 5-13. Ablation study on Luma rectifier module. (a1) to (a3) are enhanced images without Luma rectifier. (b1) to (b3) are enhanced images without using Eq. (5-14).

However, these incorrect enhancements are not entirely useless, as demonstrated in Fig. 5-13 (a1) to (a3). The contrast of blood vessels is significantly improved, with vessel branches being more easily distinguishable from the background, and the color distribution along the vessel structures is more uniform. We believe that these images could potentially benefit blood vessel segmentations.

Equation (5-14) is a simple but important component of the Luma rectifier module. It remaps the intensity distribution into a contrast value so that the final average Luma intensity distribution of all output images is normalized to a similar level. The output without Eq (5-14) is shown in Fig. 5-13 (b1) to (b3), where the image in Fig. 5-13 (b3) is distorted, while the images in Fig. 5-13 (b2) are not bright enough to provide comfortable visual performance, compared to Fig. 5-13 (b1).

6.4 On the smooth layer

In this section, we perform an ablation study on the smooth layer and smooth function in both Luma heater and cooler modules. As written in Eq. (5-8) and Eq. (5-11), the smooth layer and the smooth function play important roles in the framework of InQue. However, we will show that the InQue is not sensitive to the number of smooth layers, nor sensitive to certain image smooth functions.

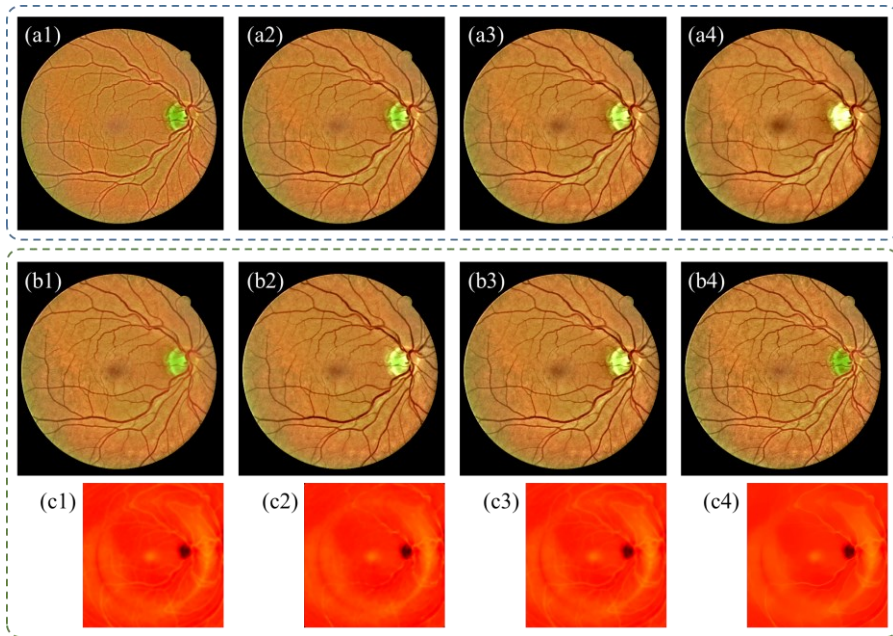


Fig. 5-14. Ablation study on smooth layers. (a1) to (a4) are enhanced images for $L = 2$, $L = 5$, $L = 8$, and $L = 16$. (b1) to (b4) are enhanced images using the guided filter, median filter, domain transfer, and total variation as smooth function. (c1) to (c4) are the first smoothed layer for corresponding outputs.

Fig. 5-14 (a1) to (a4) shows the final output of InQue for $L = 2$, $L = 5$, $L = 8$, and $L = 16$, respectively. The overall contrast enhancement effects are similar to each other

among the four images. While the intensity near the optic disk shown in Fig. 5-14 (a4) is larger than others. Note that large smooth layers contain over-smoothed images. In the extreme case, the over-smoothed images become constant white color blocks where no detailed image information appears. In such cases, these over-smoothed images lose the function of intensity correction/dehazing.

Increasing the number of smoothed layers without changing the smooth strength for each layer may not benefit the retinal image enhancement in InQue. By decreasing the initial kernel size of the Gaussian filter, we can ensure that the larger layers contain sufficient detailed information for a retinal image but increase the computational cost.

The smooth function used in InQue is not limited by a Gaussian filter. We have also tested other image smooth functions such as guided filter, median filter, domain transfer, as well as total variation. Results are shown in Fig. 5-14 (b1) to (b4). Again, the overall enhancement effects are similar except for the optic disk area. The feasibility and efficiency of different smooth functions of InQue originates from the observation that either the illumination or haze layer is spatially slow varying. Under such an assumption, we can regard the smooth function as the feature extractor that extracts the illumination/haze feature from the given images.

Different smoothing degree extracts different scales of features, as such, the illumination/haze can be efficiently represented and estimated using the linear combination of different smooth layers as shown in Eq. (5-8) and Eq. (5-11). Since image smoothing function including guided filter, median filter, domain transfer, and total variation can all smooth the image to different degrees according to the model parameters, they are in principle valid to achieve image illumination correction and dehazing as shown in Fig. 5-14. (b1) to (b4).

6.5 Limitations of InQue

InQue has in general two limitations. The first limitation is related to the tuning of its parameters. A total of three parameters, including R_{Heater} , R_{cooler} and I_0 can influent

the performance of InQue. For example, a larger value of R_{cooler} increases the contrast of the enhancement, while also decreasing the intensity. Users may need to manually adjust these parameters for better enhancement quality. Fortunately, our experiments on 2716 images show that $R_{heater} = 0.02$, $R_{cooler} = 0.01$ and $I_0 = 0.52$ are appropriate to obtain promising restoration quality.

Secondly, InQue can sometimes introduce unexpected color distortion near the optical disk area due to the color implant module. Currently, we use a pre-determined value for the background color, which is not adaptive determined based on different locations of different input images. As such, the naturalness of InQue outputs is not as good as that of learning-based methods. Further research on this topic includes using a network to achieve the color implant for naturalness preservation while using a non-learning-based method to achieve contrast enhancement for sake of interpretation.

7. Concluding remarks

In this research, we proposed InQue for color retinal image restoration. It corrects insufficient and uneven intensity distributions while suppressing haze effects. The naturalness of the retinal image is maintained.

The effectiveness of InQue doesn't rely on the simulated degraded retinal image in some certain pixel size. InQue was tested on different datasets with varying pixel sizes. The algorithm proved to be effective across 6 published datasets, demonstrating its versatility and reliability in a clinical setting. Additionally, the interpretability of each step and intermediate product of InQue allows ophthalmologists to understand and trust the algorithm's enhancement process.

We believe that InQue provides a reliable and safe way for retinal image enhancement, which could have further clinical applications. By improving the visual quality and naturalness of retinal images, InQue can assist ophthalmologists in accurately diagnosing and treating various retinal diseases, such as diabetic retinopathy, macular degeneration, and glaucoma. Furthermore, the ability of InQue to

enhance ultra-wide-field retinal images could prove valuable in detecting peripheral pathologies that may be missed with traditional imaging techniques. Particularly, the ability of how InQue can assist clinical applications, and how it can be further improved on image naturalness preservation and contrast enhancement will be investigated in the future.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding. This research is supported by China Scholarship Council (CSC) (201908340078)

References

1. E. Decenci re, G. Cazuguel, X. Zhang, G. Thibault, J. C. Klein, F. Meyer, B. Marcotegui, G. Quellec, M. Lamard, R. Danno, D. Elie, P. Massin, Z. Viktor, A. Erginay, B. Lajand A. Chabouis, "TeleOphta: Machine learning and image processing methods for teleophthalmology," *IRBM* **34**, 196-203 (2013).
2. F. Huang, B. Dashtbozorg, J. Zhang, E. Bekkers, S. Abbasi-Sureshjani, T. T. Berendschot, and B. M. Ter Haar Romeny, "Reliability of Using Retinal Vascular Fractal Dimension as a Biomarker in the Diabetic Retinopathy Detection," *J. Ophthalmol.* **2016**, 6259047 (2016).
3. M. C. V. S. Mary, E. B. Rajsingh, and G. R. Naik, "Retinal Fundus Image Analysis for Diagnosis of Glaucoma: A Comprehensive Survey," *IEEE Access* **4**, 4327-4354 (2016).
4. D. M. S. Barros, J. C. C. Moura, C. R. Freire, A. C. Taleb, R. A. M. Valentim, and P. S. G. Morais, "Machine learning applied to retinal image processing for glaucoma detection: review and perspective," *BioMedical Engineering OnLine* **19**, 20 (2020).
5. R. J. Chalakkal, W. H. Abdulla, and S. C. Hong, "3 - Fundus retinal image analyses for screening and diagnosing diabetic retinopathy, macular edema, and glaucoma disorders," in *Diabetes and Fundus OCT*, A. S. El-Baz, and J. S. Suri, eds. (Elsevier, 2020), pp. 59-111.
6. M. Trieschmann, F. J. van Kuijk, R. Alexander, P. Hermans, P. Luthert, A. C. Bird, and D. Pauleikhoff, "Macular pigment in the human retina: histological evaluation of localization and distribution," *Eye* **22**, 132-137 (2008).
7. T. S. Eliasdottir, "Retinal oximetry and systemic arterial oxygen levels," *Acta Ophthalmologica* **96**, 1-44 (2018).
8. E. Pead, R. Megaw, J. Cameron, A. Fleming, B. Dhillon, E. Trucco, and T. J. s. o. MacGillivray, "Automated detection of age-related macular degeneration in color fundus photography: a systematic review," **64**, 498-511 (2019).
9. P. Zhu, F. Huang, F. Lin, Q. Li, Y. Yuan, Z. Gao, and F. Chen, "The relationship of retinal vessel diameters and fractal dimensions with blood pressure and cardiovascular risk factors," *PloS one* **9**, e106551 (2014).
10. D. Drobnyak, I. C. Munch, C. Glumer, K. Faerch, L. Kessel, M. Larsen, and N. C. Veiby, "Retinal Vessel Diameters and Their Relationship with Cardiovascular Risk and All-Cause Mortality in the Inter99 Eye Study: A 15-Year Follow-Up," *J. Ophthalmol.* **2016**, 6138659 (2016).
11. S. P. J. P. R. Rajan, and I. Analysis, "Recognition of cardiovascular diseases through retinal images using optic cup to optic disc ratio," **30**, 256-263 (2020).
12. C. J. Heaven, J. Cansfield, and K. M. Shaw, "The quality of photographs produced by the non-mydratric fundus camera in a screening programme for diabetic retinopathy: A 1 year prospective study," *Eye* **7**, 787-790 (1993).
13. P. H. Scanlon, C. Foy, R. Malhotra, and S. J. J. D. C. Aldington, "The influence of age, duration of diabetes, cataract, and pupil size on image quality in digital photographic retinal screening," **28**, 2448-2453 (2005).
14. M. Zhou, K. Jin, S. Wang, J. Ye, and D. Qian, "Color Retinal Image Enhancement Based on Luminosity and Contrast Adjustment," *IEEE Transactions on Biomedical Engineering* (2018).
15. B. Gupta, and M. Tiwari, "Color retinal image enhancement using luminosity and quantile based contrast enhancement," *Multidimensional Systems and Signal Processing* **30**, 1829-1837 (2019).
16. L. Cao, H. Li, and Y. Zhang, "Retinal image enhancement using low-pass filtering and α -rooting," *Signal Processing* **170**, 107445 (2020).
17. J. Wang, Y.-J. Li, and K.-F. Yang, "Retinal fundus image enhancement with image decomposition and visual adaptation," *Computers in Biology and Medicine* **128**, 104116 (2021).

18. E. Peli, and T. Peli, "Restoration of retinal images obtained through cataracts," *IEEE Transactions on Medical Imaging* **8**, 401-406 (1989).
19. A. Mitra, S. Roy, S. Roy, and S. K. Setua, "Enhancement and restoration of non-uniform illuminated Fundus Image of Retina obtained through thin layer of cataract," *Computer Methods and Programs in Biomedicine* **156**, 169-178 (2018).
20. L. Xiong, H. Li, and L. Xu, "An enhancement method for color retinal images based on image formation model," *Computer Methods and Programs in Biomedicine* **143**, 137-150 (2017).
21. B. Savelli, A. Bria, A. Galdran, C. Marrocco, M. Molinara, A. e. Campilho, and F. Tortorella, "Illumination correction by dehazing for retinal vessel segmentation," *IEEE 30th International Symposium on Computer-Based Medical Systems* (2017).
22. A. Gaudio, A. Smailagic, and A. e. Campilho, "Enhancement of Retinal Fundus Images via Pixel Color Amplification," *International Conference on Image Analysis and Recognition* (2020).
23. L. Cao, and H. Li, "Detail-richest-channel based enhancement for retinal image and beyond," *Biomedical Signal Processing and Control* **69**, 102933 (2021).
24. Y. Luo, K. Chen, L. Liu, J. Liu, J. Mao, G. Ke, and M. Sun, "Dehaze of Cataractous Retinal Images using an Unpaired Generative Adversarial Network," *IEEE Journal of Biomedical and Health Informatics* (2020).
25. H. Li, H. Liu, Y. Hu, H. Fu, Y. Zhao, H. Miao, and J. Liu, "An annotation-free restoration network for cataractous fundus images," *IEEE Transactions on Medical Imaging* (2022).
26. B. Yang, H. Zhao, L. Cao, H. Liu, N. Wang, and H. Li, "Retinal image enhancement with artifact reduction and structure retention," *Pattern Recognition* **133**, 108968 (2023).
27. Z. Shen, H. Fu, J. Shen, and L. Shao, "Modeling and Enhancing Low-Quality Retinal Fundus Images," *IEEE Transactions on Medical Imaging* **40**, 996-1006 (2021).
28. C. Wan, X. Zhou, Q. You, J. Sun, J. Shen, S. Zhu, Q. Jiang, and W. Yang, "Retinal Image Enhancement Using Cycle-Constraint Adversarial Network," *Front. Med.* **8**, 793726 (2021).
29. S. Chen, Z. Qian, and Z. Hua, "A Novel Un-Supervised GAN for Fundus Image Enhancement with Classification Prior Loss," *Electronics* **11** (2022).
30. K. TSIMA, "The reproducibility issues that haunt health-care AI," *Nature* **613** (2023).
31. S. Zhang, C. A. B. Webers, and T. T. J. M. Berendschot, "A double-pass fundus reflection model for efficient single retinal image enhancement," *Signal Processing* **192**, 108400 (2022).
32. S. Zhang, "A reflective background filling algorithm for fundus images preprocessing," <https://github.com/ShuheZhang-MUMC/Fundus-image-background-filling>.
33. S. Sun, and X. Guo, "Image Enhancement Using Bright Channel Prior," in *2016 International Conference on Industrial Informatics - Computing Technology, Intelligent Technology, Industrial Information Integration (ICIICI)*(2016), pp. 83-86.
34. B. Savelli, A. Bria, A. Galdran, C. Marrocco, M. Molinara, A. Campilho, and F. Tortorella, "Illumination Correction by Dehazing for Retinal Vessel Segmentation," in *2017 IEEE 30th International Symposium on Computer-Based Medical Systems (CBMS)*(2017), pp. 219-224.
35. M. Foracchia, E. Grisan, and A. Ruggeri, "Luminosity and contrast normalization in retinal images," *Med Image Anal* **9**, 179-190 (2005).
36. E. Decencière, X. Zhang, G. Cazuguel, B. Lay, B. Cochener, C. Trone, P. Gain, R. Ordonez, P. Massin, A. Erginay, B. Charton, and J.-C. Klein, "Feedback on a publicly distributed database: the Messidor database," *2014* **33**, 4 (2014).
37. kaggle, "cataract dataset."

38. T. Kauppi, V. Kalesnykiene, J.-K. Kamarainen, L. Lensu, I. Sorri, H. Uusitalo, H. Kälviäinen, and J. Pietilä, "DIARETDB0: Evaluation database and methodology for diabetic retinopathy algorithms," *Machine Vision and Pattern Recognition Research Group, Lappeenranta University of Technology, Finland* **73**, 1-17 (2006).
39. R. Kälviäinen, and H. Uusitalo, "DIARETDB1 diabetic retinopathy database and evaluation protocol," in *Medical image understanding and analysis*(Citeseer2007), p. 61.
40. P. Porwal, S. Pachade, R. Kamble, M. Kokare, G. Deshmukh, V. Sahasrabuddhe, and F. Meriaudeau, "Indian Diabetic Retinopathy Image Dataset (IDRIID): A Database for Diabetic Retinopathy Screening Research," *Data* **3**, 25 (2018).
41. H. Li, H. Liu, H. Fu, H. Shu, Y. Zhao, X. Luo, Y. Hu, and J. Liu, "Structure-consistent restoration network for cataract fundus image enhancement," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*(Springer2022), pp. 487-496.
42. H. Li, H. Liu, Y. Hu, R. Higashita, Y. Zhao, H. Qi, and J. Liu, "Restoration Of Cataract Fundus Images Via Unsupervised Domain Adaptation," in *2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI)*(2021), pp. 516-520.
43. A. Rizzi, T. Algeri, G. Medeghini, and D. Marini, "A proposal for Contrast Measure in Digital Images," (2004).
44. K. Panetta, C. Gao, and S. Agaian, "Human-Visual-System-Inspired Underwater Image Quality Measures," *IEEE Journal of Oceanic Engineering* **41**, 1-11 (2016).
45. L. K. Choi, J. You, and A. C. Bovik, "Referenceless Prediction of Perceptual Fog Density and Perceptual Image Defogging," *IEEE Transactions on Image Processing* **24**, 3888-3901 (2015).
46. c. a. q. r. g. Video processing, "Ways of cheating on popular objective metrics: blurring, noise, super-resolution and others."
47. G. Eilertsen, S. Hajisharif, P. Hanji, A. Tsirikoglou, R. K. Mantiuk, and J. Unger, "How to cheat with metrics in single-image HDR reconstruction," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*(2021), pp. 3998-4007.
48. E. Midena, G. Marchione, S. Di Giorgio, G. Rotondi, E. Longhin, L. Frizziero, E. Pilotto, R. Parrozzani, and G. Midena, "Ultra-wide-field fundus photography compared to ophthalmoscopy in diagnosing and classifying major retinal diseases," *Scientific Reports* **12**, 19287 (2022).
49. V. Kumar, A. Surve, D. Kumawat, B. Takkar, S. Azad, R. Chawla, D. Shroff, A. Arora, R. Singh, and P. Venkatesh, "Ultra-wide field retinal imaging: A wider clinical perspective," *Indian Journal of Ophthalmology* **69**, 824-835 (2021).

Supplementary Notes for Chapter 5

InQue: An intensity-quenching scheme for large-scale fully unsupervised retinal image enhancement

Supplementary Note 5-1: Additional results for retinal images using InQue

In this section of the Supplementary Materials, additional three groups of experimental results using InQue are shown for visual assessment.

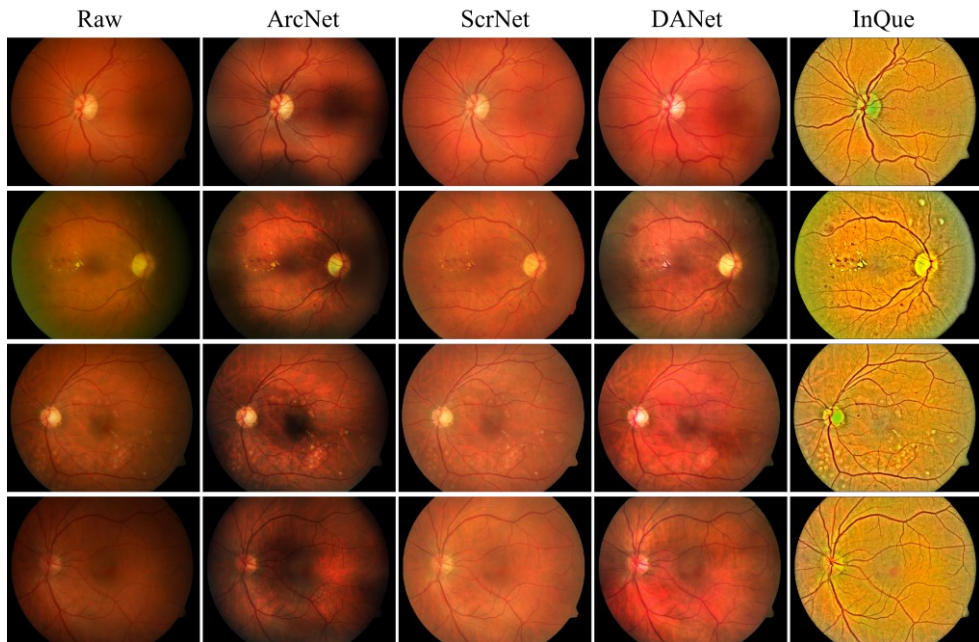


Fig. s5-1. Restoration results for images in DiaRet dataset. Image from left to right columns are raw, image and enhanced image by ArcNet, ScrNet, DANet, and InQue, respectively.

Tab. s5-1. Objective assessments for images in Fig. s5-1.

Raw image	Matrixes	Raw	Methods			
			ArcNet	ScrNet	DANet	InQue
1 st row	UIConM ↑	0.0371	0.0557	0.0611	0.0690	0.2090
	UISM ↑	2.1514	2.3798	2.5441	1.9321	4.2210
	IE ↑	6.5335	6.5843	6.7679	6.9028	6.8273
	C _{RAMM} ↑	1.1418	1.6036	1.6140	1.6176	3.9403
	FADE ↓	0.3309	0.3206	0.3285	0.2978	0.1560
2 nd row	UIConM ↑	0.0407	0.0567	0.0570	0.0769	0.2425
	UISM ↑	1.7551	2.3160	2.2620	2.0901	4.2670
	IE ↑	5.5170	6.4439	6.3437	6.8846	6.5068
	C _{RAMM} ↑	1.1285	1.5049	1.5149	1.7071	4.2711
	FADE ↓	0.3491	0.3360	0.3619	0.2910	0.1545
3 rd row	UIConM ↑	0.0440	0.0675	0.0697	0.0814	0.2385
	UISM ↑	2.5360	2.7650	2.7310	2.1407	4.3903
	IE ↑	6.5993	6.4786	6.6187	6.8607	6.8697
	C _{RAMM} ↑	1.2721	1.7605	1.7200	1.8865	4.1520
	FADE ↓	0.3201	0.2930	0.3282	0.2790	0.1547
4 th row	UIConM ↑	0.0225	0.0388	0.0629	0.0775	0.2038
	UISM ↑	2.4864	2.6003	2.5510	2.3117	4.2604
	IE ↑	6.2962	6.2189	6.5495	6.8210	6.7909
	C _{RAMM} ↑	0.9382	1.3860	1.6806	1.9038	3.7757
	FADE ↓	0.3584	0.3339	0.3368	0.2758	0.1577

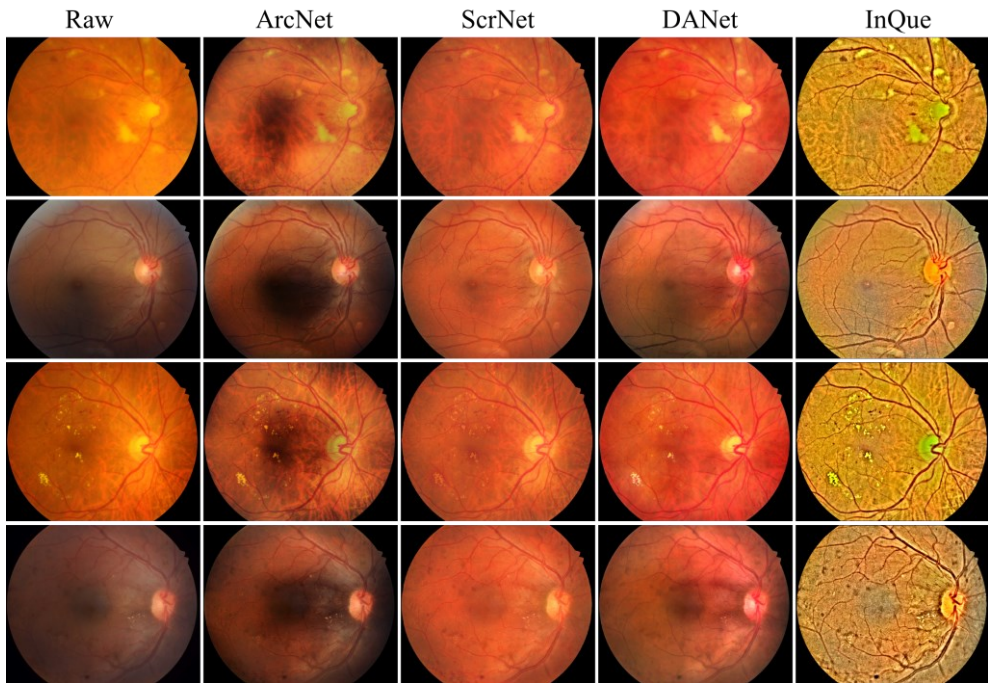


Fig. s5-2. Restoration results for images in IDRiD dataset

Tab. s5-2. Objective assessments for images in Fig. s5-2.

Raw image	Matrixes	Raw	Methods			
			ArcNet	ScrNet	DANet	InQue
1 st row	UIConM ↑	0.0651	0.0906	0.0687	0.0841	0.2557
	UISM ↑	1.4097	1.7491	1.9927	1.7678	4.2259
	IE ↑	6.6627	7.1117	6.6713	6.8535	6.9420
	C _{RAMM} ↑	1.7909	2.3043	1.8836	1.9626	5.4494
	FADE ↓	0.2994	0.2708	0.3134	0.2917	0.1462
2 nd row	UIConM ↑	0.0505	0.0742	0.0722	0.0810	0.2143
	UISM ↑	1.3749	1.8759	2.3743	1.9348	4.0983
	IE ↑	6.6037	6.8223	6.8575	7.0137	7.0826
	C _{RAMM} ↑	1.4876	1.9138	1.8828	2.0239	4.2629
	FADE ↓	0.4062	0.3054	0.3007	0.2882	0.1770
3 rd row	UIConM ↑	0.0857	0.1089	0.0765	0.1004	0.2955
	UISM ↑	2.0622	2.3654	2.7104	2.0101	4.6910
	IE ↑	6.7087	7.0537	6.6848	6.8176	6.9008
	C _{RAMM} ↑	2.0157	2.6741	2.0023	2.0897	5.8417
	FADE ↓	0.2377	0.2244	0.2746	0.2613	0.1351
4 th row	UIConM ↑	0.0485	0.0814	0.0716	0.0749	0.2731
	UISM ↑	1.3022	1.9576	2.0907	1.7464	4.4586
	IE ↑	7.3940	7.0033	6.7484	6.9294	7.1208
	C _{RAMM} ↑	1.4531	2.0553	1.8242	1.8356	5.2829
	FADE ↓	0.5698	0.3024	0.3069	0.2971	0.1572

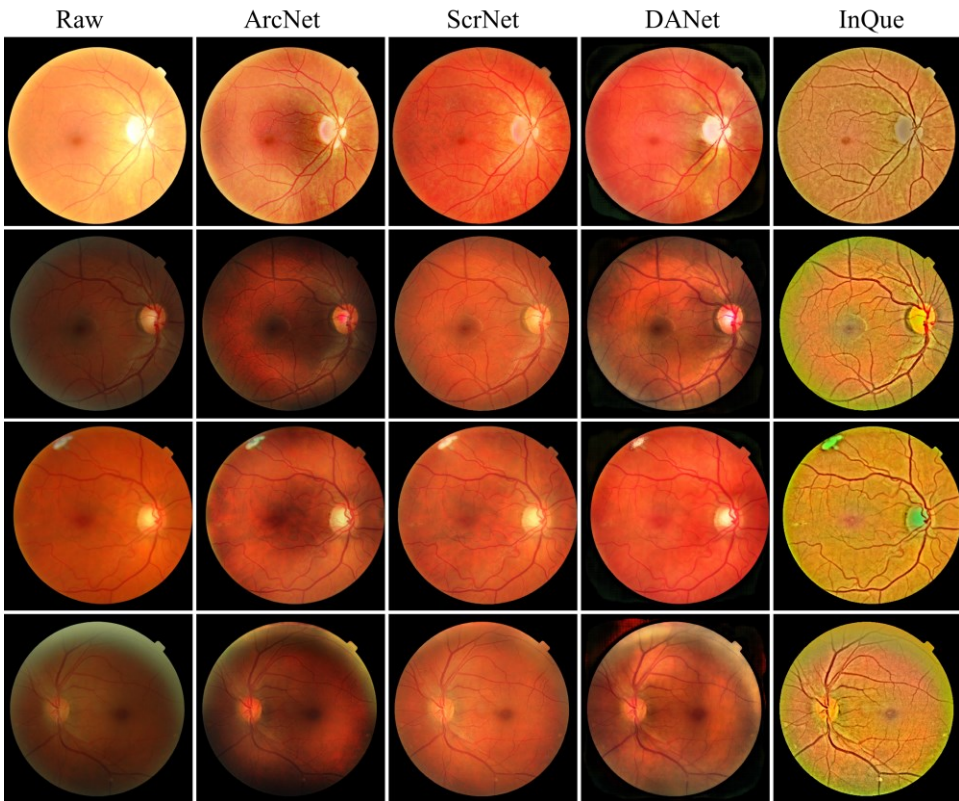


Fig. s5-3. Restoration results for images in Messidor dataset

Tab. s5-3. Objective assessments for images in Fig. s5-3.

Raw image	Matrixes	Raw	Methods			
			ArcNet	ScrNet	DANet	InQue
1 st row	UIConM ↑	0.1489	0.1505	0.1042	0.1296	0.1856
	UISM ↑	1.3450	1.4791	1.7171	2.5468	4.5321
	IE ↑	5.1925	6.1719	5.9570	6.9079	5.9742
	C _{RAMM} ↑	3.0826	3.2274	2.3463	2.8747	3.7685
	FADE ↓	0.3145	0.2165	0.2211	0.2147	0.1885
2 nd row	UIConM ↑	0.0738	0.0833	0.0865	0.0990	0.1828
	UISM ↑	1.8882	1.9131	3.2330	2.7633	4.8145
	IE ↑	5.6373	5.9396	5.7351	6.6232	5.7733
	C _{RAMM} ↑	1.9475	2.2349	2.2166	2.4314	3.8098
	FADE ↓	0.2317	0.2098	0.2307	0.2011	0.1586
3 rd row	UIConM ↑	0.0753	0.0945	0.0983	0.1044	0.2002
	UISM ↑	1.5542	1.5180	1.9053	2.4069	4.5999
	IE ↑	5.7460	6.0755	5.8994	6.5432	6.0200
	C _{RAMM} ↑	1.8926	2.3373	2.3328	2.3710	4.0795
	FADE ↓	0.2749	0.2412	0.2551	0.2240	0.1617
4 th row	UIConM ↑	0.0581	0.0689	0.0806	0.1076	0.1964
	UISM ↑	1.9748	1.9798	3.0509	2.5039	4.6448
	IE ↑	5.5756	5.5759	5.8291	6.8333	6.0338
	C _{RAMM} ↑	1.7436	1.9477	2.0706	2.8085	3.9281
	FADE ↓	0.2972	0.2470	0.2519	0.2021	0.1646

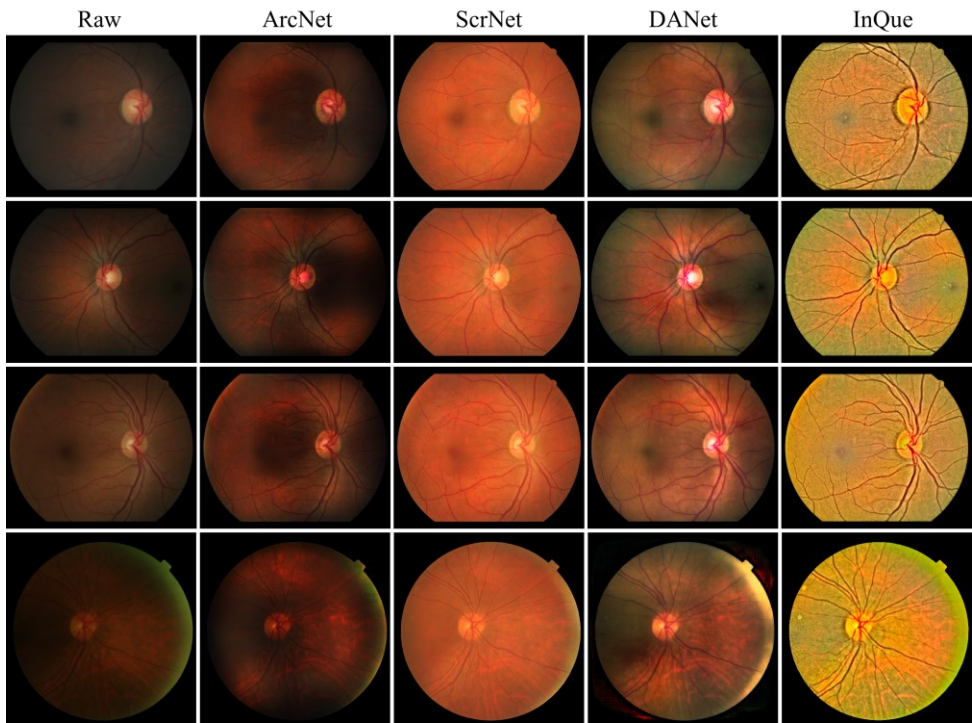


Fig. s5-4. Restoration results for images in E_optha_MA dataset

Statistical analysis of objective assessments for images from 5 published datasets are listed in Tab. s5-5. Accordingly, the proposed InQue outperforms other deep-learning methods in both contrast enhancement and hazing elimination.

In general, our proposed InQue method corrects the uneven and sufficient illumination as shown in Fig. s5-3 and Fig. s5-4 while also enhancement the contrast of the image. The color naturalness is also maintains and can be even better than that of deep-learning method as shown in the second column and fifth column in Fig. s5-2 to Fig. s5-4.

Tab. s5-4. Objective assessments for images in Fig. s5-4.

Raw image	Matrixes	Raw	Methods			
			ArcNet	ScrNet	DANet	InQue
1 st row	UIConM ↑	0.0199	0.0308	0.0688	0.0605	0.2138
	UISM ↑	1.6235	2.1624	1.5804	2.2556	4.5975
	IE ↑	5.2678	5.6683	6.2244	6.3681	6.5480
	C _{RAMM} ↑	0.9253	1.2263	1.9358	1.8735	4.3129
	FADE ↓	0.4897	0.3431	0.3179	0.2631	0.1644
2 nd row	UIConM ↑	0.0238	0.0333	0.0746	0.0753	0.2174
	UISM ↑	1.9683	2.5016	2.0489	2.3944	4.6185
	IE ↑	5.5167	5.7091	6.2558	6.4246	6.5191
	C _{RAMM} ↑	0.9982	1.2755	2.0096	2.1063	4.2110
	FADE ↓	0.3795	0.3332	0.3109	0.2399	0.1668
3 rd row	UIConM ↑	0.0299	0.0424	0.0778	0.0799	0.2016
	UISM ↑	1.8073	2.1607	1.7356	2.3650	4.4800
	IE ↑	5.7400	5.8194	6.3294	6.5339	6.5666
	C _{RAMM} ↑	1.2294	1.5497	2.1224	2.2338	4.1239
	FADE ↓	0.3364	0.2895	0.2746	0.2315	0.1728
4 th row	UIConM ↑	0.0231	0.0368	0.0741	0.0949	0.1904
	UISM ↑	1.9422	2.1761	2.6805	2.2027	4.4602
	IE ↑	5.1526	5.2168	5.7335	6.5959	5.9199
	C _{RAMM} ↑	1.0295	1.3103	1.9654	2.6409	3.9724
	FADE ↓	0.3155	0.2900	0.2781	0.2038	0.1459

Tab. s5-5. Objective assessments for images from 5 published datasets.

Databases mean \pm std.	Matrixes	Raw	Methods			
			ArcNet	ScrNet	DANet	InQue
DiaRet (219 images)	UIConM \uparrow	0.0349 \pm 0.0134	0.0533 \pm 0.0171	0.0634 \pm 0.0063	0.0751 \pm 0.0066	0.2236 \pm 0.0285
	UISM \uparrow	2.1293 \pm 0.3066	2.3928 \pm 0.1874	2.4570 \pm 0.1797	2.0575 \pm 0.1620	4.2592 \pm 0.2135
	IE \uparrow	6.2678 \pm 0.4556	6.4382 \pm 0.2704	6.6119 \pm 0.1192	6.8389 \pm 0.0550	6.7715 \pm 0.1319
	C _{RAMM} \uparrow	1.1209 \pm 0.2438	1.5309 \pm 0.2923	1.6589 \pm 0.1096	1.7941 \pm 0.1124	4.1653 \pm 0.5135
	FADE \downarrow	0.3495 \pm 0.0331	0.3264 \pm 0.0343	0.3337 \pm 0.0217	0.2814 \pm 0.0126	0.1571 \pm 0.0151
E_optha_MA (381 images)	UIConM \uparrow	0.0558 \pm 0.0224	0.0742 \pm 0.0255	0.0901 \pm 0.0111	0.0967 \pm 0.0099	0.2099 \pm 0.0235
	UISM \uparrow	1.9814 \pm 0.4603	2.3062 \pm 0.3920	2.5058 \pm 0.8442	2.4847 \pm 0.2737	4.5085 \pm 0.3208
	IE \uparrow	5.8492 \pm 0.3294	6.0748 \pm 0.3227	6.1562 \pm 0.2651	6.6559 \pm 0.1133	6.3212 \pm 0.2877
	C _{RAMM} \uparrow	1.6644 \pm 0.3896	2.0980 \pm 0.4372	2.2478 \pm 0.1546	2.4526 \pm 0.1863	4.2550 \pm 0.4096
	FADE \downarrow	0.3100 \pm 0.0432	0.2373 \pm 0.0465	0.2420 \pm 0.0437	0.2175 \pm 0.0188	0.1662 \pm 0.0165
IDRID (516 images)	UIConM \uparrow	0.0555 \pm 0.0143	0.0827 \pm 0.0177	0.0733 \pm 0.0094	0.0870 \pm 0.0081	0.2535 \pm 0.0352
	UISM \uparrow	1.6781 \pm 0.2591	2.2884 \pm 0.2801	2.4219 \pm 0.3021	1.9561 \pm 0.1723	4.3493 \pm 0.2423
	IE \uparrow	6.8749 \pm 0.4103	6.8609 \pm 0.2087	6.6792 \pm 0.1682	6.9580 \pm 0.0673	6.9900 \pm 0.1280
	C _{RAMM} \uparrow	1.5661 \pm 0.2416	2.1277 \pm 0.2810	1.8785 \pm 0.1356	2.0194 \pm 0.1133	4.9788 \pm 0.7325
	FADE \downarrow	0.4281 \pm 0.0990	0.2694 \pm 0.0262	0.2968 \pm 0.0237	0.2739 \pm 0.0148	0.1540 \pm 0.0159
Messidor (1200 images)	UIConM \uparrow	0.0707 \pm 0.0180	0.0858 \pm 0.0180	0.0887 \pm 0.0067	0.1023 \pm 0.0064	0.1979 \pm 0.0146
	UISM \uparrow	1.8860 \pm 0.3411	1.8960 \pm 0.2985	3.0062 \pm 0.6421	2.6863 \pm 0.1282	4.7216 \pm 0.1375
	IE \uparrow	5.7305 \pm 0.2147	5.9929 \pm 0.1638	5.8864 \pm 0.1144	6.6984 \pm 0.0961	5.9891 \pm 0.1064
	C _{RAMM} \uparrow	1.8749 \pm 0.2913	2.2453 \pm 0.2632	2.2321 \pm 0.1100	2.4912 \pm 0.1193	4.0391 \pm 0.2619
	FADE \downarrow	0.2717 \pm 0.0259	0.2246 \pm 0.0173	0.2361 \pm 0.0147	0.2054 \pm 0.0115	0.1611 \pm 0.0113
Normal (300 images)	UIConM \uparrow	0.0736 \pm 0.0148	0.0920 \pm 0.0148	0.0854 \pm 0.0047	0.0989 \pm 0.0061	0.2048 \pm 0.0219
	UISM \uparrow	1.5712 \pm 0.2847	1.7946 \pm 0.2679	2.0733 \pm 0.5546	2.5942 \pm 0.1415	4.7192 \pm 0.1877
	IE \uparrow	6.1005 \pm 0.2449	6.2678 \pm 0.1615	6.0993 \pm 0.1374	6.6686 \pm 0.0717	6.2734 \pm 0.1356
	C _{RAMM} \uparrow	1.9931 \pm 0.2334	2.4263 \pm 0.2377	2.2287 \pm 0.0847	2.4332 \pm 0.0997	4.2579 \pm 0.4721
	FADE \downarrow	0.3246 \pm 0.0370	0.2386 \pm 0.0194	0.2473 \pm 0.0145	0.2202 \pm 0.0111	0.1732 \pm 0.0127

Supplementary Note 5-2: Comparing against State-of-the-art non-deep-learning methods.

To show its performance, this section compares, both qualitatively and quantitatively, MUTE with that of three recently published state-of-art methods, namely the Luminosity and Contrast Adjustment (LCA) [14], Decomposition and Visual Adaptation (DVA) [17], the detail-richest-channel method (DRC) [23], the low-pass filtering and α -rooting (LPAR) [16], and the Double-pass Fundus Reflection model (DPFR) [31].

In actual implementation, the kernel parameters for DRC, LPAR and DPFR are determined adaptively according to the size of input images. Other parameters for DRC, LPAR and DPFR models such as intensity correction strength can be also automatically determined by the models. While for DVA method, it has two parameters, λ to determine the frequency band, and α to control the local contrast for the enhancement. In our implementation, $\lambda = 0.3$, $\alpha = 600$ based on [17]. All methods are tested on CATARACT dataset, and experimental results for visual assessments are shown in Fig. s5-5 to Fig. s5-7.

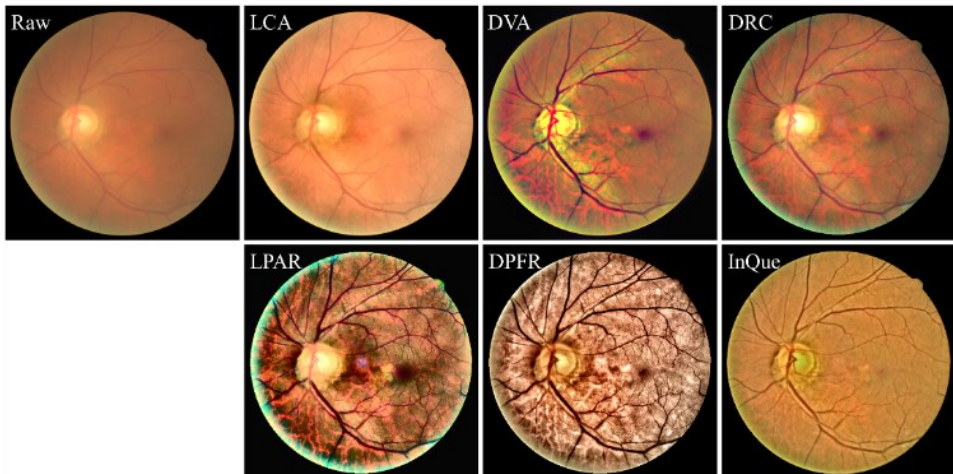


Fig. s5-5. Restoration results for image 001 in cataract dataset

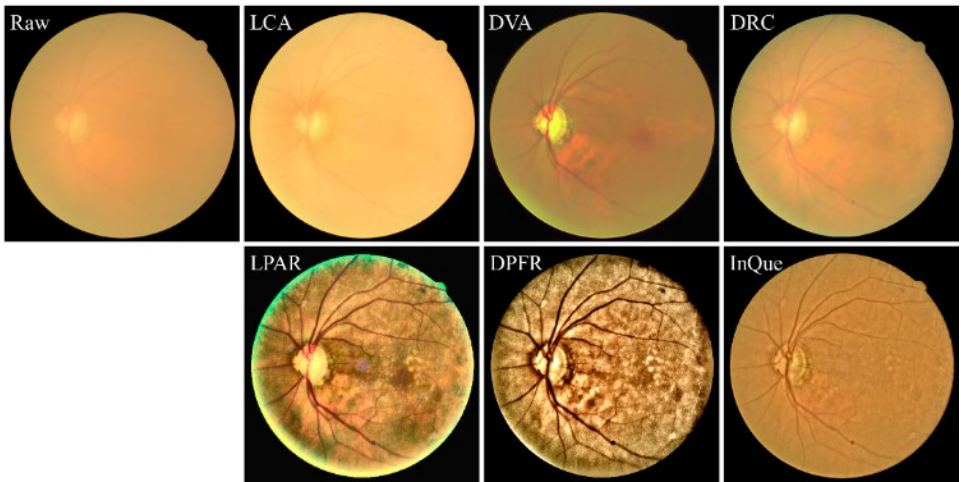


Fig. s5-6. Restoration results for image 021 in cataract dataset

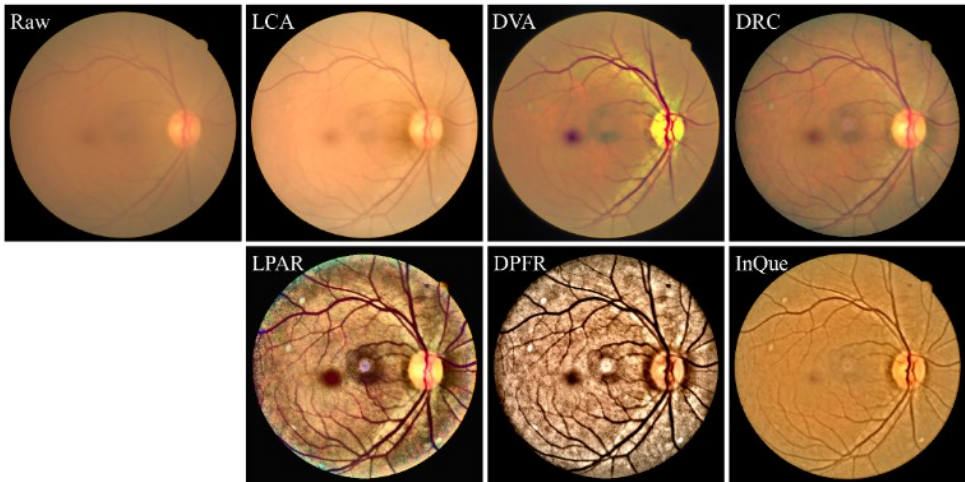


Fig. s5-7. Restoration results for image 096 in cataract dataset

Tab. s5-5. Objective assessments for images from CATARACT datasets for non-deep learning methods

Databases mean \pm std.	Matrixes	Raw	Methods					
			LCA	DVA	DRC	LPAR	DPFR	InQue
Cataract (100 images)	UIConM \uparrow	0.0689	0.1261	0.1411	0.1103	0.2767	0.3574	0.1758
		± 0.0163	± 0.0133	± 0.0376	± 0.0250	± 0.0346	± 0.0098	± 0.0293
	UISM \uparrow	1.2874	1.5189	1.7437	3.8750	4.5895	6.2512	4.3753
		± 0.1971	± 0.3059	± 0.6517	± 0.3056	± 0.6136	± 0.5429	± 0.2686
	IE \uparrow	5.8933	5.8931	6.8396	5.9145	6.3014	6.1507	6.1375
		± 0.2622	± 0.3915	± 0.2169	± 0.2169	± 0.1046	± 0.1159	± 0.2310
	C _{GRAMM} \uparrow	1.8005	2.7764	3.1611	2.5847	6.6565	8.4800	3.7292
		± 0.2608	± 0.2762	± 0.8805	± 0.5026	± 1.1019	± 0.5915	± 0.5667
	FADE \downarrow	0.4354	0.3324	0.7405	0.2584	0.1947	0.1336	0.1911
		± 0.0796	± 0.0791	± 0.2723	± 0.0553	± 0.0390	± 0.0109	± 0.0253

Statistical analysis are listed in Tab s5-5. In general, the traditional methods such as LPAR and DPFR significantly improve the contrast of the cataract retinal image, while distort the color naturalness. Methods such as DVA, DRC, and LPAR may cause overexposure effect at the bright area, optical disk, of the retinal image. Although the InQue cannot increase the contrast as high as that of LPAR and DPFR methods, as shown in Fig. s5-6, it maintains the naturalness and won't cause distortion of colors.

References

1. M. Zhou, K. Jin, S. Wang, J. Ye, and D. Qian, "Color Retinal Image Enhancement Based on Luminosity and Contrast Adjustment," *IEEE Transactions on Biomedical Engineering* (2018).
2. J. Wang, Y.-J. Li, and K.-F. Yang, "Retinal fundus image enhancement with image decomposition and visual adaptation," *Computers in Biology and Medicine* **128**, 104116 (2021).
3. L. Cao, and H. Li, "Detail-richest-channel based enhancement for retinal image and beyond," *Biomedical Signal Processing and Control* **69**, 102933 (2021).
4. L. Cao, H. Li, and Y. Zhang, "Retinal image enhancement using low-pass filtering and α -rooting," *Signal Processing* **170**, 107445 (2020).
5. S. Zhang, C. A. B. Webers, and T. T. J. M. Berendschot, "A double-pass fundus reflection model for efficient single retinal image enhancement," *Signal Processing* **192**, 108400 (2022).

Chapter 6

BDIC:

Boosting the performance of optical microscopy using imaging processing

Shuhe Zhang, Tos T. J. M. Berendschot, Meng Shao, and Jinhua Zhou

Proceedings Volume 12138, Optics, Photonics and Digital Technologies for Imaging Applications VII; 121380E (2022). DOI: 10.1117/12.2620781. SPIE Photonics Europe, 2022, Strasbourg, France

Abstract

Increasing imaging quality is of importance for microscopy. In this research, we proposed a universal method termed BDIC to boost the performance of an arbitrary optical microscopy system. The BDIC is an imaging processing technique, including blind deconvolution and none reference illumination correction. The only information that is needed for BDIC is a single raw image captured by a microscope, and the BDIC returns a blurriness-suppressed, illumination corrected enhanced images without any modification or prior knowledge of the imaging system. We demonstrate the BDIC experimentally on different samples, including USAF targets, plant root tissue, and human blood smears. Both subjective and objective assessments imply that the BDIC increases the image's visual quality including its contrast and signal-to-noise ratio.

1. Introduction

The imaging quality of a microscopy system is limited mainly by two optical subsystems: the illumination system and the detection system. In a traditional optical microscope, the illumination system aims to deliver a uniform and clear illumination pattern to illuminate the sample. While the detection system aims to form an ideal image of the sample and suppress the influence of noises.

However, no optical system is perfect. There are two agents happened in the collection and illumination system that are to be blamed for the microscopy imaging degrading. The first one is image blurriness. Blurriness is present in every image with different degrees during the imaging collection and degrades the image quality. The blurriness can be caused by optical aberrations, defocusing, accidental slight motion of the imaging platform or samples.

Fortunately, it is possible to compensate such blurriness agents as long as they can be measured or bypassed. Under the framework of adaptive optics [1], for example, the optical aberrations can be measured by wavefront sensor, and be compensated using deformable mirrors [2]. Moreover, in the framework of computational imaging, aberrations can be also bypassed using the lensless imaging method [3] including quantitative phase imaging or ptychography [4]. Meanwhile, the aberration can be measured indirectly using double-passing techniques [5, 6] or coded aperture methods like Fourier ptychography [7, 8] and other aperture synthesis methods. Defocusing and be regarded as one kind of optical aberration and thus can be corrected in the same methods. As for motion blur, it can be overcome by using coded illumination or coded exposure [9, 10].

Normally, the above-mentioned computational imaging methods rely on multiple capturing of the image of an identical sample [11, 12]. The illumination and detection system are modified to work cooperatively to produce the coded raw image data for digitally solving the clear images. Thus, these methods have special requirements to the

microscopy system and cannot be directly applied to an arbitrary optical microscopy system.

To provide a universal way of removing the blurriness of images, single-image blind deblurring is proposed [13, 14]. The blind deblurring is based on information theory and the Heavy-Tailed priors of blurred images [15]. In this case, both the raw clear image and the blur kernel are unknown, and the algorithm is aimed to find the best combination of the blur kernel and the estimation of a raw image that fits the Heavy-Tailed priors. These methods have been widely applied for medical imaging [16, 17], astronomical imaging [18, 19], remote sensing, photography deblurring [15, 20-24].

The biggest advantage of the blind deblurring methods is that they are pure mathematical methods. They have no special requirement for photographing devices and thus can be adapted for different situations. Experimental results of blind deblurring are promising in comparison with the raw blurred images but not as good as that of the computational imaging method. After all, the single blurred image is the only available information we have during the blind deblurring.

Another agent that causes the degrading of microscopy imaging is uneven illumination. Many factors, including misaligned optics, dust, nonuniform light sources, and vignetting, contribute to uneven illumination [25].

Correction of uneven illuminations for microscopy is complex, and thus their impacts are ignored in many cases of computational imaging or are carefully measured and celebrated before experiments. Severe uneven illuminations problems can be largely reduced using high-quality but expensive light sources like a laser, but mild uneven illuminations still exist.

Most illumination correction methods need reference images, and only a few research has reported on non-reference (single image) illumination correction [25, 26]. Even though, there is no report that shows the combination of single image deblurring and illumination correction on microscopy.

In order to fill this gap, in this study, we proposed an imaging processing method termed BDIC to boost the image quality of an arbitrary optical microscopy system. The BDIC method consists of two image restoration procedures - blind deconvolution (BD) and illumination correction (IC). Both procedures are based on a single image, which means that the only input of BDIC is the raw image captured during the microscopy. The rest of this article is organized as follows: in Section 2, the procedures of BDIC are detailed. The intermedia products in each step of BDIC show how the raw degraded image is becoming a deblurred, evenly illuminated image. In section 3, experimental results on different samples are demonstrated to show the robustness and universality of BDIC. Sections 4 and 5 are the discussion and conclusion.

2. Image restoration

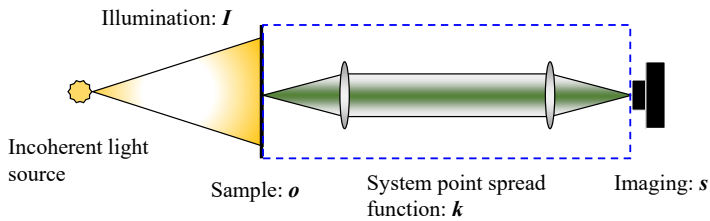


Fig. 6-1. Sketch for a typical microscopy system.

Assuming \mathbf{O} to be the 2D intensity distribution of a sample, and \mathbf{S} to be the gray-scaled image of \mathbf{O} captured by a monochromic camera. For a typical microscopy system with an incoherent illumination source as shown in Fig. 6-1, the image formation model of a thin sample can be expressed as [27-29]

$$\mathbf{S} = (\mathbf{O} \circ \mathbf{I}) \otimes \mathbf{p} + \boldsymbol{\varepsilon}, \quad (6-1)$$

where \mathbf{I} is the illumination intensity pattern projected by the illumination system, and \mathbf{p} is the shift-invariant point spread function (PSF) of the imaging system. $\boldsymbol{\varepsilon}$ is the noise term imposed on the captured image \mathbf{S} . \circ denotes the pixel-wise matrix multiplication (Hadamard product), and \otimes denotes the 2D-convolution.

Equation (6-1) implies that the degrading of a microscopy image is due to three aspects: (1) the unsatisfactory illumination condition such as uneven illumination and insufficient illumination intensity; (2) the blurriness caused by the system's point spread function (PSF) including defocusing, inadvertently motion, and optical aberrations; and (3) noises. On the other hand, a good image of \mathbf{O} can be recovered digitally using Eq. (6-1) if those degrading conditions are measured. Our goal is to solve a good estimation of \mathbf{O} from known only of \mathbf{S} .

2.1 Blind deconvolution

The BDIC consists of two procedures. The first procedure is blind deconvolution. Here we use blind deconvolution under patch-wise minimal pixels regularization [23]. Let $\mathbf{Q} = \mathbf{O} \cdot \mathbf{I}$, the blind deconvolution is transferred to an optimization problem which is

$$\arg \min_{\mathbf{Q}, \mathbf{p}} \|\mathbf{Q} \otimes \mathbf{p} - \mathbf{S}\|_2^2 + \gamma \|\mathbf{p}\|_2^2 + \beta \|\nabla \mathbf{Q} - \mathbf{G}\|_2^2 + \mu \|\mathbf{G}\|_0 + \alpha \|\mathcal{P}(\mathbf{Q})\|_0, \quad (6-2)$$

where γ , β , μ , and α are positive weight parameters for Lagrange multipliers. \mathbf{G} is an auxiliary variable for solving the regularization imposed on $\nabla \mathbf{Q}$. $\mathcal{P}(\mathbf{Q})$ is the patch-wise minimal pixels of \mathbf{Q} . For an arbitrary matrix \mathbf{X} , $\mathcal{P}(\mathbf{X})$ is given by

$$\mathcal{P}(\mathbf{X})(x, y) = \min_{(x_0, y_0) \in \Omega(x, y)} [\mathbf{X}(x_0, y_0)], \quad (6-3)$$

which first divides the matrix \mathbf{X} into many squared patches centered at (x, y) with the width of w pixels. Pixels in one small patch are denoted by $\Omega(x, y)$ and function \mathcal{P} extracts the single pixel that has the minimum value in the corresponding patch. Assuming the size of \mathbf{X} is M by N , the number of squared patches is given by $\lceil M/w \rceil \times \lceil N/w \rceil$, and $\lceil \cdot \rceil$ denotes ceiling. In this research the size of each patch is $\lceil 0.0125(M+N) \rceil$.

Eq. (6-2) implies two priors imposed on the none-blurred images: (1) The L_0 -norm of the gradient of a none-blurred image should be a sparse matrix, and (2) The L_0 -norm

of the patch-wise minimal pixels should also be a sparse matrix. Eq. (6-2) is solved in an iterative manner. Since it is not the aim of this research, readers may see [23] for its detailed description of solving for Eq. (6-2).

Fig. 6-2 demonstrated one of the results for the blind deconvolution, where a USAF resolution target is imaged by a microscopy system with uneven haze-like illumination. The blind deconvolution generates an estimation of the blind kernel as shown in Fig. 6-2 (b), and the deconvoluted image as shown in Fig. 6-2 (c) for upcoming illumination correction.

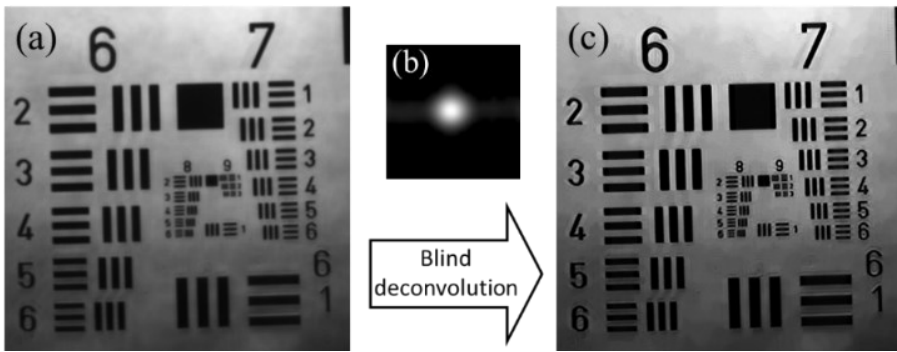


Fig. 6-2. The output of blind deconvolution procedure. (a) Raw image. (b) estimated kernel. (c) deconvoluted image.

2.2 Illumination correction

The second procedure of BDIC is illumination correction. Here, we use our own developed illumination correction algorithm, which is a combination of the Retinex [30] and the dark-channel prior [31] theories. The output of Eq. (6-2) includes Q which is a deblurred image but may still suffer from uneven or insufficient illumination conditions. Q is rewritten as

$$Q = I_{coarse} \circ I_{fine} \circ O, \quad (6-4)$$

where I_{coarse} and I_{fine} denote the coarse illumination pattern (slowly varying in the spatial domain) and fine illumination pattern (fast varying in the spatial domain)

projected on the sample, respectively. The coarse illumination pattern denotes the overall illumination pattern provided by the light source, while the fine illumination pattern can be due to the intensity fluctuations caused by floating dust.

Correction of I_{coarse} is achieved using the Retinex theory [30]. Since the I_{coarse} is slowly varying, a good estimation of I_{coarse} is given by low-pass filtering of \mathbf{Q} , and thus can be removed by [30]

$$\mathbf{Q}_0 = \mathcal{N} \left[\exp \left(\log(\mathbf{Q} + \xi) - \log \left\{ \mathbf{Q} \otimes \left[\frac{1}{2\pi\omega^2} \exp \left(-\frac{\mathbf{r}^2}{2\omega^2} \right) \right] + \xi \right\} \right) - \xi \right], \quad (6-5)$$

where $\mathbf{Q}_0 = I_{fine} \cdot \mathbf{O}$, and ξ is a small value to avoid the situation of $\log(0)$. ω is the width of the Gaussian kernel and is given by $\omega = \lceil \sqrt{MN} / 10 \rceil$, $\mathcal{N}(\mathbf{X})$ denotes the normalization function that normalizes the gray-value of the image \mathbf{X} into $[0, 1]$.

Fig. 6-3 demonstrates the coarse illumination correction results. As shown in Fig. 6-3 (a), pixels on the upper side of the image have a higher intensity than that of on the lower side of the image. After the coarse illumination correction, as shown in Fig. 6-3 (c), the overall image is in the same intensity level but still has a black-cloud effect.

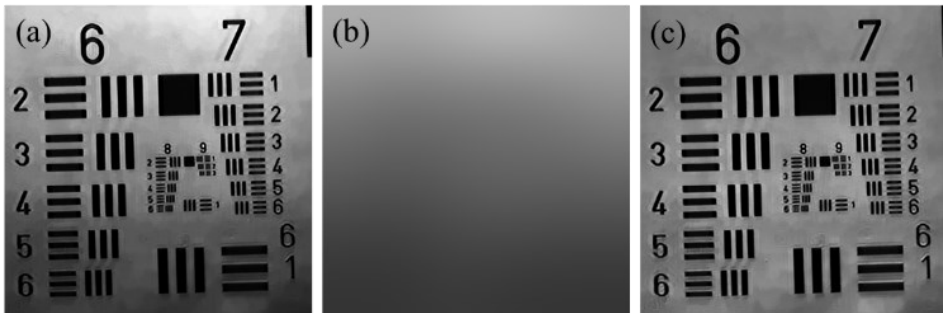


Fig. 6-3. Output of coarse illumination correction procedure. (a) Raw image in Fig. 2(c). (b) estimated coarse illumination pattern. (c) coarse illumination corrected image.

Correction of I_{fine} is achieved using dark-channel prior [32, 33], and \mathbf{O} is given by

$$\mathbf{O} = 1 - \frac{(1 - \mathbf{Q}_0) + \sigma \mathcal{D}(1 - \mathbf{Q}_0)}{1 - \sigma \mathcal{D}(1 - \mathbf{Q}_0)}. \quad (6-6)$$

$\mathcal{D}(1 - \mathbf{Q}_0)$ is the dark channel of $1 - \mathbf{Q}_0$ and is refined by the guided filter. The dark channel of a gray-scale image \mathbf{I} estimated in a local neighborhood $\Omega(\mathbf{r})$ with the size of w pixels is equal to filtering the image using the local minimum filter. $w = 5$ in this research.

The parameter $\sigma \in [0, 1]$ is used to control the degree of illumination correction. A large value of α will lighten the grayscale image and may also cause an overexposure effect where the pixel value is larger than 1. An appropriate value of σ is important for ensuring good enhancement results, we, therefore, use an adaptive manner to determine the value of σ .

To determine the value of α with low complexity but high accuracy, a global wise optimization function is designed as

$$\sigma = \arg \min \left[\frac{1}{MN} \sum_{y=1}^N \sum_{x=1}^M \mathbf{O}(x, y) - 0.65 \right]^2, \quad (6-7)$$

Equation (6-7) is a one-dimensional optimization function. In this work, the Fibonacci method (FM) is adopted to solve Eq. (6-7) since it is able to gradually narrow the search interval for a one-dimensional optimization problem until the convergence condition is satisfied [34].

Fig. 6-4 demonstrates the find illumination correction results. As shown in Fig. 6-3 (c), the black-cloud effect is significantly suppressed.

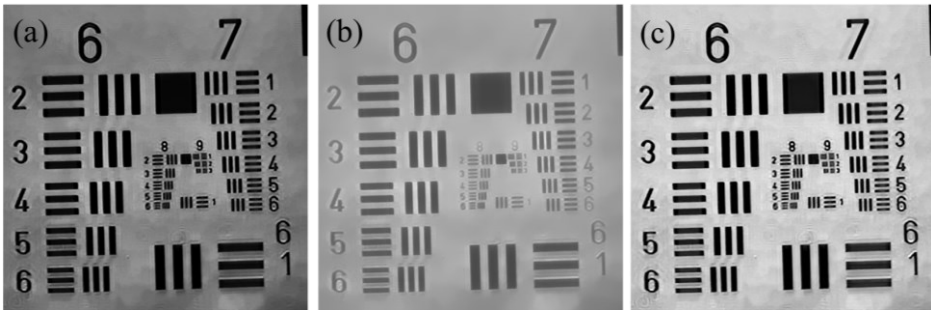


Fig. 6-4. The output of fine illumination correction procedure. (a) Raw image in Fig. 6-3 (c). (b) estimated find illumination pattern. (c) fine illumination corrected image.

3. Experimental results

In the following experiments, the size of the blur kernel is 51 by 51 pixels, and $\mu = 1 \times 10^{-3}$. Other parameters for blind deconvolution are consistent with those in [23]. The raw images are captured by a self-built microscope and a commercial Olympus microscope. Parameters of two microscopy systems are not recorded since we don't need this information for the BDIC algorithm.

3.1 USAF Target

First, we applied the BDIC to the imaging of USAF targets. Each image is obtained under unknown different illumination conditions and slightly defocusing as shown in Fig. 6-5 (a1) to (d1). Here, only Fig. 6-5 (c1) has even illumination, and the other three images suffer from uneven illuminations or insufficient illuminations (Fig. 6-5 (d1)). An overview of the BDIC enhanced images is shown in Fig. 6-5 (a2) to (d2). One can find that all images are at a similar intensity illumination level.

The partially enlarge images for Groups 8 and 9 in the USAF target are shown in Fig. 6-6. The BDIC is able to correct the illumination and suppress the blurriness of the raw image without loss of resolution. The visual quality of raw images is significantly enhanced. In some cases, as shown in Fig. 6-6 (c2) and (d2), the resolutions are slightly increased as due to the effect of deblurring and illumination correction.

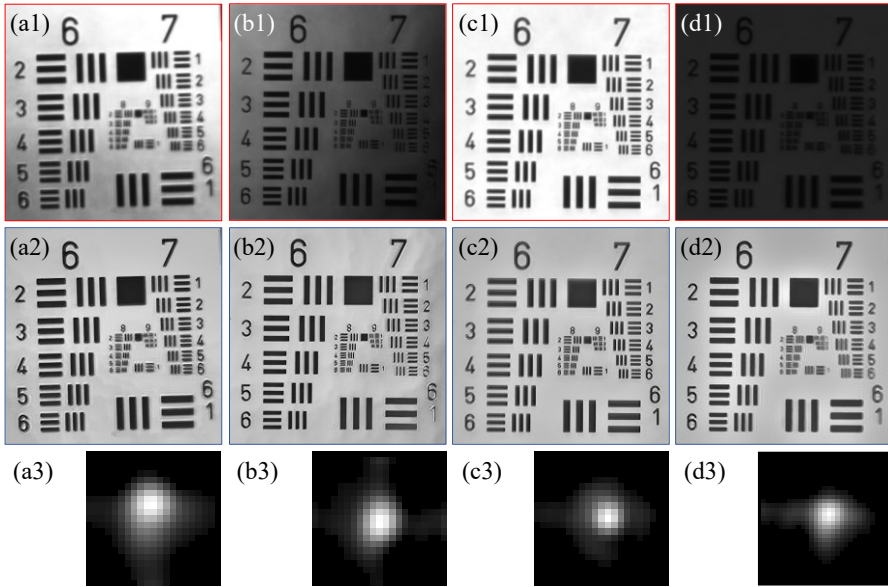


Fig. 6-5. Output of the BDIC on USAF target. (a1) to (d1) are raw images. (a2) to (d2) are enhanced results of BDIC. (a3) to (d3) are estimated blurring kernels.

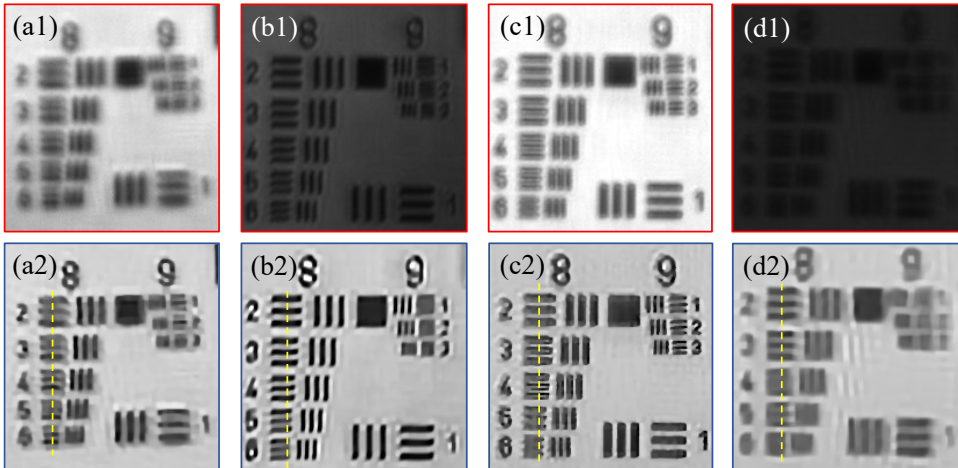


Fig. 6-6. Enlarged parts for groups 8 and 9. (a1) to (d1) are raw images. (a2) to (d2) are the output of BDIC.

Pixels values along the yellow lines in Fig. 6-6 are plotted in Fig. 6-7 in comparison with the raw images. Our BDIC method increases the contrast of the image as the distance between the peaks (pixels with large gray value) and valleys (pixels with small gray value) after BDIC treatments becomes larger than the raw images, especially for Fig. 6-6 (d1) and (d2). Moreover, the blind deblurring process also increases the resolution of the images.

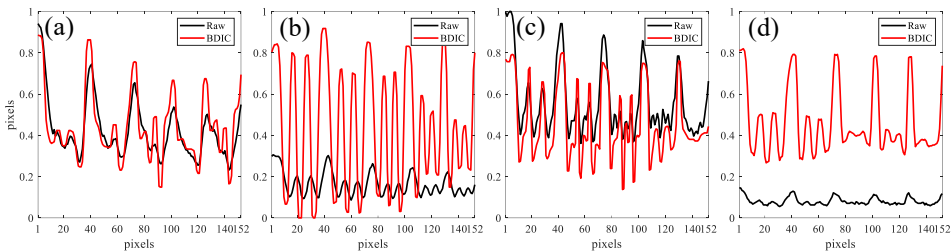


Fig. 6-7. Plots for pixels values along the yellow line in Fig. 6. Black lines are pixels values in the raw images, while red lines are pixels values after BDIC treatment.

For objective assessment, we collect 121 images under different imaging conditions, for example, imaging the same USAF target using an objective lens with different NA and magnifications. The illumination condition for each image is also diverse. We have also collected the images after careful adjustment of the imaging platform.

Tab. 6-1. The average value of objective assessment. A total of 121 images are collected.

	PSNR	SSIM	IE	C _{rmm}
Raw image	11.071	0.585	4.238	8.503
BDIC enhanced	15.501	0.673	4.327	16.809

These images are regarded as ground truth or reference images for the calculation of peak signal-to-noise ratio (PSNR) and structural similarity (SSIM). For none reference measurement, we use image entropy (IE) and image multiscale contrast (C_{rmm}) [35]. The average values are listed in Tab. 6-1, where the PSNR and C_{rmm} are

significantly increased after BDIC enhancement. The SSIM of the enhanced image is also larger than that of raw images.

3.2 Plant root tissue

We have tested the performance of BDIC on plant root tissues. The plant root tissue has a more intricate structure than that of USAF. A wide-field image of plant root tissue is obtained under uneven illumination and slightly defocusing as shown in Fig. 6-8 (a1). The BDIC output is shown in Fig. 6-8 (a2) where the uneven illumination is corrected. The image in Fig. 6-8 (a3) is captured under careful system celebration and good illumination conditions, which can be regarded as the ground truth of Fig. 6-8 (a1).

Partially enlarged images in the yellow, red, and green boxes are shown in Fig. 6-8 (b1) to (d3), respectively, to demonstrate the blind deconvolution results. Experimental results show that the BDIC is able to correct the uneven illumination over the entire field of view, and the deblurring results are also consistent with that of the ground truth image.

3.3 Human blood smear

We have tested the performance of BDIC on human blood smears as shown in Fig. 6-9 (a1). The raw image is 1920 by 1200 pixels and is divided uniformly into 5 by 5 subsets (384 by 240 pixels for each subset). The blurring kernel is estimated separately in each subset as shown in Fig. 6-9 (b). The shapes of the blurring kernel are similar to each other implying that the imaging system suffers less from spatial-variant point spread function.

Afterward, the illumination correction is applied to the entire image. The output of BDIC is shown in Fig. 6-9 (a2), where the illumination condition becomes good, and uniform. Enlarged parts in yellow boxes are shown in Fig. 6-9 (c1) and (c2).

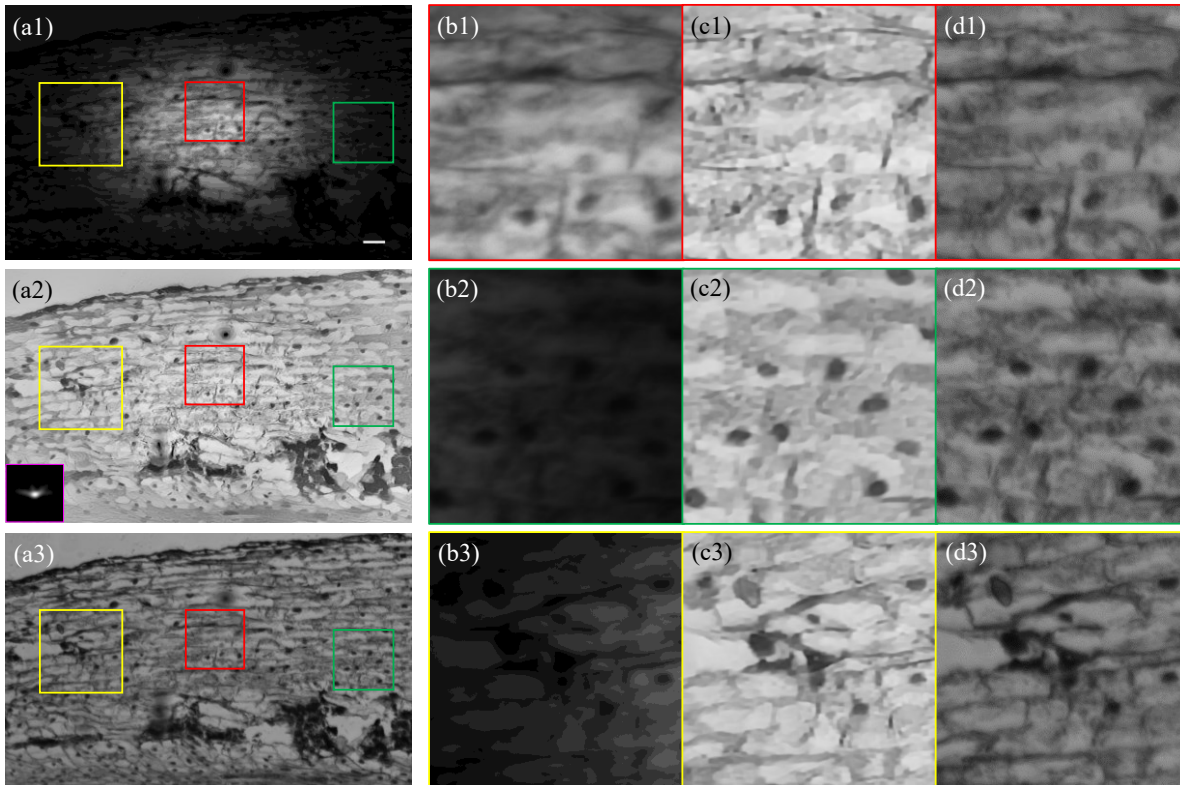


Fig. 6-8. Experimental results on plant root tissue. (a1) captured raw image. (a2) BDIC enhanced image. (a3) images with good focusing position and illumination condition. (b1) to (b3) are enlarged parts in corresponding boxes for (a1). (c1) to (c3) are enlarged parts in corresponding boxes for (a2). While (c1) to (c3) are enlarged parts (a3). The scale bar is 10 μm .

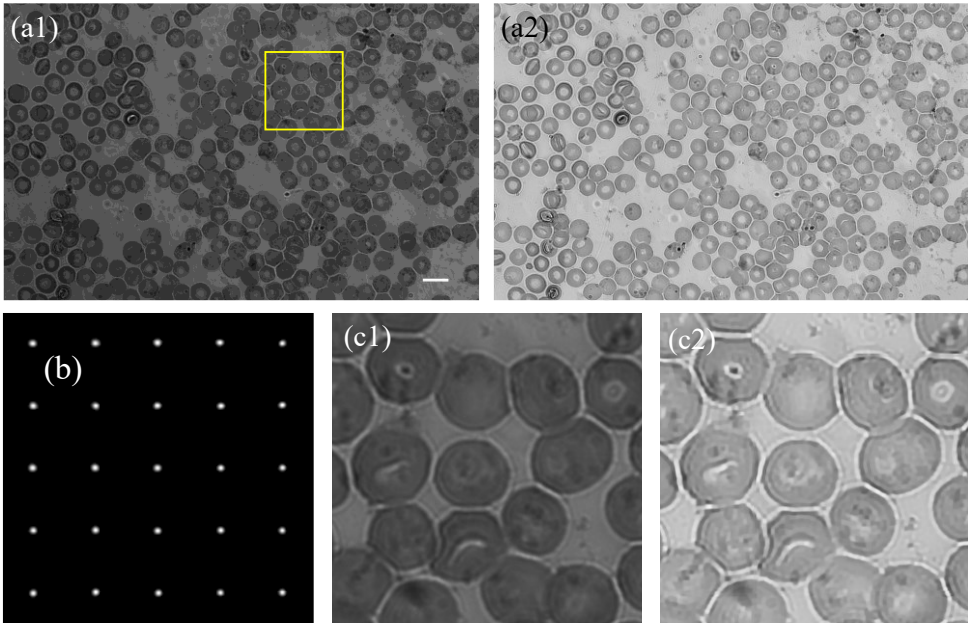


Fig. 6-9. Experimental results on the human blood smear. (a1) captured raw image. (b) Estimated kernel. (a2) BDIC output. (c1) and (c2) are enlarged parts in the yellow of (a1) and (a2), respectively. The scale bar is 10 μm .

3.4 With motion blurring

In some cases, there would be motion blurring imposed during the imaging procedure, due to unstable imaging platform, and accidental vibration of the sample. These problems usually happen to low-cost microscopy systems. Our BDIC can correct the motion blurring as well. We test the BDIC on the USAF target and the blood smear. The stage is driven and moved by a piezoelectric platform in a random direction to create motion blurring as shown in Fig. 6-10 (a). The stripe pattern is blurred due to the motion.

The BDIC output is shown in Fig. 6-10 (b) where a clear, none motion blurring image is recovered, together with the blurring kernel shown in Fig. 6-10 (c).

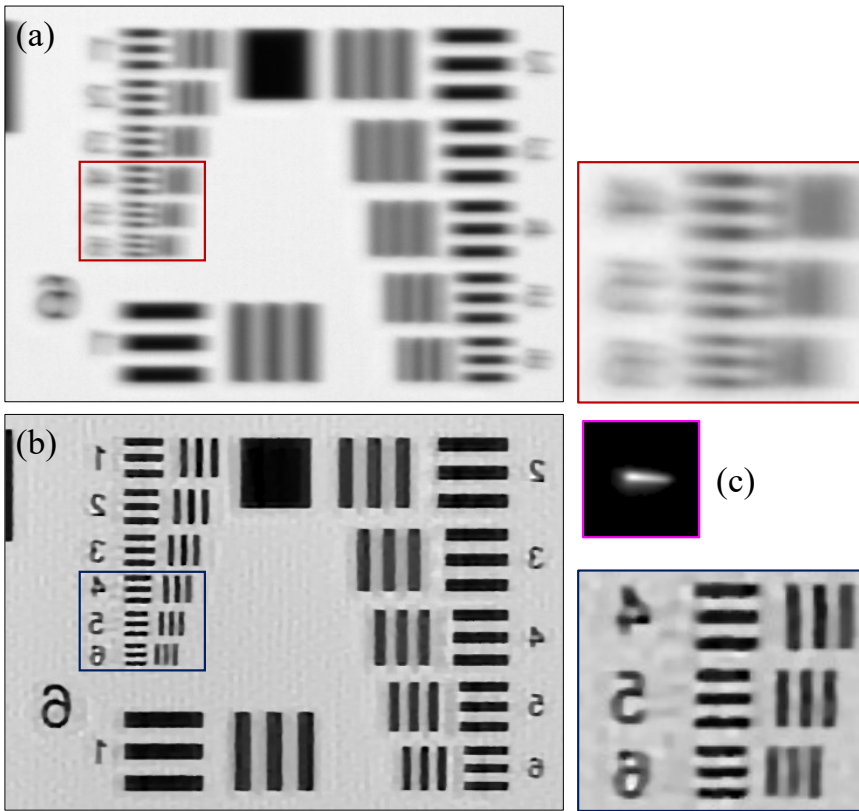


Fig. 6-10. Recovery of motion blurring on the USAF target. (a) captured raw image. (b) BDIC output. (c) Estimated kernel.

Experiments on blood smears are shown in Fig. 6-11. Images on the first row of Fig. 6-11 are the captured raw images, where they are severely blurred due to the sudden motion during the exposure. The BDIC recovery results are shown in the second row of Fig. 6-11 for corresponding images, where the motion blurriness is removed. Note that if the raw image contains objects that are stationary while the sample is moving, for example, dust particles on the lens, the objects will cause artifacts due to the deconvolution as shown in the red circles in Fig. 6-11 (c2).

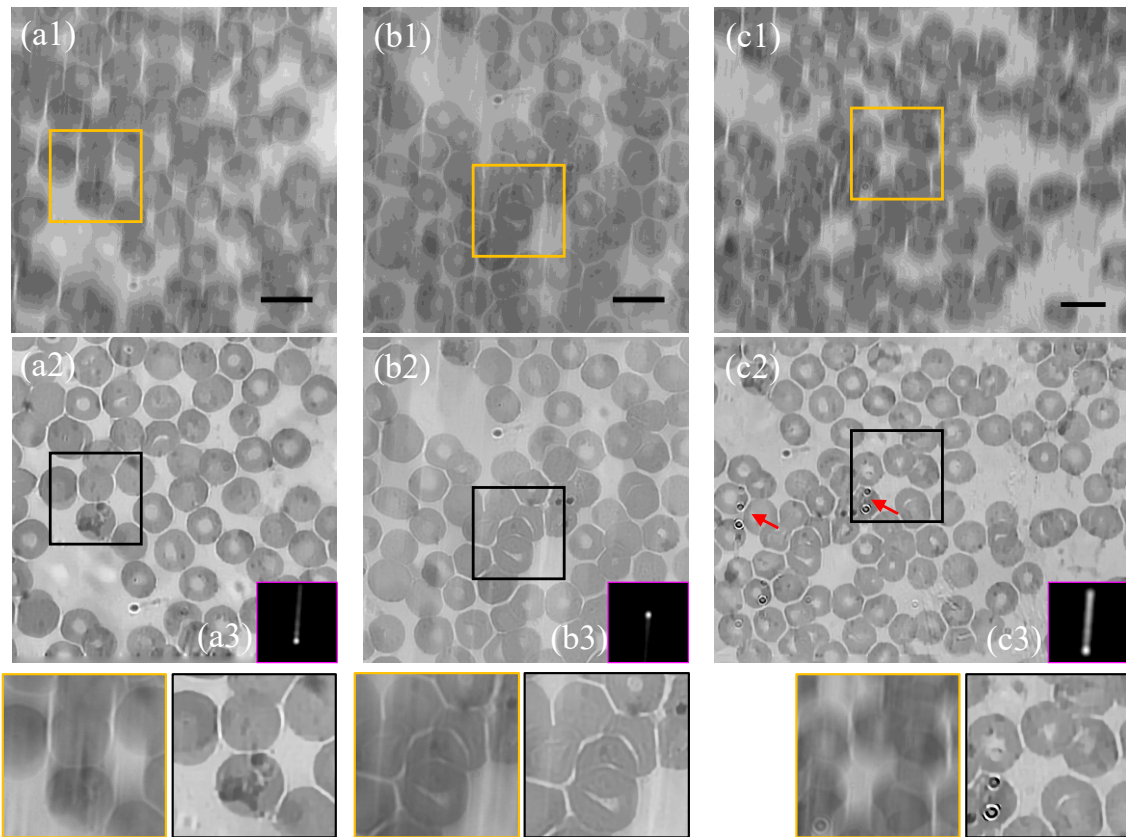


Fig. 6-11. Recovery of motion blurring on the blood smear. (a1) to (c1) captured raw image. (a2) to (c2) BDIC outputs. (a3) to (c3) Estimated kernel. The scale bar is 10 μm .

4. Discussion

The BDIC is a purely mathematical method that doesn't involve the modification of the optical microscopy system. This feature makes it very easy to be deployed and performed among different type of microscopy. The BDIC is divided into two parts, blind deconvolution and illumination correction. In our method, the blind deconvolution is achieved by patch-wise minimal pixels regularization, however, it should be noted that other blind deblurring methods are also available such as local maximum gradient regularization [21], and patch-wise maximum gradient regularization [36]. There are two reasons for choosing patch-wise minimal pixels regularization in this research. First, the patch-wise minimal pixels (PMP) regularization has a low-memory usage requirement [23]. The PMP matrix $\mathcal{P}(\mathbf{X})$ is a sparsely sampled version of raw images \mathbf{X} , which makes it easy in programming. Moreover, due to the sparsity of $\mathcal{P}(\mathbf{X})$, the execution time is also short which makes this method suitable for running on a personal computer. The second reason is that the PMP is a special case of dark-channel [20]. By using the PMP, the deconvolution part and the illumination part can be organized into an identical framework.

The reliability of the BDIC is tested by comparing its results with the ground truth images captured by the carefully calibrated imaging system. Experimental results show that the BDIC can improve the imaging quality with different degrees depending on the quality of a raw image. Note that BDIC is an imaging processing method, the only thing that it can do to make full use of the information provided by the raw image. Thus, it cannot recover the information that does not exist on the raw images, and may also enhance the noise signal as mentioned in Fig. 6-11 (c2). However, the single image requirement, cheap usage, and universal for different microscopy systems are unique advantages of the BDIC. In conclusion, the BDIC aims to provide a simple and fast solution to enhance the performance of microscopy under limited conditions and may also support the results computation imaging before and after the image reconstruct.

Although both blind deconvolution and illumination correction can be achieved through deep-learning methods [23, 37, 38], the requirement of training data and massive computational resources largely limit their application, and their inner mechanism of each layer of the network and intermedial products are also unknown. On the contrary, no training data is needed for BDIC, and it is a transparent algorithm. The procedure about how a degraded raw image becomes the better one can be shown step by step.

5. Conclusion

In conclusion, we have developed an imaging processing algorithm that suppresses the blurriness and corrects the uneven illumination of microscopy images. The proposed image formation model together with the BDIC algorithm perfectly combines the framework of blind deblurring (BD) and illumination correction (IC). To prove the effectiveness of the BDIC, we test it on different microscopy samples. The experimental results show that our algorithm greatly improves the imaging performance. Due to its simplicity and universality, the BDIC has great potential applications

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding. This research is supported by China Scholarship Council (CSC) (201908340078)

References

1. B. C. Platt, and R. Shack, "History and Principles of Shack-Hartmann Wavefront Sensing," *J. Refract. Surg.* **17**, S573-577 (2001).
2. J. Liang, D. R. Williams, and D. T. Miller, "Supernormal vision and high-resolution retinal imaging through adaptive optics," *J. Opt. Soc. Am. A* **14**, 2884-2892 (1997).
3. A. Ozcan, and E. McLeod, "Lensless Imaging and Sensing," *Annual Review of Biomedical Engineering* **18**, 77-102 (2016).
4. J. M. Rodenburg, and H. M. L. Faulkner, "A phase retrieval algorithm for shifting illumination," *Appl. Phys. Lett.* **85**, 4795-4797 (2004).
5. P. Artal, I. Iglesias, and N. L'opez-Gil, "Double-pass measurements of the retinal-image quality with unequal entrance and exit pupil sizes and the reversibility of the eye's optical system," *J. Opt. Soc. Am. A* **12**, 2358-2366 (1995).
6. I. Iglesias, N. L'opez-Gil, and P. Artal, "Reconstruction of the point-spread function of the human eye from two double-pass retinal images by phase-retrieval algorithms," *J. Opt. Soc. Am. A* **15** (1998).
7. J. Chung, G. W. Martinez, K. C. Lencioni, S. R. Sadda, and C. C. Yang, "Computational aberration compensation by codedaperture- based correction of aberration obtained from optical Fourier coding and blur estimation," *Optica* **6**, 647-661 (2019).
8. G. Zheng, R. Horstmeyer, and C. Yang, "Wide-field, high-resolution Fourier ptychographic microscopy," *Nat. Photonics* **7**, 739-745 (2013).
9. R. Raskar, A. Agrawal, and J. Tumblin, "Coded exposure photography: motion deblurring using fluttered shutter," *ACM SIGGRAPH 2006 Papers*, 795-804 (2006).
10. A. Agrawal, and Y. Xu, "Coded exposure deblurring: Optimized codes for PSF estimation and invertibility," *2009 IEEE Conference on Computer Vision and Pattern Recognition* (2009).
11. O. Cossairt, M. Gupta, and S. K. Nayar, "When Does Computational Imaging Improve Performance," *IEEE Transactions on Image Processing* **22**, 447 - 458 (2013).
12. J. N. Mait, G. W. Euliss, and R. A. Athale, "Computational imaging," *Advances in Optics and Photonics* **10**, 409-483 (2018).
13. R. Fergus, B. Singh, A. Hertzmann, S. T. Roweis, and W. T. Freeman, "Removing camera shake from a single photograph," in *ACM SIGGRAPH 2006 Papers*(2006), pp. 787-794.
14. A. Levin, Y. Weiss, F. Durand, and W. T. Freeman, "Understanding and evaluating blind deconvolution algorithms," in *2009 IEEE Conference on Computer Vision and Pattern Recognition*(IEEE2009), pp. 1964-1971.
15. J. Kotera, F. Šroubek, and P. Milanfar, "Blind Deconvolution Using Alternating Maximum a Posteriori Estimation with Heavy-Tailed Priors " *International Conference on Computer Analysis of Images and Patterns* 59-66 (2013).
16. Z. Dong, H. Jing, Z. Hua, Z. Bian, S. Niu, Z. Zhang, Q. Feng, W. Chen, and J. Ma, "Spectral CT image restoration via an average imageinduced nonlocal means filter," *IEEE Trans. Biomed. Eng.* **63**, 1044-1057 (2016).
17. U. Qidwai, and U. Qidwai, "Blind Deconvolution for Retinal Image Enhancement," *IEEE EMBS Conference on Biomedical Engineering & Sciences* (2010).
18. W. Gao, X. Zhao, J. Zou, Y. Yang, R. Xu, R. Zhang, and X. Xuebin, "Parametric blur estimation for blind restoration of atmospherically degraded images: Class G," *Opt. Rev.* **24**, 1-13 (2017).
19. Y. V. J. A. o. Zhulina, "Multiframe blind deconvolution of heavily blurred astronomical images," **45**, 7342-7352 (2006).

20. J. Pan, D. Sun, H. Pfister, and M.-H. Yang, "Blind Image Deblurring Using Dark Channel Prior," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2016).
21. L. Chen, F. Fang, T. Wang, and G. Zhang, "Blind Image Deblurring with Local Maximum Gradient Prior," 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2019).
22. L. Huang, Y. Xia, and T. Ye, "Effective Blind Image Deblurring Using Matrix-Variable Optimization," IEEE Trans. Image Process **30** (2021).
23. F. Wen, R. Ying, P. Liu, and T.-K. Truong, "A Simple Local Minimal Intensity Prior and An Improved Algorithm for Blind Image Deblurring," IEEE Trans Circuits and Systems for Video Technology **3**, 2923-2937 (2021).
24. H. Zhou, Y. Chen, H. Feng, G. Lv, Z. Xu, and Q. Li, "Rotated rectangular aperture imaging through multi-frame blind deconvolution with Hyper-Laplacian priors," Opt. Express **29**, 12145-12159 (2021).
25. K. Smith, Y. Li, F. Piccinini, G. Csucs, C. Balazs, A. Bevilacqua, and P. Horvath, "CIDRE: an illumination-correction method for optical microscopy," Nature Method **12**, 404-410 (2015).
26. J. Yu, J. Zhao, H. Shao, Y. Lu, Y. He, and L. Zhang, "Illumination compensation for microscope images based on illumination difference estimation," The Visual Computer (2021).
27. M. Born, and E. Wolf, *Principles of Optics* (Cambridge University Press, 1999).
28. J. W. Goodman, *Introduction to Fourier optics* (Roberts and Company Publishers, 2005).
29. G. Zheng, *Fourier Ptychographic Imaging A MATLAB® tutorial* (Morgan & Claypool Publishers, 2016).
30. D. J. Jobson, Z.-u. Rahman, and G. A. Woodell, "Properties and Performance of a Center/Surround Retinex," IEEE Trans. Image Process. **6**, 451-462 (1997).
31. K. He, J. Sun, and X. Tang, "Single Image Haze Removal Using Dark Channel Prior," (2009).
32. Y. Wang, S. Zhuo, D. Tao, J. Bu, and N. Li, "Automatic local exposure correction using bright channel prior for under-exposed images," Signal Processing **93**, 3227-3238 (2013).
33. B. Savelli, A. Bria, A. Galdran, C. Marrocco, M. Molinara, A. e. Campilho, and F. Tortorella, "Illumination correction by dehazing for retinal vessel segmentation," IEEE 30th International Symposium on Computer-Based Medical Systems (2017).
34. M. Ju, C. Ding, Y. J. Guo, and D. Zhang, "IDGCP: Image Dehazing Based on Gamma Correction Prior," IEEE Trans. Image Process. **29**, 3104 - 3118 (2021).
35. A. Rizzi, T. Algeri, G. Medeghini, and D. Marini, "A proposal for Contrast Measure in Digital Images," (2004).
36. Z. Xu, H. Chen, and Z. Li, "Fast blind deconvolution using a deeper sparse patch-wise maximum gradient prior," Signal Processing: Image Communication **90**, 116050 (2021).
37. J. Sun, W. Cao, Z. Xu, and J. Ponce, "Learning a convolutional neural network for non-uniform motion blur removal," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*(2015), pp. 769-777.
38. O. Kupyn, T. Martyniuk, J. Wu, and Z. Wang, "Deblurgan-v2: Deblurring (orders-of-magnitude) faster and better," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*(2019), pp. 8878-8887.

Chapter 7

LR²L:

Luminosity rectified blind Richardson-Lucy deconvolution for single retinal image restoration

Shuhe Zhang, Carroll A. B. Webers, and Tos T. J. M. Berendschot

Computer Methods and Programs in Biomedicine, 2023 (229), pp 107297.

Doi: [10.1016/j.cmpb.2022.107297](https://doi.org/10.1016/j.cmpb.2022.107297).

Abstract

Background and Objective: Due to imperfect imaging conditions, retinal images can be degraded by uneven/insufficient illumination, blurriness caused by optical aberrations and unintentional motions. Degraded images reduce the effectiveness of diagnosis by an ophthalmologist. To restore the image quality, in this research we propose the luminosity rectified Richardson-Lucy (LRRL) blind deconvolution framework for single retinal image restoration. **Methods:** We established an image formation model based on the double-pass fundus reflection feature and developed a differentiable non-convex cost function that jointly achieves illumination correction and blind deconvolution. To solve this non-convex optimization problem, we derived the closed-form expression of the gradients and used gradient descent with Nesterov-accelerated adaptive momentum estimation to accelerate the optimization, which is more efficient than the traditional half quadratic splitting method. **Results:** The LRRL was tested on 1719 images from three public databases. Four image quality matrixes including image definition, image sharpness, image entropy, and image multiscale contrast were used for objective assessments. The LRRL was compared against the state-of-the-art retinal image blind deconvolution methods. **Conclusions:** Our LRRL corrects the problematic illumination and improves the clarity of the retinal image simultaneously, showing its superiority in terms of restoration quality and implementation efficiency.

Keywords

Retinal image; Image restoration; Blind deconvolution; Illumination correction; Adaptive momentum estimation.

1. Introduction

Retinal imaging is widely used by ophthalmologists all over the world, where it plays a key and fundamental role in the diagnosis and management of ophthalmologic disorders [1], such as diabetic retinopathy [2, 3], glaucoma [4], and age-related macular degeneration [5, 6]. Moreover, since the retina has a high metabolism rate and is the place in the human body where the blood vessels and neuro branches can be non-invasively observed in-vivo, some psychiatric diseases including schizophrenia and bipolar disorder can be detected according to the tortuosity features of the blood vessels and neurons [7, 8]. Certain cardiovascular risk factors can be also determined based on the morphology of retinal blood vessels [9, 10] and optic cup-optic disk ratio [11]. Retinal oxygen saturation can be measured using retinal oximetry [12] and biomarkers for early Alzheimer's disease diagnosis can be found using hyperspectral retinal imaging [13].

Diagnosing efficiency and precision are deeply related to the quality of retinal imaging. However, not every retinal image is perfect and low-quality image occurrence is not a minor fact. Heaven et al. found 9.5% of all acquired images to be entirely unsatisfactory in a prospective study of 981 diabetic retinopathy patients [14]. Scanlon and Stephen found the ungradable image rate to be between 19.7% for nonmydriatic photography and 3.7% for mydriatic photography study of 3650 diabetic patients [15].

Retinal images can be severely degraded by opacities in the optical media of cataract eyes [16-18], retinal images for non-cataract subjects can as well be degraded by poor illumination conditions including uneven or insufficient illuminations. The quality of retinal imaging can be improved either by including high-end fundus cameras and using adaptive optics to tackle the optical aberrations [19], or by using image enhancement processing to correct for illumination artifacts [20-22], to enhance contrast [23-25] and to use dehazing algorithms [26-28]. Most of these methods aim to improve the contrast of the retinal image by adjusting pixel values, either using histogram adjustment or using image formation models (IFM). For the IFM-based methods, since they digitally

inverse the progress of how a degraded image was formed, they are also known as retinal image restoration methods.

In addition, retinal images can be also degraded by optical aberrations such as defocus, and motion blur caused by unintended movement of a subject's head. Such degenerations are modeled by a convolution process between the ideal retinal image and a blur kernel, and cannot be restored using the contrast enhancement methods mentioned above. To tackle the blurriness degenerations, the blind deconvolution method is needed as the kernels are usually unknown. However, according to our investigation, only few studies have reported on single retinal image blind deconvolution [29-31], which rather aimed to correct the blurriness caused by aberrations and motions during the retinal imaging.

Image blind deconvolution has been greatly developed and is mainly used for natural scene image deburring. Much prior knowledge, including but not limited to the heavy-tail prior [32, 33], the dark-channel prior [34] and the local maximum gradient prior [35], has been explored to facilitate the single image blind deconvolution tasks. Nevertheless, blind deconvolution for retinal images is still problematic and challenging since there are a large number of retinal images suffering from poor illumination conditions which hide the structure (edge) information that is essential for proper deconvolutions.

Traditional methods consider the illumination correction and blind deconvolution as two unrelated problems. Both problems are ill-posed since the number of unknown parameters is more than known parameters. Andrés et al. proposed a two-steps retinal image blind deconvolution method [30], in which the first step is estimating and compensating for the uneven illumination using a fourth-order polynomial. The second step is blind deconvolution with TV-regularization corresponding to the Heavy-tail-prior to natural scene deburring. However, this method requires at least two paired retinal images of one identical subject. Francisco et al. limit the shape of the convolution kernel to a Gaussian shape and perform a line search to determine the size of the

Gaussian kernel corresponding to the peak image quality score [31]. This method doesn't correct the illumination pattern of the retinal image, In addition, not all retinal images are degraded by a simple Gaussian kernel. Therefore, there is a need for methods that can solve two ill-posed problems in one scene, and achieve blind illumination correction and deconvolution simultaneously, which can potentially benefit clinical application on retinal images.

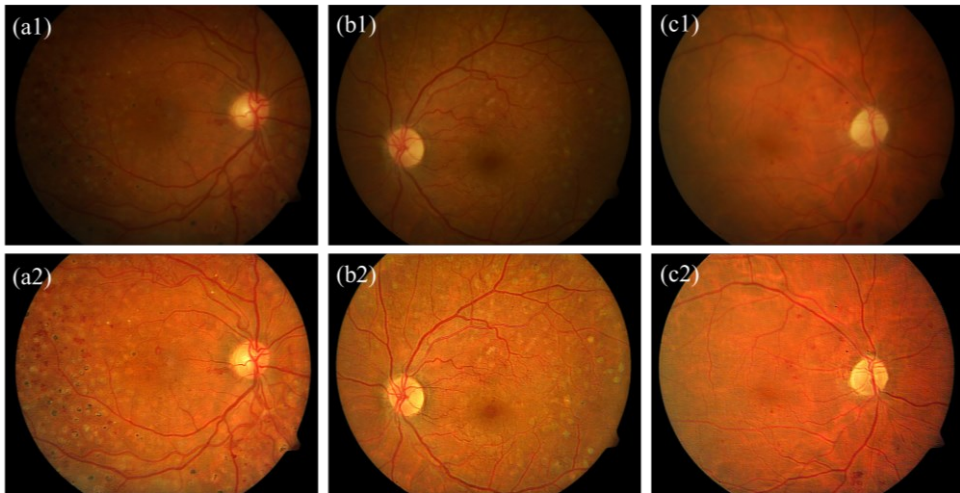


Fig. 7-1. Demonstration of LRRL on retinal images from DiaRet database. Images in the first row are raw images, and images in the second row are outputs of LRRL.

In this study, we proposed the luminosity rectified blind Richardson-Lucy deconvolution (LRRL) algorithm. Our research has the following four contributions: (1) we proposed a new cost function for the blind deconvolution problem that can correct the problematic illumination of the retinal image while achieving image blind deconvolution at the same time. (2) we proposed a differentiable approximation to efficiently solve the L_0 -norm problem, and the closed-form gradient is deduced. (3) we proposed an embedded illumination correction cost function with Hessian regularization for the non-blind deconvolution task to obtain the final output of our model. (4) our proposed model outperforms other state-of-the-art retinal image blind

deconvolution methods. Fig. 7-1 shows the results of our LRRL approach on retinal image deconvolutions.

The rest of this manuscript is arranged as follows: Section 2 introduces the double-pass fundus reflection model and its simplified version for single retinal image deconvolution. Section 3 proposes the blind Richardson-Lucy deconvolution framework with an extended cost function for the estimation of the latent image, the latent illumination pattern, and the convolution kernel. With knowledge of the convolution kernel and illumination pattern, we designed a non-blind luminosity rectified deconvolution framework under Richardson-Lucy deconvolution with Hessian regularization. Section 4 compares LRRL against state-of-the-art blind deconvolution methods using both visual and objective assessment. Section 5 demonstrates the potential clinical applications of our proposed method for computer-aided diagnosing. Section 6 discussed the contributions and limitations of the LRRL method, and concluding remarks make up the final section.

2. Double-pass fundus reflection model

Here we consider the double-pass fundus reflection process of retinal image formation and ignore the intraocular scattering problem caused by opacities in the optical media.

Fig. 7-2 shows a schematic sketch of the light path in a fundus camera, where $p_0(\mathbf{r})$ is the point spread function of the eye, $\mathbf{o}(\mathbf{r})$ the retinal reflectance (an M -by- N matrix) and $\mathbf{r} = (x, y)$ the coordinate vector.

i_0 is the illumination pattern generated by the light source at a certain distance and is imaged by the ocular system forming \mathbf{i} — the image of i_0 that illuminates the retinal surface. Therefore $\mathbf{i} = p_0 \otimes i_0$ as the fundus camera uses incoherent white light for illuminations. Then, the reflectance from the retina surface is imaged by the imaging system which is the combination of the ocular system and fundus camera, yielding the

final observed retinal images. The double-pass fundus reflection model is written as [36, 37]:

$$s(\mathbf{r}) = \mathbf{p}(\mathbf{r}) \otimes \{[\mathbf{p}_0(\mathbf{r}) \otimes \mathbf{i}_0(\mathbf{r})] \circ \mathbf{o}(\mathbf{r})\} + \varepsilon, \quad (7-1)$$

where $\mathbf{p}(\mathbf{r})$ is the point spread function of the combined imaging system of the fundus camera and eye, $s(\mathbf{r})$ is the observed retinal image, and ε is an unavoidable noise signal. \otimes is the 2D-convolution and \circ the pixel-wise matrix multiplication.

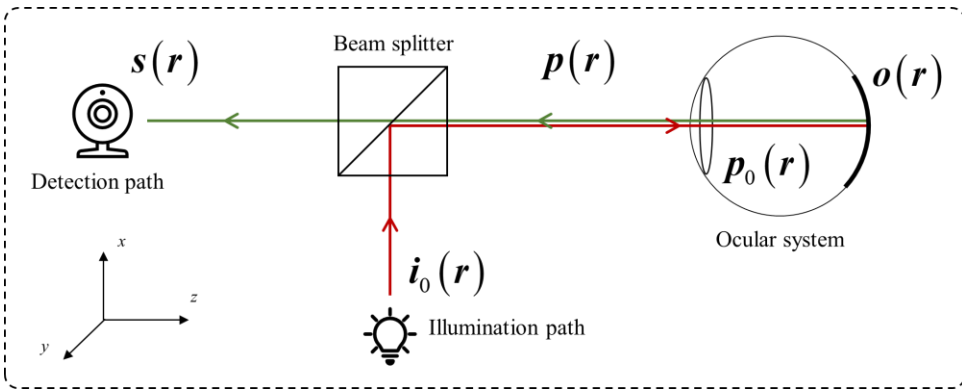


Fig. 7-2. Sketch for optical path in fundus camera. A beam splitter is used to deliver the illumination light into the eyes.

Since $\mathbf{p}_0(\mathbf{r}) \otimes \mathbf{i}_0(\mathbf{r})$ works as a new illumination component cast on the retina, we are able to further simplify Eq. (1) to:

$$s(\mathbf{r}) = \mathbf{p}(\mathbf{r}) \otimes [\mathbf{i}(\mathbf{r}) \circ \mathbf{o}(\mathbf{r})] + \varepsilon, \quad (7-2)$$

Our LRRL approach aims to simultaneously recover the estimations of $\mathbf{p}(\mathbf{r})$, $\mathbf{i}(\mathbf{r})$ and $\mathbf{o}(\mathbf{r})$ with knowledge only of $s(\mathbf{r})$.

Before entering the retinal image blind deconvolution, an image background padding process was applied in order to remove the boundary effect during the deconvolution process. Further, throughout the manuscript, pixel values were normalized within the range of [0, 1] (divided by 255 for uint8 format images).

3. Luminosity rectified blind deconvolution

Assuming that the convolution kernel for three color channels is identical, we can convert the RGB color image into the gray-scale image, and perform blind deconvolution. Our luminosity rectified blind deconvolution model is given by an optimization problem, where the cost function to be minimized is

$$E(\mathbf{o}, \mathbf{i}, \mathbf{p}) = \|\mathbf{p} \otimes (\mathbf{i} \circ \mathbf{o}) - \mathbf{s}\|_2^2 + \alpha \varphi(\mathbf{o}) + \beta \omega(\mathbf{i}) + \gamma \psi(\mathbf{p}). \quad (7-3)$$

The first term is a data fidelity term, which restricts the estimation of $(\mathbf{o}, \mathbf{i}, \mathbf{p})$ to be consistent with the observed image \mathbf{s} . This data fidelity term is consistent with the Richardson-Lucy deconvolution under the assumption of Gaussian noise.

Since the single image blind illumination correction and deconvolution task are severely ill-posed as there exist many different solution groups of $(\mathbf{o}, \mathbf{i}, \mathbf{p})$ that give rise to the same \mathbf{s} , the $\varphi(\mathbf{o})$, $\omega(\mathbf{i})$ and $\psi(\mathbf{p})$ are penalty functions that restrict the solution of $\mathbf{o}, \mathbf{i}, \mathbf{p}$, respectively, in order to make the problem well-posed. Here, α , β , and γ are penalty parameters.

$\varphi(\mathbf{o})$ is the l_0 -norm of the gradient of \mathbf{o} which retains large gradients and removes tiny details [35, 38], and $\psi(\mathbf{p})$ is the l_2 -norm of \mathbf{p} . We assume that \mathbf{i} is spatially slow varying and therefore we use the total-variation to ensure this property of \mathbf{i} . By vectorizing the \mathbf{o} and \mathbf{s} we obtain the following optimization problem

$$E(\vec{\mathbf{O}}, \mathbf{I}, \mathbf{K}) = \|\mathbf{K}\mathbf{I}\vec{\mathbf{O}} - \mathbf{S}\|_2^2 + \alpha \|\nabla \vec{\mathbf{O}}\|_0 + \beta \|\nabla \mathbf{I}\|_1 + \gamma \|\mathbf{K}\|_2, \quad (7-4)$$

where \mathbf{K} is an MN -by- MN matrix denoting the convolution operation. \mathbf{I} is an MN -by- MN diagonal matrix denoting the pixel-wise illumination operation. Eq. (7-4) is further converted into two sub-problems (1) with knowledge of \mathbf{K} , the latent image \mathbf{O} , and illumination pattern \mathbf{I} are solved. (2) With knowledge of $\mathbf{I}\mathbf{O}$, an estimation of \mathbf{K} is obtained. These two steps are iteratively implemented among the coarse-to-fine multi-scale of the image.

3.1 Estimation of latent image and illumination pattern

As Eq. (7-4) involves the L_0 -norm term which is an NP-hard problem, it can be solved using the Half Quadratic Splitting (HQS) methods [39] by introducing two auxiliary variables with respect to $\nabla \vec{\mathbf{O}}$ and $\nabla \mathbf{I} \cdot \vec{\mathbf{O}}$ or \mathbf{I} can be solved by setting the derivative of (4) with respect to $\vec{\mathbf{O}}$ or \mathbf{I} to zero. However, as we have shown in **Supplementary Note 7-1**, the half quadratic splitting method requires a large amount of computation and is also time-consuming due to the presence of \mathbf{I} .

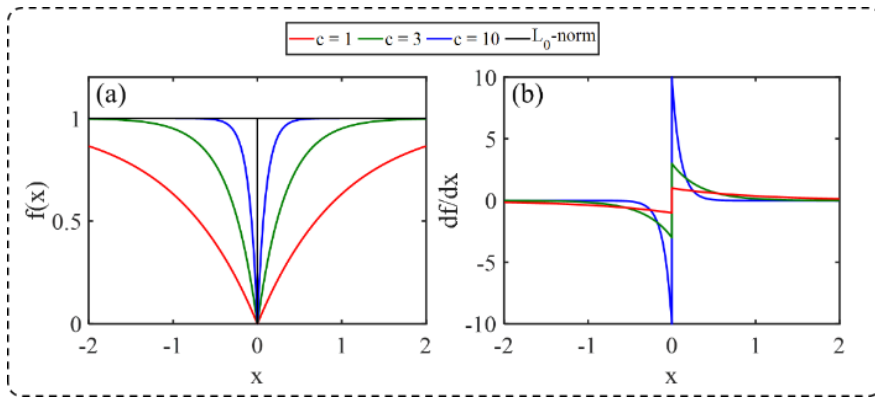


Fig. 7-3. Approximation of L_0 -norm using exponential functions. (a) plots of Eq. (5) with respect to parameters c and x . (b) derivative of Eq. (5) with respect to x .

To tackle the L_0 -norm in Eq. (7-4) and maintain computational efficiency, here we propose to approximate the L_0 -norm of a vector/matrix using an element-wise function:

$$f(x) = 1 - \exp(-c|x|), \quad f(x) \geq 0, \quad \forall x \in R. \quad (7-5)$$

As shown in Fig. 7-3 (a), Eq. (7-5) can approximate the L_0 -norm by assigning a large value to parameter c . We, therefore, are able to approximate the L_0 -norm of a vector/matrix \mathbf{X} by calculating the element-wise summation of the vector/matrix

$f(\mathbf{X})$. **Supplementary Note 7-2** shows the details of this approximation. The derivative of $f(\mathbf{X})$ with respect to x is shown in Fig. 7-3 (b).

The sub-problem for latent image and illumination pattern is then given by

$$E_1(\mathbf{o}, \mathbf{i}, \mathbf{p}) = \|\mathbf{p} \otimes (\mathbf{i} \circ \mathbf{o}) - \mathbf{s}\|_2^2 + \alpha \|1 - \exp(-c|\nabla \mathbf{o}|\|) + \beta \|\nabla \mathbf{i}\|_1, \quad (7-6)$$

we can use the L₁-norm to represent the element-wise summation since $f(\mathbf{X})$ is a non-negative function. Moreover, since Eq. (7-6) is differentiable everywhere except the origin, we are able to optimize it using gradient descent methods, and the (sub)gradient of E_1 with respect to \mathbf{o} is given by

$$\frac{\partial E_1}{\partial \mathbf{o}} = 2\mathbf{i} \circ [\mathbf{p}^* \otimes \mathbf{p} \otimes (\mathbf{i} \circ \mathbf{o}) - \mathbf{p}^* \otimes \mathbf{s}] + \alpha c \nabla^T \left[\exp(-c|\nabla \mathbf{o}|) \circ \frac{\nabla \mathbf{o}}{|\nabla \mathbf{o}|} \right], \quad (7-7)$$

while the (sub)gradient with respect to \mathbf{i} is given by

$$\frac{\partial E_1}{\partial \mathbf{i}} = 2\mathbf{o} \circ [\mathbf{p}^* \otimes \mathbf{p} \otimes (\mathbf{i} \circ \mathbf{o}) - \mathbf{p}^* \otimes \mathbf{s}] + \beta \nabla^T \left(\frac{\nabla \mathbf{i}}{|\nabla \mathbf{i}|} \right). \quad (7-8)$$

Here $\mathbf{p}^* = \mathbf{p}(-r)$. Without \mathbf{i} , the first term in Eq. (7-7) is consistent with the standard Richardson-Lucy deconvolution with Gaussian noise likelihood [40]. (also in **Supplementary Note 7-5**.)

Equation (7-6) is non-convex, and the gradient of the exponential term is very small when x is large according to Fig. 7-3. While when the x becomes small, the gradient of the exponential term increases rapidly. As such, the normal gradient descent method with static step-size may not ensure the convergence of Eq. (7-6), especially for the exponential term. It is hoped that the optimization can be accelerated when the gradient is flat, while can be also slowed down a bit when the gradient is extremely large. Here, in this research, we choose the Nesterov accelerated adaptive momentum estimation (N-Adam) [41] to boost the gradient descent method and avoid the gradient bursting

problem in Eq. (7-7). For detailed validation and comparison between different optimizers, please see **Supplementary Note 7-2**.

Algorithm 7-1: Estimation of latent image and illumination pattern

Input: kernel \mathbf{p} ; image \mathbf{s} ; V-channel \mathbf{v} ; parameters α and β .

$t = 1$

Calculate \mathbf{i}^0 using Eq. (7-9)

$\mathbf{o}^o = \mathbf{s} / \mathbf{i}^0$

While $t < t_{\max}$ **do**

 Calculate the gradient $\partial_o E_1$ using Eq. (7-7)

 Calculate the gradient $\partial_i E_1$ using Eq. (7-8)

 Updating \mathbf{o}^t and \mathbf{i}^t using N-Adam optimizer

$t \leftarrow t + 1$

End while

Output: Latent image \mathbf{o} and latent illumination pattern \mathbf{i} .

The initial guess of \mathbf{i} is given by

$$\mathbf{i}^0 = \left[\frac{1}{2\pi\sigma^2} \exp\left(-\frac{\mathbf{r}^2}{2\sigma^2}\right) \right] \otimes \mathbf{v}, \quad (7-9)$$

according to the Retinex theory in which the illumination pattern is obtained by low-pass filtering of the raw image [42]. $\sigma = \sqrt{MN/10}$ controls the size of the Gaussian kernel, and \mathbf{v} is the V-channel of the raw image by converting the RGB image of corresponding scales into HSV color space. While the initial guess of \mathbf{o} is given by \mathbf{s}/\mathbf{i}^0 . This algorithm is summarized in **Algorithm 7-1**, which contains only one loop and is more efficient than HQS methods. t_{\max} is the maximum iteration number for solving \mathbf{o} and \mathbf{i} , and $t_{\max} = 25$ in our experiments.

3.2 Estimation of the convolution kernel

With the combination of estimated \mathbf{o} and \mathbf{i} , we enter our next step: the kernel estimation. Similar to other existing state-of-the-art methods, the kernel is estimated based on the gradient of \mathbf{s} and $\mathbf{i} \circ \mathbf{o}$ which is given by the following optimization problem [34, 35, 38, 43, 44]

$$E_2(\mathbf{p}) = \|\mathbf{p} \otimes [\nabla(\mathbf{i}\mathbf{o})] - \nabla\mathbf{s}\|_2^2 + \gamma \|\mathbf{p}\|_2^2. \quad (7-10)$$

The image gradient in Eq. (7-10) benefits the estimation of the kernel by utilizing the edge information. Eq. (7-10) is quadratic, and the solution of \mathbf{p} can be directly calculated by setting the derivative of Eq. (7-10) with respect to \mathbf{p} to 0. The convolution in Eq. (7-10) can be accelerated by the Fourier transform, and \mathbf{p} is then given by

$$\mathbf{p} = \text{Re} \left\{ \mathcal{F}^{-1} \left\{ \frac{\text{conj}[\mathcal{F}(\nabla_x \mathbf{i}\mathbf{o})] \mathcal{F}(\nabla_x \mathbf{s}) + \text{conj}[\mathcal{F}(\nabla_y \mathbf{i}\mathbf{o})] \mathcal{F}(\nabla_y \mathbf{s})}{|\mathcal{F}(\nabla_x \mathbf{i}\mathbf{o})|^2 + |\mathcal{F}(\nabla_y \mathbf{i}\mathbf{o})|^2 + \gamma} \right\} \right\}. \quad (7-11)$$

Here, \mathcal{F} denotes the 2D Fourier transform, and $\text{conj}(x)$ denotes the complex conjugate of variable x . $\text{Re}(x)$ denotes the real path of the complex variable x . Furthermore, the estimated kernel is refined by setting the negative elements to zero and normalization so that $\sum_r \mathbf{p}(r) = 1$.

Fig. 7-4 shows the latent image and illumination estimation for two groups of experiments of retinal image blind deconvolutions. As shown in Fig. 7-4 (a1) and Fig. 7-4 (a2) both raw images have uneven and insufficient problematic illumination conditions, while our LRRL deconvolution generates the illumination-corrected latent images as shown in Fig. 7-4 (b1) and Fig. 7-4 (b2), together with the illumination patterns as shown in Fig. 7-4 (c1) and Fig. 7-4 (c2). Accordingly, the estimation of illumination patterns corresponds to the raw images as regions near the optical disk have a higher intensity than other places.

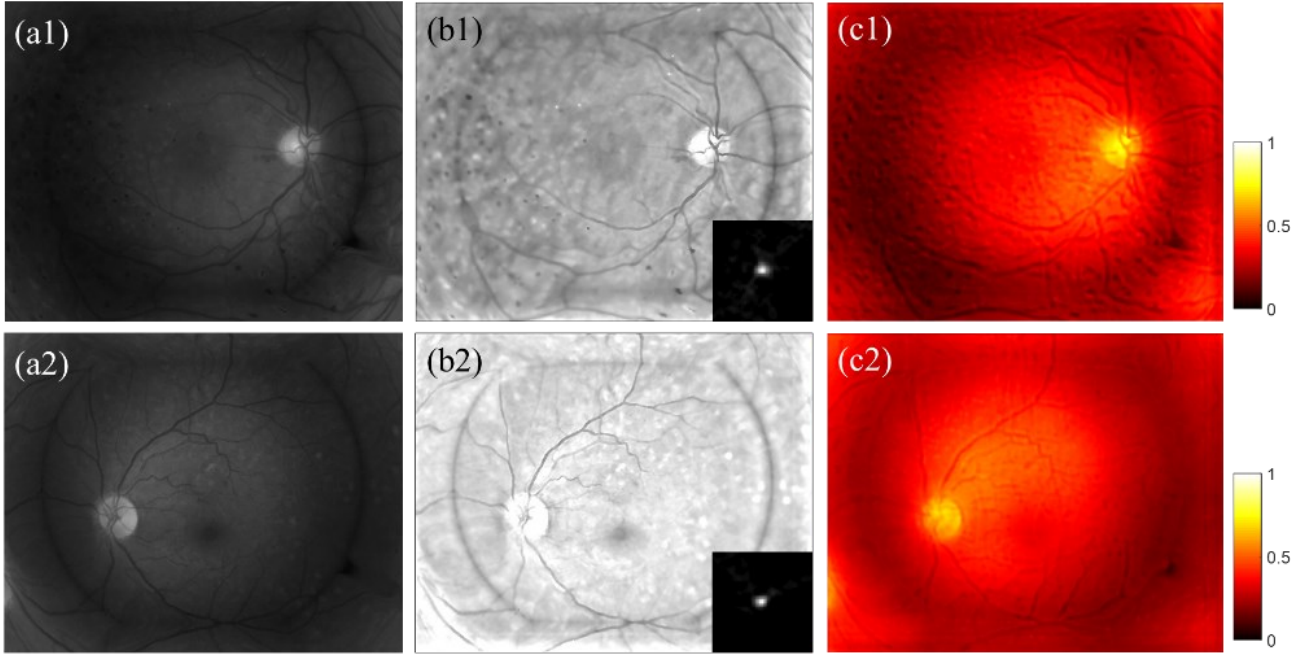


Fig. 7-4. Latent images and illumination patterns for blind illumination correction and deconvolution. (a1) and (a2) are raw images in grayscale corresponding to the images in Fig. 7-1 (a1) and (b1), respectively. The convolution kernels are estimated in 51-by-51 pixels and are up-sampled to 128-by-128 in Fig. 7-4 for better observation. (b1) and (b2) are latent images. (c1) and (c2) are latent illumination estimations.

According to Fig. 7-4 (b1) and Fig. 7-4 (b2), the latent images are illumination corrected, however, their intensity levels are different as Fig. 7-4 (b2) is brighter than that of Fig. 7-4 (b1), therefore an intensity refinement needs to be applied to the final non-blind deconvolution process as we will show in the next subsection.

In general, the two-step blind deconvolution is summarized in **Algorithm 7-2**.

Algorithm 7-2: Luminosity rectified blind Richardson-Lucy deconvolution

Input: Raw retinal image \mathbf{s} , an initial guess of kernel \mathbf{p} from the estimation in the last coarser-scale

While $\kappa < \text{max_iter}$ **do**

Updating \mathbf{i} and \mathbf{o} using **Algorithm 7-1**

Updating \mathbf{p} using Eq. (7-11)

$\kappa \leftarrow \kappa + 1$

End While

Output: Latent image \mathbf{o} , illumination pattern \mathbf{i} , convolution kernel \mathbf{p}

Notes: \mathbf{s} , \mathbf{i} and \mathbf{p} are then the inputs of subsequent Richardson-Lucy Hessian deconvolution.

3.3 Non-blind deconvolution: RL-Hessian cost

The latent image is not the final output of LRRL since it is over-smoothed. With the estimation of illumination pattern \mathbf{i} and convolution kernel, a final non-blind deconvolution process is applied to the raw RGB retinal image. We design the new cost function which is

$$E_3(\mathbf{o}, \eta) = \|\mathbf{p} \otimes (\mathbf{I} \circ \mathbf{o}) - \mathbf{s}\|_2^2 + \zeta \|\mathcal{H} \mathbf{o}\|_1 + \left\| \frac{\mathbf{s}}{\mathbf{I}(\eta)} - \bar{\mathbf{I}} \right\|_2^2, \quad (7-12)$$

to simultaneously achieve illumination correction, deconvolution, and noise suppression, where

$$\mathbf{I}(\eta) = 1 - \eta \times (1 - \mathbf{i}). \quad (7-13)$$

$E_3(\mathbf{o})$ in Eq. (7-12) is our proposed Richardson-Lucy-Hessian cost, and the data fidelity term is derived from Richardson-Lucy deconvolution under the assumption of Gaussian noise distribution.

The second penalty function in Eq. (7-12) suppresses the noise signal during the deconvolution process, \mathcal{H} denotes the Hessian gradient of image \mathbf{o} , viz, the second-order gradient. We are not using the TV regularization since it uses the first-order gradient and will cause staircase-like artifacts. Retinal images differ from natural scene images as they do not contain sharp edges, while natural scenes do. Therefore, the staircase-like artifacts will largely decrease the naturalness of biomedical images [45] including retinal images.

The third term restricts the average intensity distribution of \mathbf{o} approaching the target intensity \bar{I} . If the average intensity of the raw image is larger than \bar{I} , no illumination correction is applied. We are not using the original illumination pattern \mathbf{i} as it may cause color distortion for RGB images. Instead, we use Eq. (7-13) as a replacement for the illumination pattern and η is a scalar parameter that controls the illumination correction degree since not all retinal images need to be corrected.

Again, we use gradient descent with N-Adam to solve Eq. (7-12), and the derivative (sub-gradient) of E_3 with respect to \mathbf{o} is

$$\frac{\partial E_3}{\partial \mathbf{o}} = 2\mathbf{l} \circ \left[\mathbf{p}^* \otimes \mathbf{p} \otimes (\mathbf{l} \circ \mathbf{o}) - \mathbf{p}^* \otimes \mathbf{s} \right] + \zeta \cdot \mathcal{H}^T \frac{\mathcal{H} \mathbf{o}}{|\mathcal{H} \mathbf{o}|}, \quad (7-14)$$

and the derivative of E_3 with respect to η is

$$\frac{\partial E_3}{\partial \eta} = \sum_r \left[(1 - \mathbf{i}) \circ \left(\frac{\mathbf{s}}{\mathbf{l}} - \bar{I} \right) \circ \frac{\mathbf{s}}{\mathbf{l}^2} \right]. \quad (7-15)$$

Algorithm 7-3: Richardson-Lucy Hessian deconvolution

Input: Raw retinal image s , estimation of illumination pattern i , convolution kernel p from **Algorithm 7-2**

While $\kappa < max_iter$ **do**

 Updating o using Eq. (7-14)

 Updating η using Eq. (7-15)

$\kappa \leftarrow \kappa + 1$

End While

Output: Restoration image o

The non-blind deconvolution procedure is summarized in **Algorithm 7-3**. In summary, a flow chart for a complete LRRL method has two main steps, which are summarized in Fig. 7-5. First, with the input image, we use the blind-deconvolution to obtain the estimation of illumination pattern and blurry kernel (Section 3.1 and 3.2). Second, with the illumination pattern and kernel, we apply non-blind deconvolution (Section 3.3) to the input image to obtain the luminosity rectified, deconvoluted images.

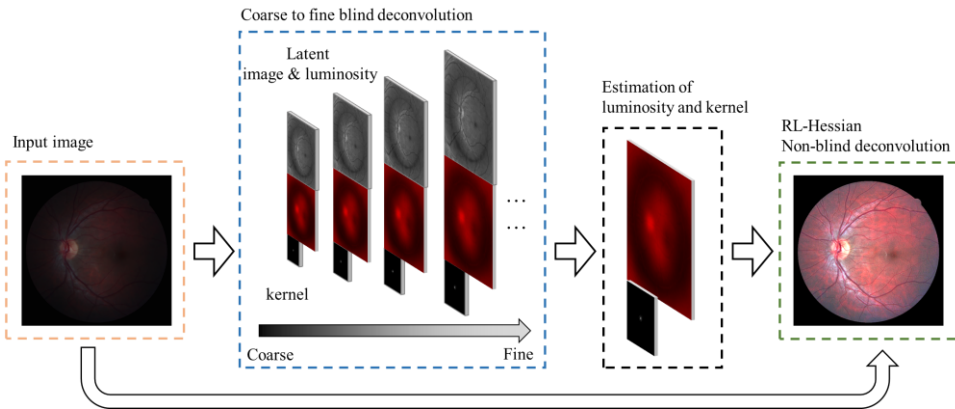


Fig. 7-5. Flow chart for the complete LRRL method including both blind deconvolution and non-blind deconvolution.

3.4 Model parameters

In our proposed model, we have a total of 6 parameters that need to be adjusted, which are α , β , γ , ζ , \bar{I} , and kernel size. Parameters α , β , and γ are for the blind deconvolution and $\gamma = 1$ according to historical literature. In the following experiments, we set $\beta = 2$ to ensure the illumination pattern is spatially smoothly varying, and $\alpha = 0.0025$ according to experimental trials. The kernel size should be manually adjusted like other blind deconvolution methods. In this study, the default kernel size is 51 by 51.

For the non-blind deconvolution part, $\bar{I} = 0.55$ to ensure maintain the naturalness of the output retinal images and the final parameter ζ is chosen according to the noise level and we use a noise-estimator [46]

$$\zeta = \frac{1}{50} \frac{1}{C} \frac{1}{MN} \sum_{n=1}^C \sum_{x=1}^M \sum_{y=1}^N |s_n(x, y) \otimes \mathcal{L}|, \quad (7-16)$$

to adaptively adjust the value of ζ . Where C is the number of color channels, M and N are the size of the image in x and y directions. $\mathcal{L} = [-1, 2, -1; 2, -4, 2; -1, 2, -1]$ is the Laplacian operator as it is sensitive to noise signals. For noise-free images, the Laplacian operator extracts the edge information. On the contrary, for noise-corrupted images, the Laplacian operator enlarges the noise signals. Therefore, the average absolute pixel value of the filtered image can be regarded as a measurement of the noise level.

Fig. 7-6 shows the final output of LRRL for two retinal images from the DiaRet database. The LRRL corrects the uneven illumination and improves the clarity of the images as shown in Fig. 7-6 (b1) to Fig. 7-6 (c2). It is worth noting that in many retinal image enhancement methods, illumination correction may cause intensity saturation for regions near the optical disk which usually has a large intensity distribution. Thanks to our proposed average intensity strategy in Eq. (7-12) and Eq. (7-13), the LRRL will overcome this problem and will maintain the naturalness of the retinal images.

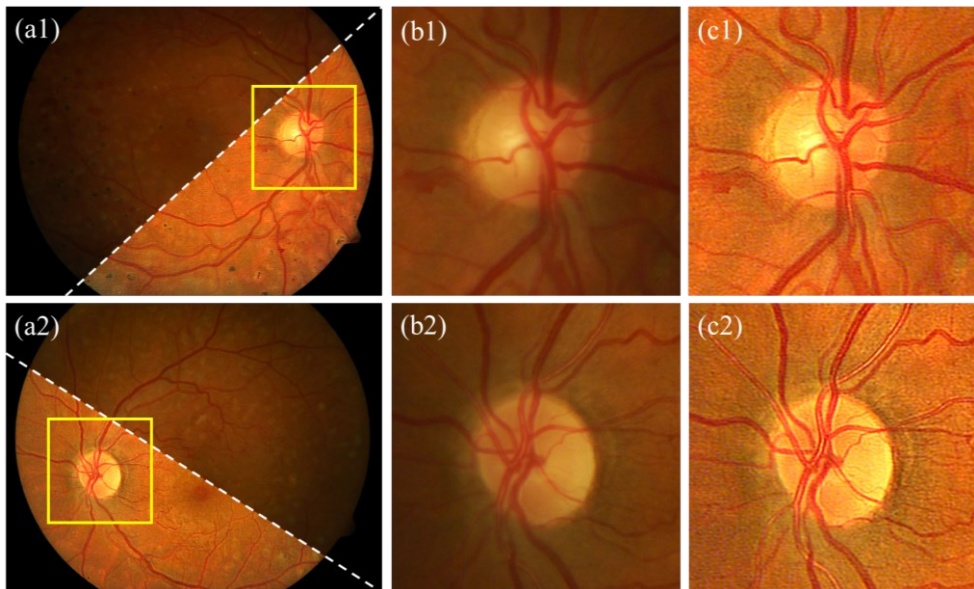


Fig. 7-6. Montage of raw and output retinal images for the image in Fig. 7-1 (a1) and Fig. 7-1 (a2). (b1) and (c1) are enlarged parts in the yellow box in (a1). (b2) and (c2) are enlarged parts in the yellow box in (a2).

4. Experimental results

We test the LRRL on DiaRet retinal image database (219 images included) [47, 48], NORMAL database (300 images included) [49] and Messidor database (1200 images included) [50]. A total of 1719 images are processed. We compare the results of LRRL against two state-of-the-art blind deconvolution methods, which we refer the method in [44] as PMP, and the method in [31] as SLC. Method in Ref. [30] is not included since it requires paired two retinal images where in our case we only have a single measurement. The kernel size is 51 by 51.

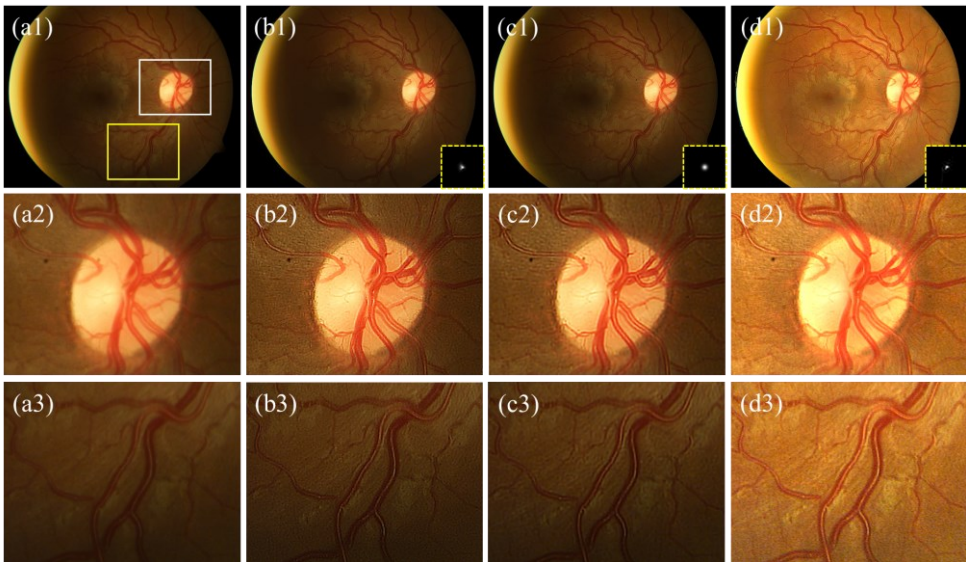


Fig. 7-7. Experimental results on image047 from DiaRetDB01 dataset. (a1) to (d1) are raw image, and output of PMP, SLC, and LRRL. (a2) to (d2) are zoomed-in parts in the region of the white box. (a3) to (d3) are zoomed-in parts in the region of the yellow box.

4.1 Qualitative evaluations

Fig. 7-7 shows the experimental results on image047 from DiaRetDB01 dataset. Where the raw image has uneven and insufficient problematic illumination conditions as shown in Fig. 7-7 (a1). The raw image is also blurred as shown in Fig. 7-7 (a2). All three methods correct the blurriness as shown in Fig. 7-7 (b2) to (d3), where the clarity of the retinal structures including blood vessels and tissue textures are improved in comparison with the raw image shown in Fig. 7-7 (a2) and (a3).

SLC estimates the Gaussian kernel as shown in Fig. 7-7 (c1). However, since the kernel cannot be simply approximated by the Gaussian distribution, SLC slightly overestimates the size of the Gaussian kernel yielding the ringing effect, especially at the edge of the blood vessels as shown in Fig. 7-7 (c2).

The LRRL can further correct the illumination pattern as shown in Fig. 7-7 (d1), and the visual quality of Fig. 7-7 (d1) is better than that of Fig. 7-7 (b1) due to the

illumination correction, while the blurriness correction abilities of all three methods are similar according to the visual assessment.

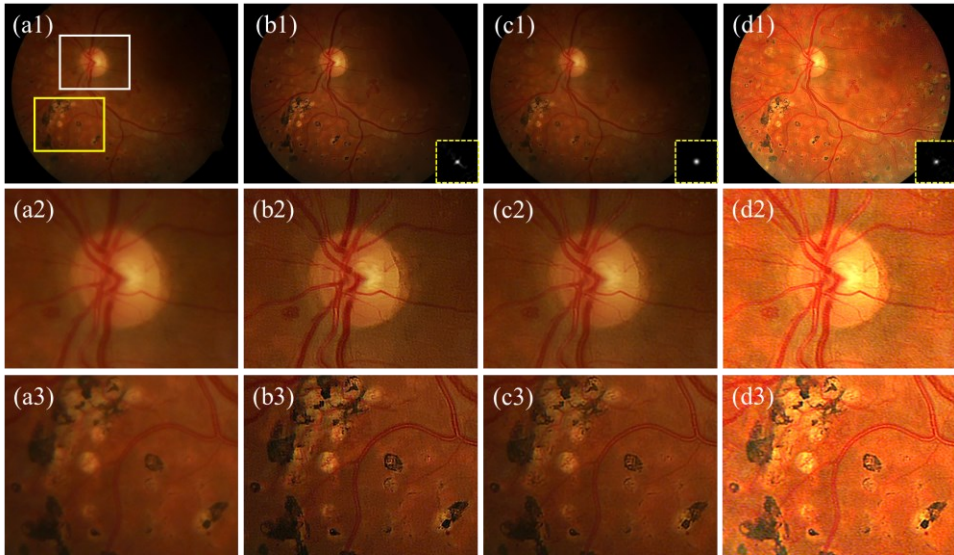


Fig. 7-8. Experimental results on image088 from DiaRetDB01 dataset. (a1) to (d1) are raw image, output of PMP, SLC, and LRRLL. (a2) to (d2) are zoomed-in parts in the region of the white box. (a3) to (d3) are zoomed-in parts in the region of the yellow box.

Fig. 7-8 and Fig. 7-9 show other two groups of experimental results on image037 from the DiaRetDB01 dataset, and image 2 from the MESSIDOR dataset, respectively. PMP shows its ability on blind deconvolution as the clarity of the output images becomes better than the raw images as shown in Fig. 7-8 (b1) and Fig. 7-9 (b1), however, the visual quality of the raw images is still limited by the poor illumination condition.

Although SLC also achieves deconvolution as the clarity of the resultant images in Fig. 7-8 (c1) to Fig. 7-8 (c3) are improved, it fails to estimate the small structure of the kernel since it limits the shape of the kernel to be Gaussian, while in this case, the kernel has side lobes as shown in Fig. 7-8 (b2) and Fig. 7-8 (d2). Moreover, according to Fig.

7-9 (c1), SLC overestimates the size of the kernel causing ringing effects along the tube structures like blood vessels.

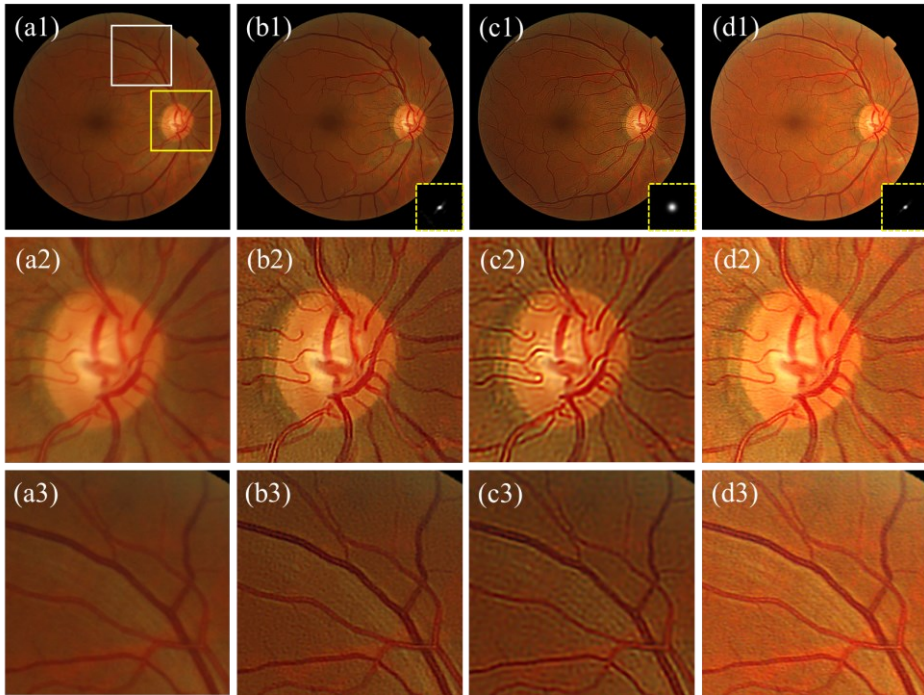


Fig. 7-9. Experimental results on image 2 from MESSIDOR dataset. (a1) to (d1) are raw image, output of PMP, SLC, and LRRL. (a2) to (d2) are zoomed-in parts in the region of the white box. (a3) to (d3) are zoomed-in parts in the region of the yellow box.

On the contrary, the LRRL method corrects the illumination pattern output images, as shown in Fig. 7-8 (d1) and Fig. 7-9 (d1) which have good intensity distribution. The visual quality is therefore improved as the visual perception is more sensitive to intensity than to structures. The zoomed-in images are shown in group Fig. 7-8 (d2) and Fig. 7-8 (d3), and in group Fig. 7-9 (d2) and Fig. 7-9 (d3). The LRRL method also improves image clarity by blind-deconvolution.

In general, the LRRL achieves state-of-the-art deconvolution results for retinal images while also correcting poor illumination which largely increases the visual

quality of the restored retinal images. More experimental results for visual assessment are found in **Supplementary Note 7-4**.

4.2 Quantitative evaluations

In this sub-section, we conducted quantitative evaluations of the enhancement results in terms of (1) image definition [51], (2) image sharpness [52], (3) image entropy, and (4) image multiscale contrast [53]. All images are formatted in 8-bit unsigned integers (uint8, the gray-value is ranged in [0, 255]). **Supplementary Note 7-3** includes the details on the calculation of these four quality matrixes.

Tab. 7-1 lists the quantitative evaluations for the image restoration quality. Here, high values of DE, UISM, IE, and C_{RAMM} indicate high image quality. The raw images have low DE and UISM scores due to the blurriness of the image – convolution of a clear image with the blurry kernel reduces the edge information of the images, leading to the decrease in images' sharpness. Since the DE is calculated from the first-order gradient of the image, it is also reduced by the blurriness of the image.

As listed in Tab. 7-1, both PMP and LRRL show their ability of success in deconvolution as DE and UISM scores for resultant images are increased by at least two-fold. PMP gains a better UISM score than that of LRRL. The C_{RAMM} of LRRL method is better than those of PMP and SLC, which is mainly due to the intensity correction. Especially for Fig. 7-7, the $C_{RAMM} = 0.6061$ for the raw image, while the LRRL enhancement gains $C_{RAMM} = 2.0022$, which is three times larger than that of the raw images and results of SLC.

The SLC method may fail to achieve good deconvolution results since the DE for Fig. 7-7 and Fig. 7-8 decrease after the deconvolution. Additional experimental results for both qualitative and quantitative assessments are available in **Supplementary Note 7-4**.

Tab. 7-1. Quantitative assessment of the enhanced images among Fig. 7-7 to Fig. 7-9 with four metrics.

Raw image	Matrixes	Raw	Methods		
			PMP	SLC	LRRL
Fig. 7-7 (a1)	DE	1.4871	3.4013	1.6955	5.5548
	UISM	1.4613	4.9876	2.6535	4.3643
	IE	6.3506	6.4752	6.4126	6.7060
	C _{RAMM}	1.6499	2.0181	1.8385	2.9068
Fig. 7-8 (a1)	DE	1.1784	2.3878	0.7949	5.8679
	UISM	2.5338	4.7432	2.6968	5.0562
	IE	5.7798	5.8742	5.8065	6.5662
	C _{RAMM}	0.6061	0.8611	0.6291	2.0022
Fig. 7-9 (a1)	DE	2.2186	4.0115	2.1513	6.1596
	UISM	1.9077	5.3364	4.3909	5.2341
	IE	6.0099	6.2245	6.1172	6.3419
	C _{RAMM}	1.8655	2.1924	2.0474	2.5251
Average	DE	1.6280	3.2669	1.5472	5.8609
	UISM	1.9676	5.0224	3.2471	4.8849
	IE	6.0468	6.1913	6.1121	6.5380
	C _{RAMM}	1.3738	1.6905	1.5050	2.4780

Tab. 7-2 demonstrates the enhancement results of the quantitative evaluation for a total of 1719 images from three databases. In general, the DPFR model can achieve a better quality of image restoration including image definition, image entropy, and image contrast than current state-of-art methods. The blurriness and problematic illumination are corrected.

Note that we are not using any contrast enhancement processes such as HE or CLAHE. The improvement of image contrast is directly related to the effect of deconvolution and illumination correction. The image contrast of the LRRL-restored images can be further improved by using subsequent contrast enhancement algorithms like CLAHE, and others.

Tab. 7-2. Quantitative assessment of the enhanced images for 3 databases.

Datasets	Matrixes	Raw	Methods		
			PMP	SLC	LRRL
DiaRet db00 (130 images)	DE	1.3650± 0.1474	2.5955± 0.3577	1.2925± 0.3388	4.5285± 0.9611
	UISM	2.3573± 0.2036	4.4178± 0.3807	2.6097± 0.2301	4.0793± 0.4590
	IE	6.0696± 0.3549	6.2002± 0.3294	6.1135± 0.3564	6.3766± 0.2444
	C _{RAMM}	0.9829± 0.1805	1.3394± 0.2405	1.1162± 0.2711	1.9423± 0.3148
DiaRet db01 (89 images)	DE	1.3188± 0.1356	2.5146± 0.3643	1.1637± 0.2307	4.6585± 1.1380
	UISM	2.4132± 0.1980	4.3910± 0.4989	2.6688± 0.2039	4.2663± 0.5751
	IE	5.9016± 0.3546	5.9968± 0.3244	5.9400± 0.3467	6.3437± 0.2527
	C _{RAMM}	0.8910± 0.1916	1.1232± 0.1982	0.9889± 0.2212	1.9254± 0.3168
NORMAL (300 images)	DE	1.2650± 0.1811	3.2126± 0.5307	2.0208± 0.4667	3.3397± 0.9278
	UISM	1.5047± 0.2625	4.5128± 0.4582	3.4600± 0.4224	3.9061± 0.5383
	IE	6.1324± 0.2962	6.2475± 0.3068	6.2548± 0.2766	6.2010± 0.3186
	C _{RAMM}	1.7388± 0.1994	2.0199± 0.2413	2.0241± 0.2688	2.2926± 0.3495
Messidor (1200 images)	DE	1.4641± 0.2676	3.6115± 0.7298	2.1037± 0.5575	4.4378± 1.0826
	UISM	1.8228± 0.3220	5.2395± 0.5439	4.0469± 0.4322	4.6991± 0.5196
	IE	5.7119± 0.2882	6.0221± 0.2896	5.9131± 0.3068	6.1544± 0.2630
	C _{RAMM}	1.6444± 0.2455	2.0343± 0.3258	1.9001± 0.3114	2.1564± 0.3004
Average	DE	1.4144± 0.2540	3.4083± 0.7510	1.9792± 0.5894	4.2645± 1.1345
	UISM	1.8383± 0.3828	5.0066± 0.6268	3.7644± 0.6344	4.4914± 0.6141
	IE	5.8221± 0.3455	6.0736± 0.3116	5.9893± 0.3352	6.1892± 0.2801
	C _{RAMM}	1.5718± 0.3348	1.9320± 0.3996	1.8153± 0.4163	2.1520± 0.3260

4.3 Comparing with non-deconvolution methods

The LRRL aims to correct the blurriness and problematic intensity imposed on the retinal images. It won't increase the image contrast as much as that of contrast enhancement methods such as CLAHE, and other retinal image enhancement methods. To show this, we compare the output of LRRL against Luminosity and Contrast adjustment (LCA) [54], and DPFR methods as shown in Fig. 7-10, Fig. 7-11, and Fig. 7-12.

Tab. 7-3 Quantitative assessment of the enhanced images for Fig. s9 and s10.

Raw image	Matrixes	Raw	Methods			
			LCA	DPFR	LRRL	LRRL+CLAHE
Fig. 7-10	DE	1.4171	4.1591	9.7376	3.9061	9.0067
	UISM	1.5993	2.5288	6.5334	5.0967	6.3662
	IE	5.8493	7.0311	7.4273	6.5383	7.2597
	C _{RAMM}	1.7189	3.8928	7.3846	2.8848	5.1714
Fig. 7-11	DE	1.8946	4.1273	6.8521	5.7311	12.0209
	UISM	3.5776	3.5587	5.5637	5.3548	7.6116
	IE	6.1559	6.1436	6.3648	6.6497	6.5854
	C _{RAMM}	1.8915	2.8682	5.8095	2.1537	3.5687
Fig. 7-12	DE	1.5401	4.6584	9.4748	5.5939	12.7236
	UISM	3.7516	5.3423	5.5333	4.9767	6.3695
	IE	5.4914	6.6040	7.2645	6.2261	7.0355
	C _{RAMM}	1.5086	3.6568	8.0187	2.6189	5.5024

Tab. 7-3 shows the quantitative evaluation for the results in Fig. 7-10 to Fig. 7-12. As shown in Fig. 7-10 (b1) to Fig. 7-10 (d3), and according to Tab. 7-3, both LCA and DPFR methods increase the image's contrast as the color in the blood vessel region is deeper than in other background regions, the DE, UISM, IE, and C_{RAMM} scores are also largely improved. But the enhanced images still have blurriness problems, for example, the small blood vessel in the optical disk in Fig. 7-10 (b2) and Fig. 7-10 (c2) are not as clear as that in Fig. 7-10 (d2). Similar results can be found in Fig. 7-11 and Fig. 7-12.

While for the LRRL method as shown in Figs. 7- 10, 11, 12 (d1), the blurriness is corrected compared with the raw image, but the improvement in the image's contrast is not as much as that of LCA and DPFR outputs as listed in Tab. 7-3.

The main reason is that the LRRL models the image degrading process differently from the CLAHE or DPFR. In LRRL, the image degrading is described by a convolution process, while the contrast enhancement method such as CLAHE models the image degrading as changes in the statistical distribution of pixel values. For other model-based enhancement methods including the DPFR, the image degrading is modeled by a hazing formation process (matrix element-wise production). Note that the DPFR model distorts the color of retinal images.

Since the physical and mathematical insights are different, we would consider the combination of LRRL and contrast enhancements to further improve the image's contrast. Fig. 7-10, 7-11, and 7-12 (e1) show the LRRL output with the enhancement of CLAHE in L^* channel in the LAB color space. Comparing Fig. 7-10 (e2) and (e3) to Fig. 7-10 (b2) and (e2), for example, the CLAHE further improves the image's contrast.

Furthermore, with the combination of deconvolution and contrast enhancement, the final output can have better visual quality as listed in Tab. 3, scores for all four evaluation matrices are improved after the CLAHE is applied. In addition, as shown in Fig. 7-9(e2) [also in Fig. 7-10 (e2), and Fig. 7-12 (e2)], the blood vessels in the optical disk are more clear than that in Fig. 7-9 (a2), (b2), (c2), and (d2).

In summary, the LRRL aims to correct the blurriness caused by the convolution process. Its unique deblurring ability together with further contrast enhancement can achieve promising image restoration results for retinal images that have blurriness such as Fig. 7-10 (a1) to Fig. 7-12 (a1).

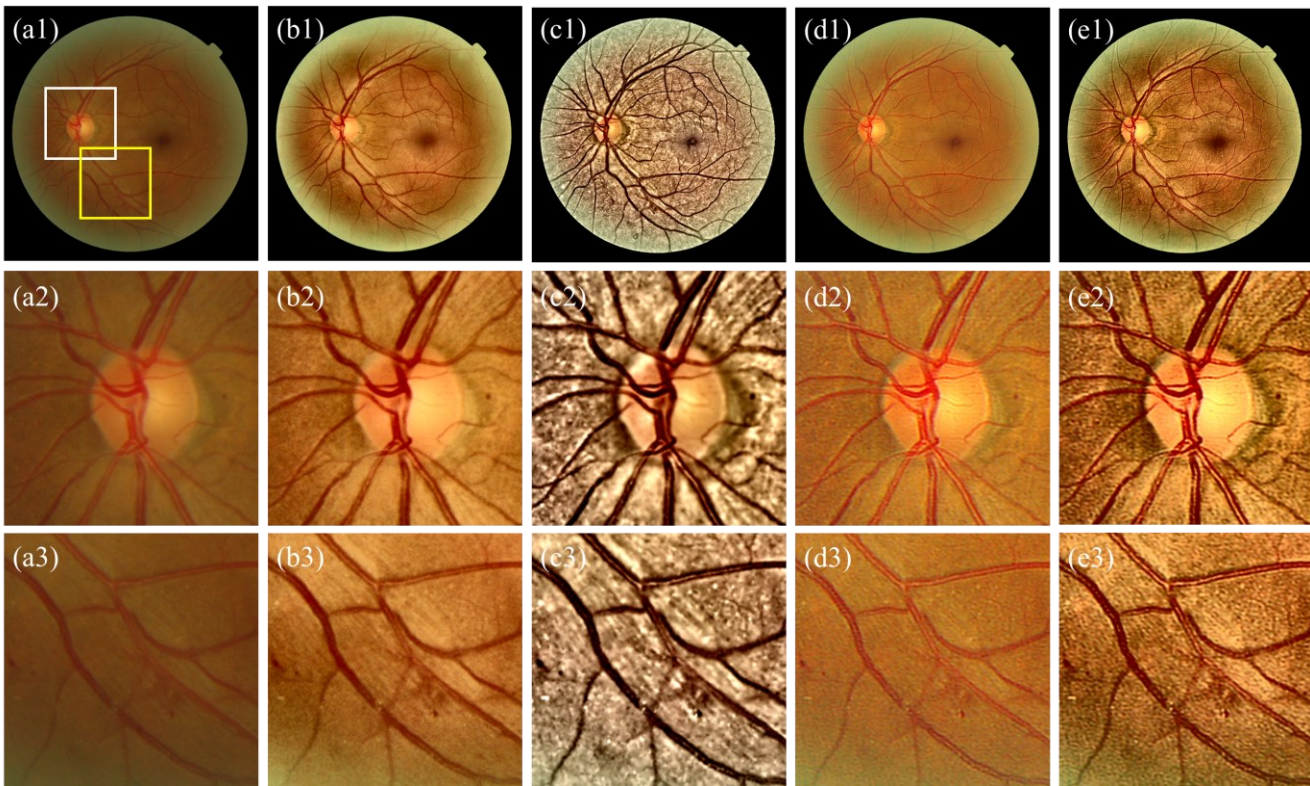


Fig. 7-10. Comparison with non-deconvolution methods. (a1) to (a3) are raw, and zoomed-in images. (b1) to (b3): LCA enhancement. (c1) to (c3), DPFR enhancements. (d1) to (d3) LRR output. (e1) to (e3), CLAHE enhancements based on (d1) to (d3).

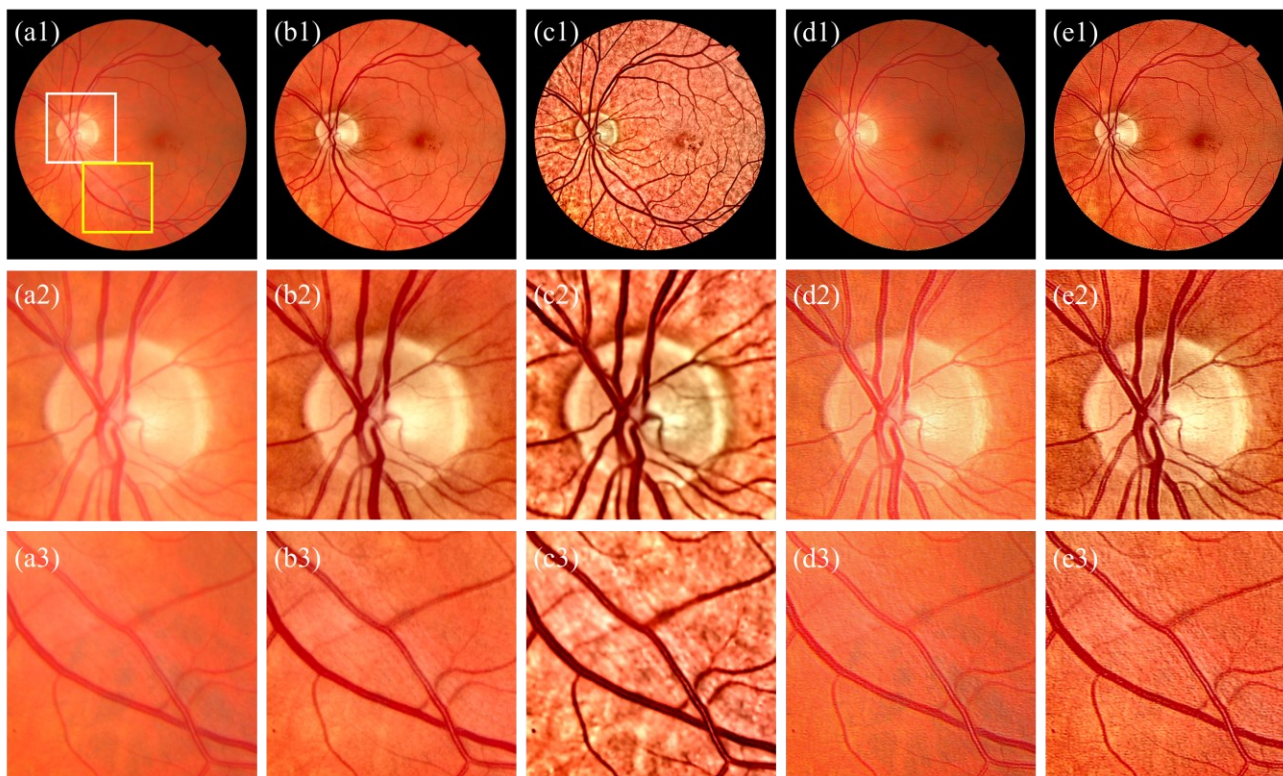


Fig. 7-11. Comparison with non-deconvolution methods. (a1) to (a3) are raw, and zoomed-in images. (b1) to (b3): LCA enhancement. (c1) to (c3), DPF enhancements. (d1) to (d3) LRR output. (e1) to (e3), CLAHE enhancements based on (d1) to (d3).

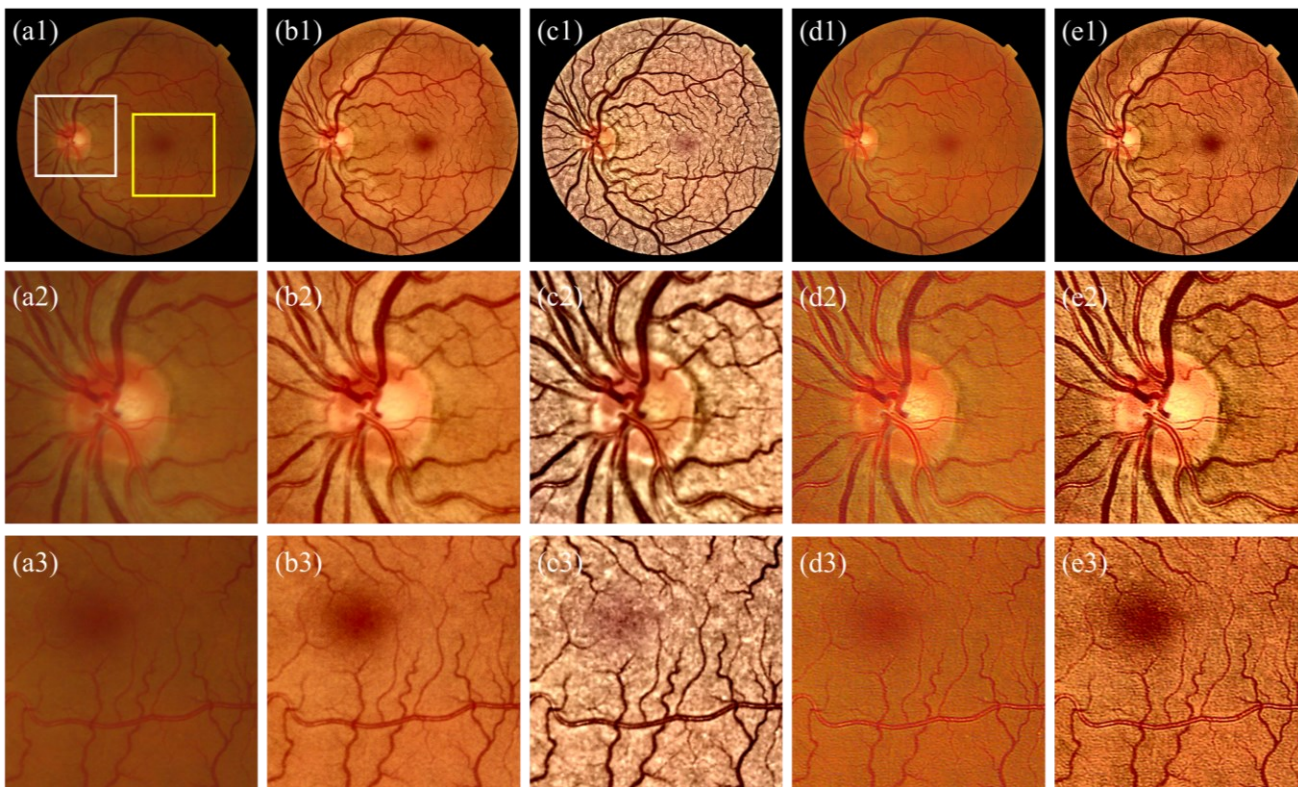


Fig. 7-12. Comparison with non-deconvolution methods. (a1) to (a3) are raw, and zoomed-in images. (b1) to (b3): LCA enhancement. (c1) to (c3), DPFR enhancements. (d1) to (d3) LRR output. (e1) to (e3), CLAHE enhancements based on (d1) to (d3).

5. Retinopathy diagnosis using LRRL restored images

As LRRL restores image clarity and corrects for uneven illuminations, the restored images can benefit clinical applications like retinopathy diagnosis. In this section, we focus on the restoration results in areas that indicate retinopathy, including hard exudates and hemorrhages. The restoration results should increase the visual quality of the retinopathy area, without creating unexpected artifacts to guarantee structure fidelity.

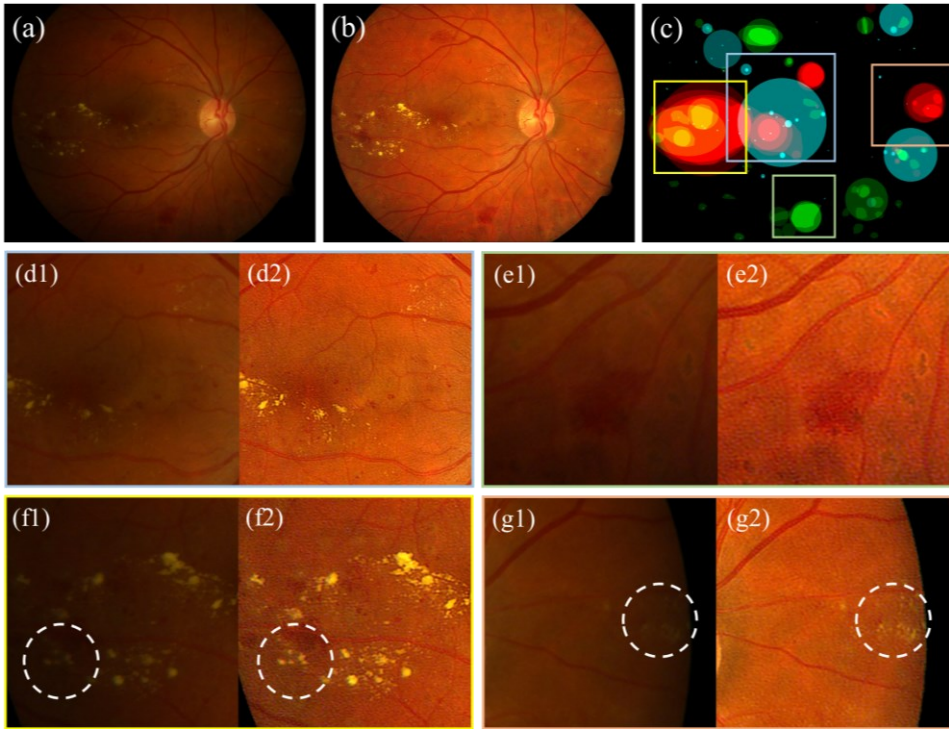


Fig. 7-13. Enhancement of retinopathy areas using LRRL. (a) Raw image. (b) Restored image. (c) Labels of retinopathy areas. Red: Hard exudates; Green: Hemorrhages; Cyan: red small dots. (d1) to (g2) are zoomed-in pictures for regions in blue, green, yellow, and red boxes.

We tested LRRL on the DiaRetDB1 dataset, since the retinopathies were already marked by human specialists [shown in Fig. 7-13 (c)]. The raw image shown in Fig. 7-13 (a) has problematic illumination conditions which improved after LRRL as shown in Fig. 7-13 (b).

As shown in Fig. 7-13 (d1), (d2), (f1), and (f2), LRRL increases the visual quality of the hard exudates, as the intensity of the corresponding area is better than that in the raw images.

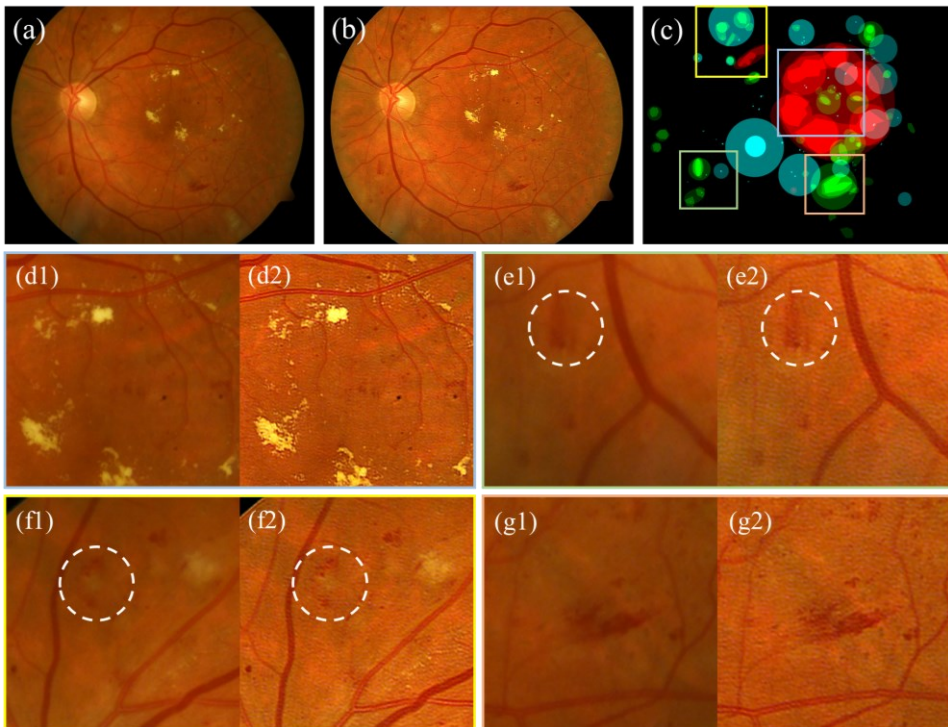


Fig. 14. Enhancement of retinopathy areas using LRRL. (a) Raw image. (b) Restored image. (c) Labels of retinopathy areas. Red: Hard exudates; Green: Hemorrhages; Cyan: red small dots. (d1) to (g2) are zoomed-in pictures for regions in blue, green, yellow, and red boxes.

Blurriness was also corrected, as the white spot of the hard exudates becomes clear. Some hard exudates which can hardly be observed in the raw image shown in Fig. 7-13

(d3) can be clearly found in the restored images in Fig. 7-13 (f1) and (f2). The visual quality of hemorrhages as shown in Fig. 7-13 (e1) and Fig. 7-13 (e2) also increased.

Another group of comparisons is shown in Fig. 14. As shown in Fig. 14 (d1) and 14 (d2), the blurriness was significantly corrected, while no artifacts were introduced. Hemorrhages are shown in Fig. 14 (e1), 14 (e2), 14 (g1) and 14 (g2). The hemorrhage areas and blood vessels in the neighborhood can be clearly observed. The red small spots in the raw image shown in Fig. 14 (f1) are blurred and can hardly be found as they are tiny structures. After the restoration as shown in Fig. 14 (f2), the visibility of the red small spots increased.

In general, LRRL increases the clarity of retinal images which may potentially facilitate the detection of retinopathy. Meanwhile, by comparing LRRL restored images against raw images, we found that LRRL did not lead to additional artifact structure on enhanced images and guarantees the information fidelity in both structure and color aspects.

6. Discussion

6.1 Contributions of LRRL

Our LRRL embeds the illumination correction in the framework of deconvolution, and efficiently avoids the problem where the kernel estimation may be degraded by problematic image intensity. Our experimental results on retinal image blind deconvolution prove the efficiency and correctness of this L_0 -approximation strategy. Our adaptive parameter adjustment largely simplifies the parameter adjustment, and only 2 parameters including α , and \bar{I} need to be manually adjusted for better blind deconvolution results.

In the non-blind deconvolution process, we also adopt extended Richardson-Lucy deconvolution with the Hessian penalty to suppress the impact of noise signals and achieve deconvolution and illumination correction simultaneously. Experimental results show the superiority of LRRL over the other state-of-the-art retinal image blind

deconvolution methods in terms of restoration quality and implementation efficiency, which enables and facilitates the detection of features like drusen, hard exudates, bleeding hemorrhages, or microaneurysm, as these tasks may use more color or local detail information.

6.2 Limitation of LRRL

Currently, the LRRL has the following two limitations. First, being a blind deconvolution algorithm, the LRRL shares limitations with non-learning-based blind-deconvolution methods, where the hyperparameters α , β , and γ for blind deconvolution need to be manually adjusted. According to published literature in blind deconvolution, α is manually adjusted according to the shape and size of the kernels, and also related to the raw blurred images. Fortunately, our experiments on 1719 images show that $\alpha = 0.0025$, $\beta = 2$, and $\gamma = 1$ are appropriate to obtain good deconvolution quality. These values can be used as a reference for further parameter adjustments.

Secondly, the LRRL can be time-consuming as it requires several iterations for solving the latent images, especially for retinal images with large resolutions. In this case, the execution time of LRRL can be significantly increased. As listed in the following tables for kernel size of 51 by 51. (Tested on i9-7940X, 64 GB RAM).

Tab. 7-4. Computational duration for LRRL and PMP methods.

	Pixel size	512 ²	1024 ²	1580 ²	2160 ²
Execution Time (s)	LRRL	34.65	96.82	188.04	423.54
	PMP	46.34	143.79	277.05	914.07

7. Conclusion

In this research, we developed the luminosity rectified blind Richardson-Lucy deconvolution (LRRL) to simultaneously achieve illumination correction and deconvolution. The LRRL corrects the insufficient and uneven intensity distribution and

increases the clarity of input retinal images, benefiting the visual quality of retinal images from both objective and subjective aspect. Although the LRRL has its limitations such as parameters tuning and being time-consuming due to the iterative deconvolution, it is believed that the LRRL supports the clinical diagnosis, and can be improved in future studies.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding. This research is supported by China Scholarship Council (CSC) (201908340078)

References

1. R. Rajalakshmi, V. Prathiba, S. Arulmalar, and M. Usha, "Review of retinal cameras for global coverage of diabetic retinopathy screening," *Eye* **35**, 162-172 (2021).
2. T. Kauppi, V. Kalesnykiene, J.-K. Kamarainen, L. Lensu, I. Sorri, H. Uusitalo, H. Kälviäinen, and J. Pietilä, "DIARETDBO: Evaluation Database and Methodology for Diabetic Retinopathy Algorithms," (2007).
3. F. Huang, B. Dashtbozorg, J. Zhang, E. Bekkers, S. Abbasi-Sureshjani, T. T. Berendschot, and B. M. Ter Haar Romeny, "Reliability of Using Retinal Vascular Fractal Dimension as a Biomarker in the Diabetic Retinopathy Detection," *J Ophthalmol* **2016**, 6259047 (2016).
4. D. M. S. Barros, J. C. C. Moura, C. R. Freire, A. C. Taleb, R. A. M. Valentim, and P. S. G. Morais, "Machine learning applied to retinal image processing for glaucoma detection: review and perspective," *BioMedical Engineering OnLine* **19**, 20 (2020).
5. M. Trieschmann, F. J. van Kuijk, R. Alexander, P. Hermans, P. Luthert, A. C. Bird, and D. Pauleikhoff, "Macular pigment in the human retina: histological evaluation of localization and distribution," *Eye* **22**, 132-137 (2008).
6. E. Pead, R. Megaw, J. Cameron, A. Fleming, B. Dhillon, E. Trucco, and T. J. s. o. MacGillivray, "Automated detection of age-related macular degeneration in color fundus photography: a systematic review," **64**, 498-511 (2019).
7. A. Appaji, B. Nagendra, D. M. Chako, A. Padmanabha, A. Jacob, C. V. Hiremath, S. Varambally, M. Kesavan, G. Venkatasubramanian, S. V. Rao, C. A. B. Webers, T. T. J. M. Berendschot, and N. P. Rao, "Retinal vascular tortuosity in schizophrenia and bipolar disorder," *Schizophr. Res.* **212**, 26-32 (2019).
8. A. Appaji, B. Nagendra, D. M. Chako, A. Padmanabha, A. Jacob, C. V. Hiremath, S. Varambally, M. Kesavan, G. Venkatasubramanian, S. V. Rao, C. A. B. Webers, T. T. J. M. Berendschot, and N. P. Rao, "Examination of retinal vascular trajectory in schizophrenia and bipolar disorder," *Psychiatry Clin. Neurosci.* **73**, 738-744 (2019).
9. P. Zhu, F. Huang, F. Lin, Q. Li, Y. Yuan, Z. Gao, and F. Chen, "The relationship of retinal vessel diameters and fractal dimensions with blood pressure and cardiovascular risk factors," *PloS one* **9**, e106551 (2014).
10. D. Drobnjak, I. C. Munch, C. Glumer, K. Faerch, L. Kessel, M. Larsen, and N. C. Veiby, "Retinal Vessel Diameters and Their Relationship with Cardiovascular Risk and All-Cause Mortality in the Inter99 Eye Study: A 15-Year Follow-Up," *J Ophthalmol* **2016**, 6138659 (2016).
11. S. P. J. P. R. Rajan and I. Analysis, "Recognition of cardiovascular diseases through retinal images using optic cup to optic disc ratio," **30**, 256-263 (2020).
12. E. Stefánsson, O. B. Olafsdottir, T. S. Eliasdottir, W. Vehmeijer, A. B. Einarsdottir, T. Bek, T. L. Torp, J. Grauslund, T. Eysteinnsson, and R. A. Karlsson, "Retinal oximetry: metabolic imaging for diseases of the retina and brain," *Progress in retinal and eye research* **70**, 1-22 (2019).
13. X. Hadoux, F. Hui, J. K. H. Lim, C. L. Masters, A. Pebay, S. Chevalier, J. Ha, S. Loi, C. J. Fowler, C. Rowe, V. L. Villemagne, E. N. Taylor, C. Fluke, J. P. Soucy, F. Lesage, J. P. Sylvestre, P. Rosa-Neto, S. Mathotaarachchi, S. Gauthier, Z. S. Nasreddine, J. D. Arbour, M. A. Rheaume, S. Beaulieu, M. Dirani, C. T. O. Nguyen, B. V. Bui, R. Williamson, J. G. Crowston, and P. van Wijngaarden, "Non-invasive in vivo hyperspectral imaging of the retina for potential biomarker use in Alzheimer's disease," *Nat Commun* **10**, 4227 (2019).
14. C. J. Heaven, J. Cansfield, and K. M. Shaw, "The quality of photographs produced by the non-mydratric fundus camera in a screening programme for diabetic retinopathy: A 1 year prospective study," *Eye* **7**, 787-790 (1993).

15. P. H. Scanlon, C. Foy, R. Malhotra, and S. J. J. D. C. Aldington, "The influence of age, duration of diabetes, cataract, and pupil size on image quality in digital photographic retinal screening," **28**, 2448-2453 (2005).
16. E. Peli and T. Peli, "Restoration of retinal images obtained through cataracts," *IEEE Transactions on Medical Imaging* **8**, 401-406 (1989).
17. C. W. Yang, X. U. Liang, S. Wang, and H. Yang, "The evaluation of screening for cataract needed surgery with digital nonmydriatic fundus camera," *Ophthalmol. China* **19**, 46-49 (2010).
18. L. Xiong, H. Li, and L. Xu, "An Approach to Evaluate Blurriness in Retinal Images with Vitreous Opacity for Cataract Diagnosis," *Journal of Healthcare Engineering* **2017** 1-16 (2017).
19. S. A. Burns, A. E. Elsner, K. A. Sapoznik, R. L. Warner, and T. J. Gast, "Adaptive optics imaging of the human retina," *Progress in retinal and eye research* **68**, 1-30 (2019).
20. M. Foracchia, E. Grisan, and A. Ruggeri, "Luminosity and contrast normalization in retinal images," *Medical Image Analysis* **9**, 179-190 (2005).
21. B. Savelli, A. Bria, A. Galdran, C. Marrocco, M. Molinara, A. e. Campilho, and F. Tortorella, "Illumination correction by dehazing for retinal vessel segmentation," *IEEE 30th International Symposium on Computer-Based Medical Systems* (2017).
22. Y. Ma, J. Liu, Y. Liu, H. Fu, Y. Hu, J. Cheng, H. Qi, Y. Wu, J. Zhang, and Y. Zhao, "Structure and illumination constrained GAN for medical image enhancement," *IEEE Transactions on Medical Imaging* **40**, 3955-3967 (2021).
23. A. W. Setiawan, T. R. Mengko, O. S. Santoso, and A. B. Suksmono, "Color Retinal Image Enhancement using CLAHE," *International Conference on ICT for Smart Society* (2013).
24. L. Cao and H. Li, "Detail-richest-channel based enhancement for retinal image and beyond," *Biomedical Signal Processing and Control* **69**, 102933 (2021).
25. J. Wang, Y.-J. Li, and K.-F. Yang, "Retinal fundus image enhancement with image decomposition and visual adaptation," *Computers in Biology and Medicine* **128**, 104116 (2021).
26. Y. Luo, K. Chen, L. Liu, J. Liu, J. Mao, G. Ke, and M. Sun, "Dehaze of Cataractous Retinal Images using an Unpaired Generative Adversarial Network," *IEEE Journal of Biomedical and Health Informatics* (2020).
27. S. Zhang, C. A. B. Webers, and T. T. J. M. Berendschot, "A double-pass fundus reflection model for efficient single retinal image enhancement," *Signal Processing* **192**, 108400 (2022).
28. H. Li, H. Liu, Y. Hu, H. Fu, Y. Zhao, H. Miao, and J. Liu, "An annotation-free restoration network for cataractous fundus images," *IEEE Transactions on Medical Imaging* (2022).
29. U. Qidwai and U. Qidwai, "Blind Deconvolution for Retinal Image Enhancement," *IEEE EMBS Conference on Biomedical Engineering & Sciences* (2010).
30. A. G. Marrugo, M. S. Millan, M. Sorel, and F. Sroubek, "Retinal image restoration by means of blind deconvolution," *J. of Biomedical Optics* **16**, 116016 (2011).
31. F. J. Ávila, J. Ares, M. C. Marcellán, M. V. Collados, and L. Remón, "Iterative-Trained Semi-Blind Deconvolution Algorithm to Compensate Straylight in Retinal Images," *J. Imaging* **7**, 1-15 (2021).
32. J. Kotera, F. Sroubek, and P. Milanfar, "Blind Deconvolution Using Alternating Maximum a Posteriori Estimation with Heavy-Tailed Priors," *International Conference on Computer Analysis of Images and Patterns* 59-66 (2013).
33. A. Levin, Y. Weiss, F. Durand, and W. T. Freeman, "Understanding and evaluating blind deconvolution algorithms," in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, (IEEE, 2009), 1964-1971.
34. J. Pan, D. Sun, H. Pfister, and M.-H. Yang, "Blind Image Deblurring Using Dark Channel Prior," *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2016).

35. L. Chen, F. Fang, T. Wang, and G. Zhang, "Blind Image Deblurring with Local Maximum Gradient Prior," 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2019).
36. P. Artal, I. Iglesias, and N. L'opez-Gil, "Double-pass measurements of the retinal-image quality with unequal entrance and exit pupil sizes and the reversibility of the eye's optical system," *J. Opt. Soc. Am. A* **12**, 2358-2366 (1995).
37. P. Artal, S. Marcos, R. Navarro, and D. R. Williams, "Odd aberrations and double-pass measurements of retinal image quality," *J. Opt. Soc. Am. A* **12**, 195-201 (1995).
38. J. Pan, Z. Hu, Z. Su, and M.-H. Yang, "Deblurring text images via L0-regularized intensity and gradient prior," in *CVPR*, (2014), pp. 2901–2908.
39. L. Xu, C. Lu, Y. Xu, and J. Jia, "Image smoothing via L₀ gradient minimization," in *Proceedings of the 2011 SIGGRAPH Asia conference*, 2011), 1-12.
40. N. Dey, L. Blanc-F'eraud, C. Zimmer, P. Roux, Z. Kam, J.-C. Olivo-Marin, and J. Zerubia, "3D Microscopy Deconvolution using Richardson-Lucy Algorithm with Total Variation Regularization," (INRIA, 2004).
41. T. Dozat, "Incorporating Nesterov momentum into Adam," (2016).
42. D. J. Jobson, Z.-u. Rahman, and G. A. Woodell, "Properties and Performance of a Center/Surround Retinex," *IEEE Trans. Image Process.* **6**, 451-462 (1997).
43. L. Huang, Y. Xia, and T. Ye, "Effective Blind Image Deblurring Using Matrix-Variable Optimization," *IEEE Trans. Image Process* **30**(2021).
44. F. Wen, R. Ying, P. Liu, and T.-K. Truong, "A Simple Local Minimal Intensity Prior and An Improved Algorithm for Blind Image Deblurring," *IEEE Trans Circuits and Systems for Video Technology* **3**, 2923-2937 (2021).
45. X. Huang, J. Fan, L. Li, H. Liu, R. Wu, Y. Wu, L. Wei, H. Mao, A. Lal, P. Xi, L. Tang, Y. Zhang, Y. Liu, S. Tan, and L. Chen, "Fast, long-term, super-resolution imaging with Hessian structured illumination microscopy," *Nat. Biotechnol.* **36**, 451-459 (2018).
46. J. Immerkær, "Fast Noise Variance Estimation," *Comput. Vis. Image Understand* **64**, 300-302 (1996).
47. T. Kauppi, V. Kalesnykiene, J.-K. Kamarainen, L. Lensu, I. Sorri, J. Pietilä, H. Kälviäinen, and H. Uusitalo, "DIARETDB0 - Standard Diabetic Retinopathy Database Calibration level 0" (2007), retrieved <https://www.it.lut.fi/project/imageret/diaretdb0/>.
48. T. Kauppi, V. Kalesnykiene, J.-K. Kamarainen, L. Lensu, I. Sorri, A. Raninen, R. Voutilainen, H. Uusitalo, H. Kälviäinen, and J. Pietilä, "DIARETDB1 diabetic retinopathy database and evaluation protocol," in *Medical Image Understanding and Analysis*, (Citeseer, 2007), p. 61.
49. "Cataract Dataset" (2019), retrieved <https://www.kaggle.com/jr2ngb/cataractdataset>.
50. E. Decencière, X. Zhang, G. Cazuguel, B. Lay, B. Cochener, C. Trone, P. Gain, R. Ordonez, P. Massin, A. Erginay, B. Charton, and J.-C. Klein, "Feedback on a publicly distributed database: the Messidor database," 2014 **33**, 4 %J *Image Analysis & Stereology* (2014).
51. A. Mitra, S. Roy, S. Roy, and S. K. Setua, "Enhancement and restoration of non-uniform illuminated Fundus Image of Retina obtained through thin layer of cataract," *Computer Methods and Programs in Biomedicine* **156**, 169-178 (2018).
52. K. Panetta, C. Gao, and S. Agaian, "Human-Visual-System-Inspired Underwater Image Quality Measures," *IEEE Journal of Oceanic Engineering* **41**, 1-11 (2016).
53. A. Rizzi, T. Algeri, G. Medeghini, and D. Marini, "A proposal for Contrast Measure in Digital Images," (2004).
54. M. Zhou, K. Jin, S. Wang, J. Ye, and D. Qian, "Color Retinal Image Enhancement Based on Luminosity and Contrast Adjustment," *IEEE Transactions on Biomedical Engineering* (2018).

Supplementary Notes for Chapter 7

LR²L: Luminosity rectified blind Richardson-Lucy deconvolution for single retinal image restoration

Supplementary Note 7-1: Solving latent images in Eq. (7-4) using the Half-Quadratic Splitting method, and its shortages

Recall Eq. (7-4) in the main text, and the problem of the latent image is given by the following

$$E(\vec{\mathbf{O}}, \mathbf{I}) = \|\mathbf{KI}\vec{\mathbf{O}} - \mathbf{S}\|_2^2 + \alpha \|\nabla \vec{\mathbf{O}}\|_0 + \beta \|\nabla \mathbf{I}\|_1. \quad (\text{s7-1})$$

In this case, we have two latent images - \mathbf{O} and \mathbf{I} that need to be solved.

In a traditional blind-deconvolution framework, the latent images are solved using the Half Quadratic Splitting (HQS) method by introducing auxiliary variables corresponding to each latent image. In our case, we introduce two auxiliary variables $\vec{\mathbf{G}}$ with respect to $\nabla \vec{\mathbf{O}}$ and $\vec{\mathbf{L}}$ with respect to $\nabla \mathbf{I}$, and Eq. (s7-1) is converted into the following two groups of optimization problems:

$$\begin{cases} E_o(\vec{\mathbf{O}}) = \|\mathbf{KI}\vec{\mathbf{O}} - \mathbf{S}\|_2^2 + \alpha_0 \|\nabla \vec{\mathbf{O}} - \vec{\mathbf{G}}\|_2^2 \\ E_g(\vec{\mathbf{G}}) = \alpha \|\nabla \vec{\mathbf{O}} - \vec{\mathbf{G}}\|_2 + \alpha_0 \|\vec{\mathbf{G}}\|_0 \end{cases}, \quad (\text{s7-2})$$

and

$$\begin{cases} E_i(\vec{\mathbf{I}}) = \|\mathbf{K}\mathbf{O}\vec{\mathbf{I}} - \mathbf{S}\|_2^2 + \beta_0 \|\nabla \vec{\mathbf{I}} - \vec{\mathbf{L}}\|_2^2 \\ E_l(\vec{\mathbf{L}}) = \beta \|\nabla \vec{\mathbf{I}} - \vec{\mathbf{L}}\|_2 + \beta_0 \|\vec{\mathbf{L}}\|_1 \end{cases}. \quad (\text{s7-3})$$

Where α_0 and β_0 are sufficient large penalty parameters such that they enforce the L_2 -norm terms approach to 0.

The L_0 -norm problem in Eq. (s7-2) can be approximately solved using the hard-threshold method, while the L_1 -norm problem in Eq. (s7-3) can be solved using the soft-threshold method. Since solving Eq. (s7-2) and Eq. (s7-3) are quite similar, in the following analysis, we consider only the problem in Eq. (s7-2).

In a typical HQS framework, \mathbf{O} and \mathbf{G} are iteratively solved. For example, first, we assume an initial value of \mathbf{O} and use it to solve the value of \mathbf{G} using the hard-threshold method [1, 2]:

$$\mathbf{G} = \begin{cases} \nabla \mathbf{O}, & |\nabla \mathbf{O}|^2 \geq \alpha / \alpha_0 \\ 0, & \text{else} \end{cases} \quad (\text{s7-4})$$

With the known of \mathbf{G} we solve the E_o setting the derivative of E_o with respect to \mathbf{O} to zero and set several iterations in which α_0 gradually becomes larger and larger to ensure the solution convergence [1, 3]. However, in our case the derivation of E_o with respect to \mathbf{O} is

$$\begin{aligned} \frac{\partial E}{\partial \vec{\mathbf{O}}} &= 2(\mathbf{I}^T \mathbf{K}^T \mathbf{K} \mathbf{I} + \alpha_0 \nabla^T \nabla) \vec{\mathbf{O}} - 2\mathbf{I}^T \mathbf{K}^T \mathbf{S} - 2\alpha_0 \nabla^T \vec{\mathbf{G}} = 0, \\ \Rightarrow (\mathbf{I}^T \mathbf{K}^T \mathbf{K} \mathbf{I} + \alpha_0 \nabla^T \nabla) \vec{\mathbf{O}} &= \mathbf{I}^T \mathbf{K}^T \mathbf{S} + \alpha_0 \nabla^T \vec{\mathbf{G}} \end{aligned} \quad (\text{s7-5})$$

and cannot be calculated using Fourier transform methods due to the existence of \mathbf{I} , and also the calculation of the inverse matrix of $(\mathbf{I}^T \mathbf{K}^T \mathbf{K} \mathbf{I} + \alpha_0 \nabla^T \nabla)$ is hard since the matrix $\mathbf{I}^T \mathbf{K}^T \mathbf{K} \mathbf{I}$ is very large that it requires large storage and calculation durations [4]. For most of the retinal images, the size is larger than 1024 by 1024, therefore, the size of \mathbf{K} is larger than 1048576 by 1048576.

An alternative method to solve Eq. (s7-5) is using the conjugate gradient, and at least 20 iterations are needed to guarantee convergence. Therefore, the process for solving \mathbf{O} can be summarized in **Algorithm s7-1**.

The executive efficiency of **Algorithm s7-1** is much lower than that of **Algorithm 7-1** in the main text. For example, $\alpha_{\max} = 10^5$ and $a = 2$ is a comment setting [1]. If $\alpha_0 = 0.01$, it takes 24 iterations, and therefore the total number of iterations is at least $24 \times 20 = 480$. While in **Algorithm 7-1** in the main text only 25 iterations are needed. A similar analysis applies to the solution of Eq. (s7-3).

Algorithm s7-1: Half quadratic splitting for Eq. (s7-2)

Input: Raw image S , illumination pattern I , point spread function K , penalty parameters α , $\alpha_0 \leftarrow \alpha$,

While $\alpha_0 < \alpha_{\max}$ **do**

Solving G via Eq. (s7-4) with given O

$\beta_0 \leftarrow \beta$

For $t=1, 2, 3, 4 \dots \dots \max_iteration$

Solving O using Eq. (s7-5) and conjugate gradient method.

End For

$\alpha_0 \leftarrow a \cdot \alpha_0$

End While

Output: Latent image O

In conclusion, the HQS can be used for solving the latent images, but the computational efficiency is way lower than our proposed Richardson-Lucy method.

7

Supplementary Note 7-2: Exponential function as the approximation of L_0 -norm

7-2.1 Principle

The L_0 -norm of a matrix denotes the accounting number of non-zero elements. Since the L_0 -norm is performed in an element-wise manner, the function for an individual element x can be given by

$$y = L_0(x) = \begin{cases} 0, & x = 0 \\ 1, & \text{else} \end{cases} . \quad (\text{s7-6})$$

The shape of the function (s7-6) is a “T” as shown in Fig. 7-3(a) in the main text. Our goal is to find a function that is differentiable and also approaches the T-shape L_0 -norm function, and here we use $f(x) = 1 - \exp(-c|x|)$.

This function is able to approximate the L_0 -norm with a given value of c . For example, as the c increases, the curve of $f(x)$ gets closer and closer to the black “T” as shown in Fig. 7-3(a) in the main text. With an appropriate value of c we are able to perform $f(x)$ to every element of the matrix to approximately represent its L_0 -norm. Figure s1 represents how the $f(x)$ can approximate the L_0 -norm of the matrix \mathbf{X} . Since $f(x)$ is a non-negative function, the element-wise summation of $\mathbf{Y} = f(\mathbf{X})$ equals to $\|\mathbf{Y}\|_1$.

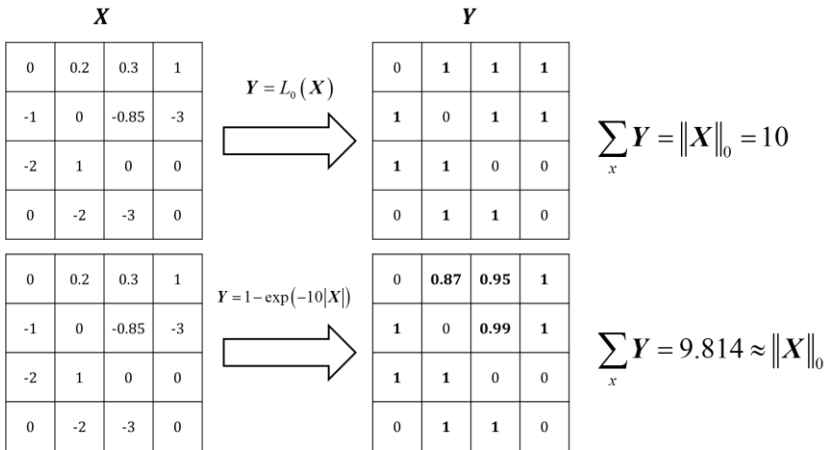


Fig. s7-1, representation of the approximation of L_0 -norm using $f(x) = 1 - \exp(-c|x|)$. $c = 10$.

According to Fig. 7-3(b) in the main text, the gradient for $f(x)$ becomes large when x approaches 0, therefore a very small step-size needs to be used for conventional gradient descent methods. However, the small step-size also limits the convergent speed of the optimization. In order to achieve fast convergence while avoiding the gradient busting problem, we propose to use the adaptive moment estimation with Nesterov Acceleration (N-Adam). Parameter θ is updated according to

$$\left\{ \begin{array}{l} \mathbf{v}_t = \rho_1 \cdot \mathbf{v}_{t-1} + (1 - \rho_1) \cdot \frac{\partial E}{\partial \mathbf{o}} \\ \mathbf{v}_t = \rho_2 \cdot \mathbf{v}_{t-1} + (1 - \rho_2) \cdot \left(\frac{\partial E}{\partial \mathbf{o}} \right)^2 \\ \hat{\mathbf{v}} = \mathbf{v}_t / (1 - \rho_1^t) \\ \hat{\mathbf{v}} = \mathbf{v}_t / (1 - \rho_2^t) \\ \mathbf{o}_{t+1} = \mathbf{o}_t - \delta \cdot \frac{1}{\sqrt{\hat{\mathbf{v}} + \xi}} \left[\rho_1 \cdot \hat{\mathbf{v}} + (1 - \rho_1) \cdot \frac{\partial E}{\partial \mathbf{o}} \right] \end{array} \right. , \quad (\text{s7-7})$$

where \mathbf{v}_t ($\mathbf{v}_0 = 0$) is the first momentum estimation, and \mathbf{v}_t ($\mathbf{v}_0 = 0$) is the second-momentum estimation. $\hat{\mathbf{v}}$ and $\hat{\mathbf{v}}$ are bias-corrected of \mathbf{v}_t and \mathbf{v}_t respectively. ρ_1 and ρ_2 are decay rates for the momentum estimations. $\rho_1 = 0.9$ and $\rho_2 = 0.999$ are used in the following experiments. $\delta = 0.05$ is the step size (learning rate). ξ is a small value to avoid the condition of dividing by 0.

7-2.2 Validation

To validate our proposed L_0 -norm approximation strategy, we considering a 2D image smoothing task using gradient L_0 -regularization:

$$E(\mathbf{o}) = \|\mathbf{o} - \mathbf{s}\|_2^2 + \alpha \|\nabla \mathbf{o}\|_0, \quad (\text{s7-8})$$

where \mathbf{s} denotes the observed image, and \mathbf{o} is the smoothed image. The first term is the data fidelity term, and the second term denotes the gradient L_0 -regularization.

The classical solution for Eq. (s7-8) is using half quadratic splitting method and is transformed into two sub-problems by introducing an additional variable, while the L_0 -norm can be achieved using the Hard-threshold method. In addition, Eq. (s7-8) can be also solved using our proposed L_0 -norm approximation function using adaptive moment gradient descent with Nesterov Acceleration (N-Adam). The cost function is rewritten as

$$E(\mathbf{o}) = \|\mathbf{o} - \mathbf{s}\|_2^2 + \alpha \left[1 - \exp(-c|\nabla \mathbf{o}|) \right], \tag{s7-9}$$

and the (sub)gradient of Eq. (s9) is calculated as

$$\frac{\partial E}{\partial \mathbf{o}} = 2(\mathbf{o} - \mathbf{s}) + \alpha \cdot c \cdot \sum_{z \in \{x,y\}} \nabla_z^T \left[\exp(-c|\nabla_z \mathbf{o}|) \circ \text{sign}(\nabla_z \mathbf{o}) \right], \tag{s7-10}$$

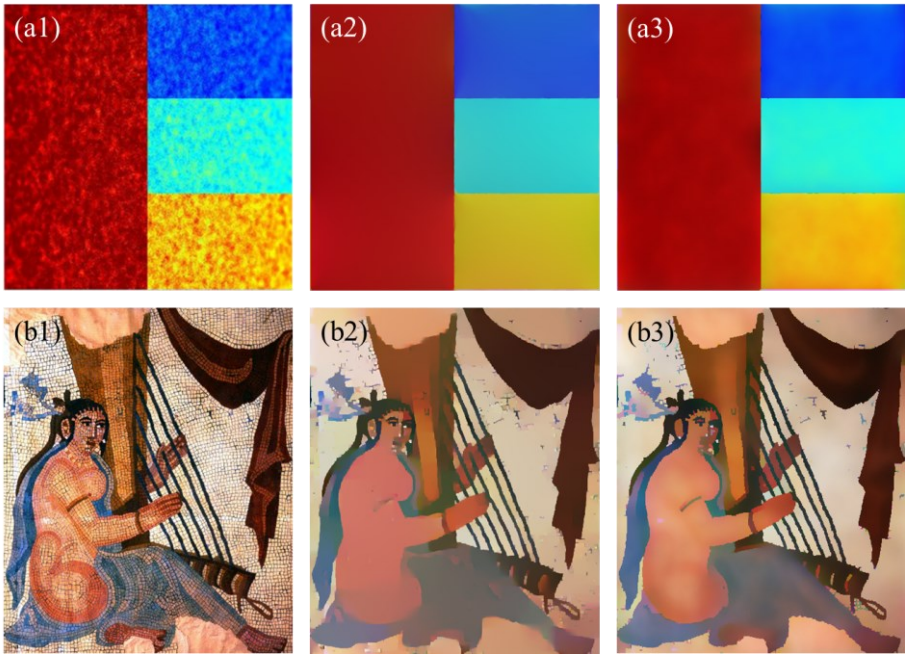


Fig. s7-2. Image smoothing using gradient L_0 -regularization. (a1) and (b1) are raw images. (a2) and (b2) are results using HQS methods. $\alpha = 0.1$ for (a2), $\alpha = 1$ for (a3). (a3) and (b3) are results using our proposed approximation. $\alpha = 0.04$ for (a3), $\alpha = 0.4$ for (b3).

Fig. s7-2 (a1) and Fig. s7-2 (b1) are the raw images while tiny structures exist. The results using the traditional HQS method are shown in Fig. s7-2 (a2) and Fig. s7-2 (b2) where the small gradient pixels are erased, remaining the color block structures. Our proposed L_0 -approximation achieves similar results to the traditional HQS method [1]. Since the penalty strength for the hard-threshold and proposed approximation method

are different as the hard-threshold is stronger than the approximations, a large value of the penalty parameter α is assigned to the approximation methods in order to achieve a similar regularization effect.

7-2.3 NAdam VS Other optimizers including SGD, AdaDelta, Adam.

In this section, we compare the performance of different optimizers including SGD, AdaDelta, Adam, and NAdam on our proposed L_0 -norm approximation method with different values of α .

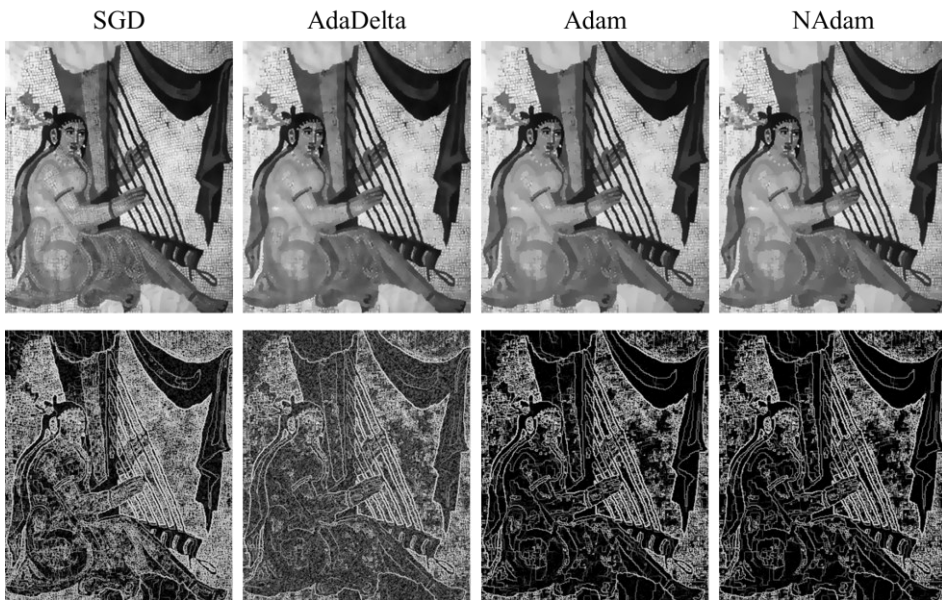


Fig. s7-3. Solving problem in (s7-8) using different optimizers. First row: output images. Second row: gradient maps. $\alpha = 0.01$.

When $\alpha = 0.01$, the contribution of gradient from the exponential term becomes less. The output images are shown in the first row of Fig. s7-3 for different optimizers. The learning rate is 0.01 for SGD, Adam, and NAdam. A total of 50 iterations. The gradient maps for each image are also shown in the second row of Fig. s7-3. The

gradient map is calculated as $\log(|D_x o| + |D_x o| + \eta)$, and $\eta = 10^{-5}$ to avoid $\log(0)$. Pixel values that are less than -5 are not shown. According to the gradient map, both NAdam and Adam generate a sparser image gradient, while SGD and AdaDelta optimizers fail to obtain the sparse gradient map and more iterations are needed.

When $\alpha = 0.2$, the contribution of gradient from the exponential term becomes prominent and the gradient exploding problem can happen. In these cases, the SGD optimizer cannot obtain stable results as shown in the first column of Fig. s7-4. The convergence of AdaDelta is not good as shown in the second column of Fig. s7-4. Both Adam and NAdam obtain promising results as the gradient maps, as shown in the third and fourth columns of Fig. s4, are sparser than that of AdaDelta.

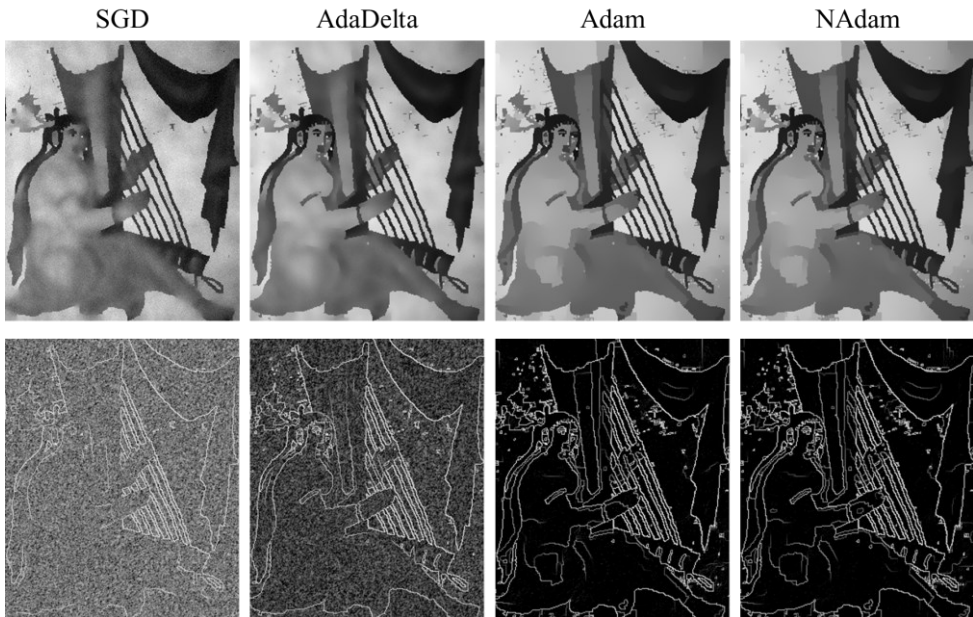


Fig. s7-4. Solving problem in (s7-8) using different optimizers. First row: output images. Second row: gradient maps. $\alpha = 0.2$.

We also plotted the change of cost function with respect to iterations for Fig. s7-3 and Fig. s7-4. As shown in Fig. s7-5, the NAdam has the fastest convergence speed among these four optimizers, while both Adam and NAdam can converge to a similar value of cost function which is smaller than that of AdaDelta and SGD. Accordingly, we choose NAdam in our research due to its faster convergent speed, better convergence properties (lower energy), and implementation efficiency.

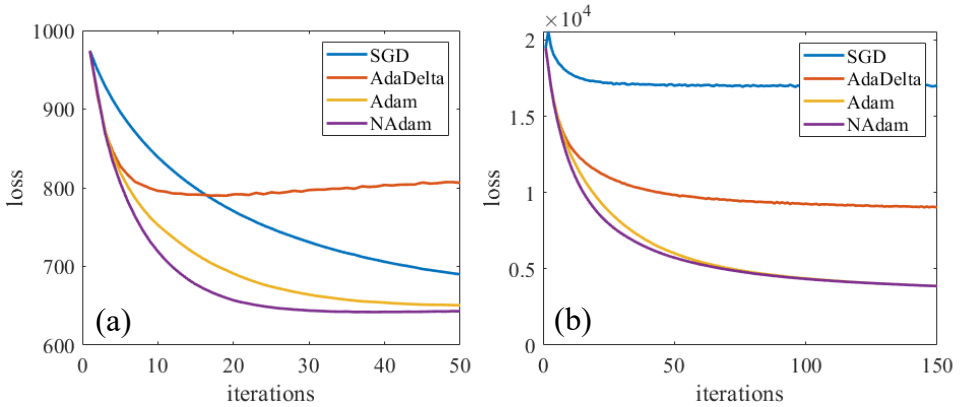


Fig. s7-5. Change of cost function with respect to iterations. (a) for Fig. s7-3, $\alpha = 0.01$. (b) for Fig. s7-4, $\alpha = 0.2$.

Supplementary Note 7-3: Calculation of four quality matrixes

The image definition [5] implies the richness of texture information which is given by the first-order gradient of the image:

$$DE = \sum_{c \in \{R, G, B\}} \lambda_c \cdot \frac{1}{MN} \sum_{x=1}^M \sum_{y=1}^N \sqrt{\frac{|\nabla I_c|^2}{2}}, \quad (s7-11)$$

where I_c is the c channel of input images. The DEs are evaluated in each RGB channel and are combined linearly with coefficients λ_c since the different color channel has different visual response. $\lambda_r = 0.299$, $\lambda_g = 0.587$ and $\lambda_b = 0.114$ are used according to the relative visual responses of the red, green, and blue channels.

Image sharpness is the attribute related to the preservation of fine details and edges. Blurred images have low sharpness score than clear images. In this study, we adopted the underwater image sharpness measure (UISM) to measure the sharpness of the retinal image [6]. In UISM, the Sobel edge detector is first applied on each RGB color channel. The resultant edge map is then multiplied with the original image to get the grayscale edge map in the corresponding color channel. By doing this, only the pixels on the edges of the original retinal image are preserved. The UISM is then given by

$$UISM = \sum_{c \in \{R, G, B\}} \lambda_c \cdot \left[\frac{2}{k_1 k_2} \sum_{l=1}^{k_1} \sum_{k=1}^{k_2} \log \left(\frac{I_{\max, k, l}^{edge}}{I_{\min, k, l}^{edge}} \right) \right], \quad (s7-12)$$

where the image is divided into $k_1 k_2$ blocks, $I_{\max, k, l}^{edge} / I_{\min, k, l}^{edge}$ indicates the relative contrast ratio within each block of the grayscale edge map.

Image entropy (IE) describes the randomness distribution of the image and its value denotes the amount of image information, which is given by

$$IE = \sum_{c \in \{R, G, B\}} \lambda_c \sum_{g=0}^{255} P(x_g) \log(x_g), \quad (s7-13)$$

where $P(x_g)$ is the probability of the appearance of the pixels that have gray-value g in the gray-scaled image. Image entropy can be used to characterize the texture of the image and determine the amount of image information.

The multiscale-contrast of the image, C_{RAMM} , is calculated with a pyramidal multi-resolution representation of luminance. C_{RAMM} is defined as [7]

$$C_{RAMM} = \frac{1}{8(\# \text{ pixel})(\# \text{ levels})} \sum_{\forall \text{ level}} \left[\sum_{\forall \text{ pixel}} \left(\sum_{8\text{-neigh}} |P_i - P_j| \right) \right], \quad (s7-14)$$

where $\# \text{ pixel}$ denotes the total pixel numbers of the image, $\forall \text{ level}$ denotes the total level of down-sampling, in each level the image is halved without pre-filtering. In this paper, we use 6 levels of down-sampling. The pixel numbers of retinal images in the

simulation and databases are large enough for 6 levels of down-sampling. During the calculation, the down-sampling process is stopped if either the row pixel numbers, or the column pixel numbers of the next down-sampled image are less than 3.

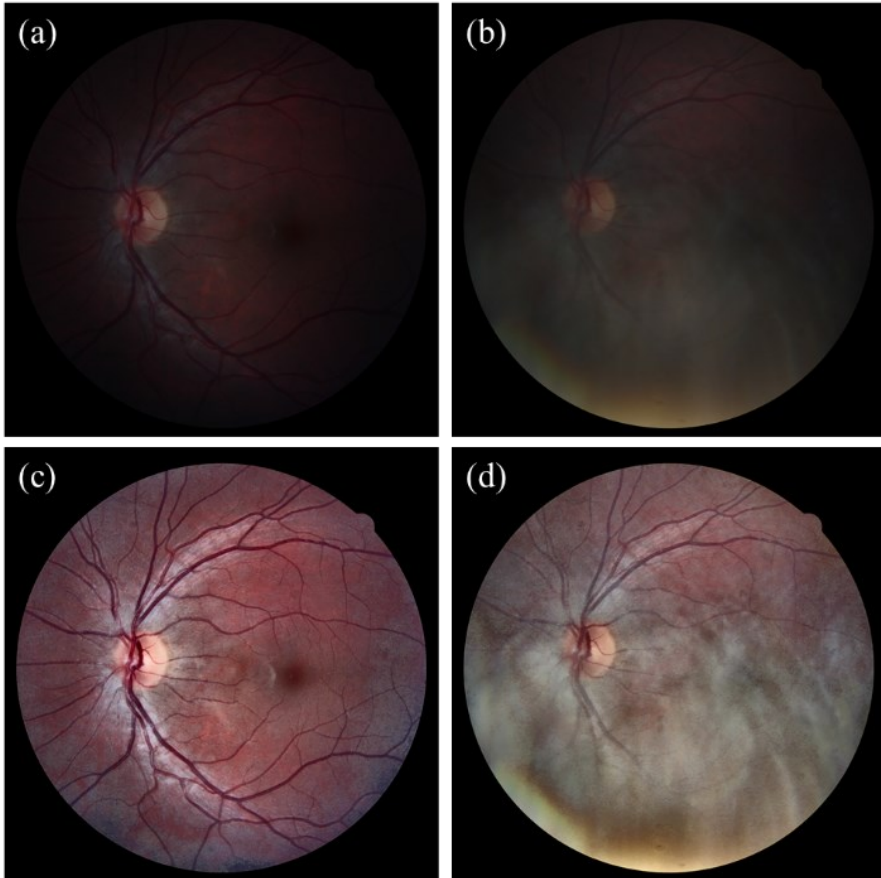


Fig. s7-6. (a) GQ image; (b) BG image. (c) enhanced image for (a); (d) enhanced image for (b)

We use the non-reference-based evaluation and don't use the reference-based evaluation such as PSNR and SSIM due to following reasons:

- 1) Unlike natural scene images, a clear un-degraded retinal image can hardly be obtained in real life. For example, images in the HRF dataset

(<https://www5.cs.fau.de/research/data/fundus-images/>) that are regarded as being of good quality still suffer from problematic illumination as shown in the following picture. Using reference-based evaluation can lead to the wrong conclusion with can conflict with a human's visual perspective.

- 2) The non-reference-based evaluation Metrics used in our study are designed based on a human-visual perspective, as the evaluation scores are related to the human's visual feelings. In this way, we are able to evaluate the quality of retinal images that are more related to human experts. These metrics are well-defined and do not depend on the statistical feature of certain retinal images dataset.

To illustrate the problem for reference-based evaluation, we collect the good quality (GQ) images (18 images) and corresponding bad quality images (BQ) (18 images) from the HRF dataset. We also perform illumination correction and contrast enhancement using CLAHE on both GQ and BQ images. One example is shown in Fig. s7-6.

Then we calculate the PSNR and SSIM between (1) GQ images and BQ images, (2) GQ images and enhanced GQ images, (3) GQ images and enhanced BQ images as listed in the Tab. s7-1. In all these calculations, the GQ images are set as the reference images.

Tab. s7-1. PSNR and SSIM for images in Fig. s7-6

Image pairs (input, ref)	PSNR	SSIM
(b), (a)	22.9488	0.8116
(c), (a)	12.8905	0.5078
(d), (a)	11.1248	0.5053

The table implies that Fig. s7-6 (c) and Fig. s7-6 (d) have worse image quality than that of Fig. s7-6 (b), which conflicts with human visual assessment. In this case, the decrease in PSNR and SSIM is largely due to the change of the image's intensity level, especially for SSIM as it depends on the intensity level between the given image and the reference image.

In addition, we have analyzed the statistical distribution of average intensity level for all 1200 images from the Messidor retinal image dataset (the black background of each retinal image is not taken into account). As shown in Fig. s7-7, almost 50 % of retinal images have intensity problems of different degrees as their average intensities are less than 0.4.

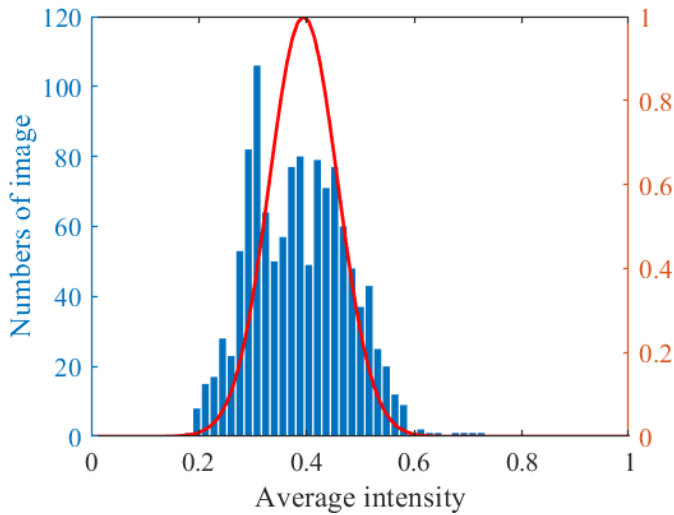


Fig. s7-7. Average intensity of retinal images from Messidor dataset

Since there are a large among of retinal images in our experimental data that have illumination problems, using PSNR and SSIM to assess the performance of our proposed model can lack fairness and practicality, especially for medical images where the real ground truth is not well-defined.

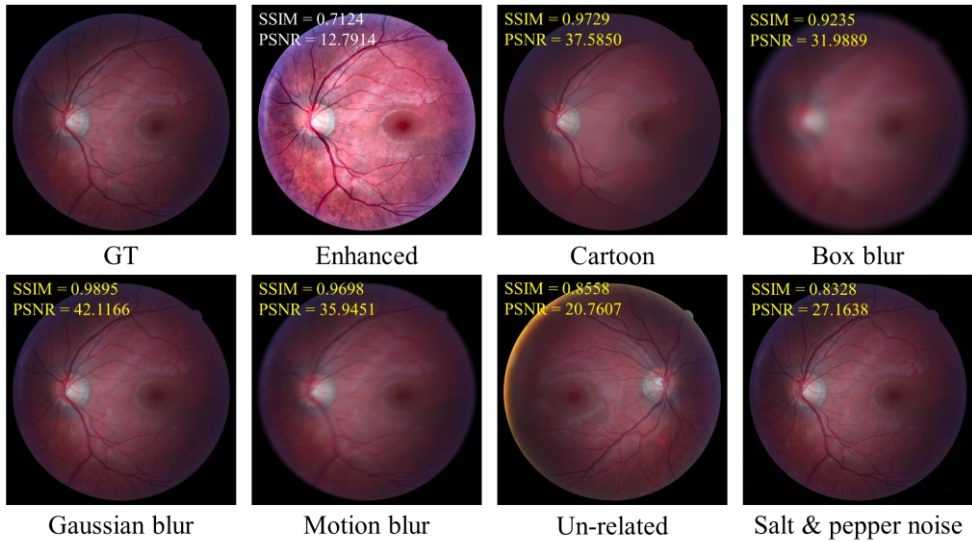


Fig. s7-8. Comparison of images with different types of distortions and their SSIM and PSNR scores.

Additional examples that demonstrate the unpredictable behavior of SSIM and PSNR are shown in Fig. s7-8. Even the image is degraded by some major types of distortions such as Gaussian blur, motion blur, and noises, the related SSIM and PSNR scores can be larger than the enhanced image, which far contradicts to human visual prospect. More related works showing the drawback of SSIM and PSNR can be found in [8-11].

Supplementary Note 7-4: Additional experimental results

7-4.1 Visual assessments

In this section, we show additional four groups of experimental results as shown in Fig. s7-9 (a1) to (d1), where the raw images are blurred and have illumination problems including uneven illumination, as shown in Fig. s7-9 (a1) and Fig. s7-9 (b1) or insufficient illuminations, as shown in Fig. s7-9 (c1). The latent illumination patterns are shown in the second row of Fig. s7-9. The last row shows the output of LRRL

method, where the images have good illumination conditions, and the clarities are also improved. Fig. s7-10 shows the estimation of kernels for using LRRL (first row), PMP (second row), and SLC (last row) methods. For each image, we enlarge parts in the white and yellow boxes to show detailed changes of structures.

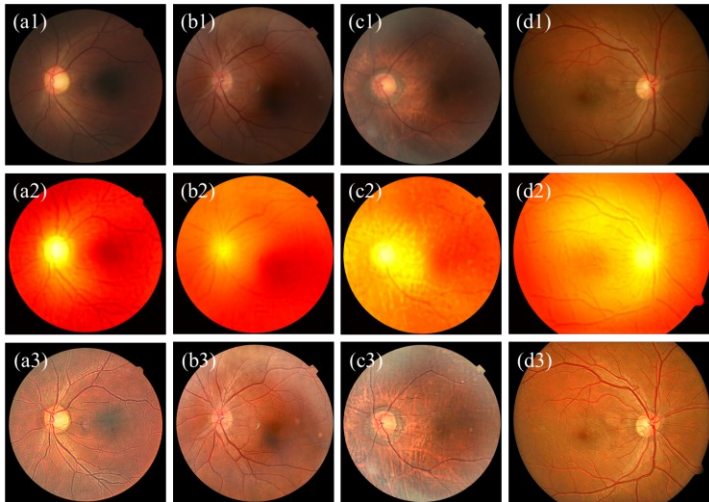


Fig. s7-9. (a1) to (d1) are raw images. (a2) to (d2) are latent illumination pattern. (a3) to (d3) are outputs of LRRL methods.

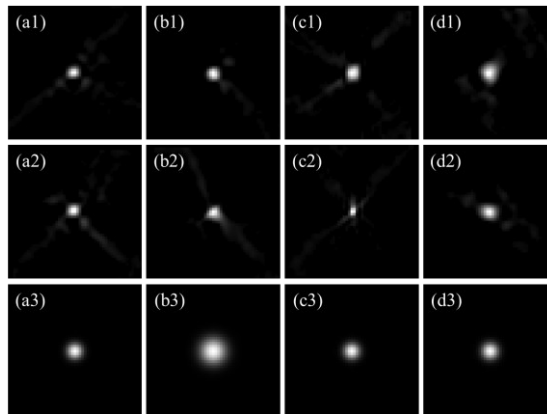


Fig. s7-10. Estimation of blurriness kernel using (a1)-(d1) LRRL, (a2)-(d2) PMP, (a3)-(d3) SLC.

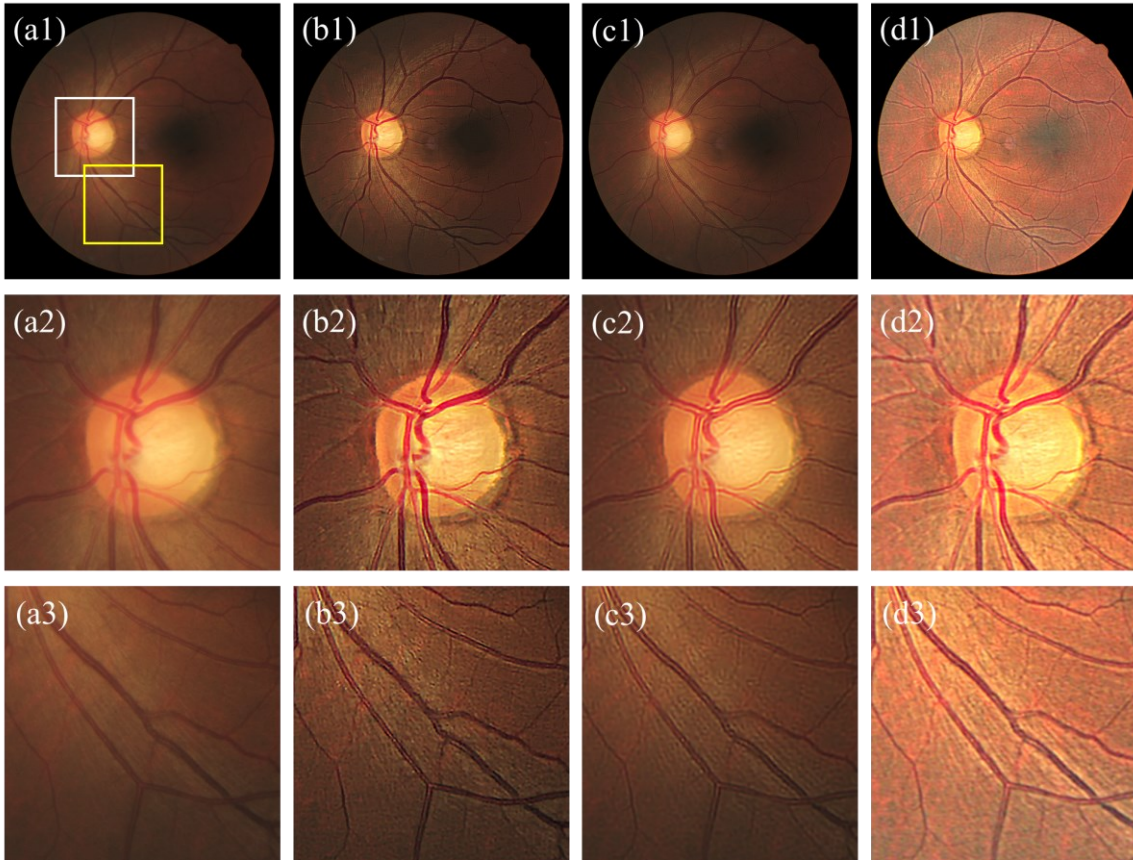


Fig. s7-11. Results of image in Fig. s7-9 (a1). (a1) to (d1) are raw image, results of PMP, SLC, and LRRL methods, respectively. (a2) to (d2) are zoomed-in parts for the white box. (a3) to (d3) are for the yellow box.

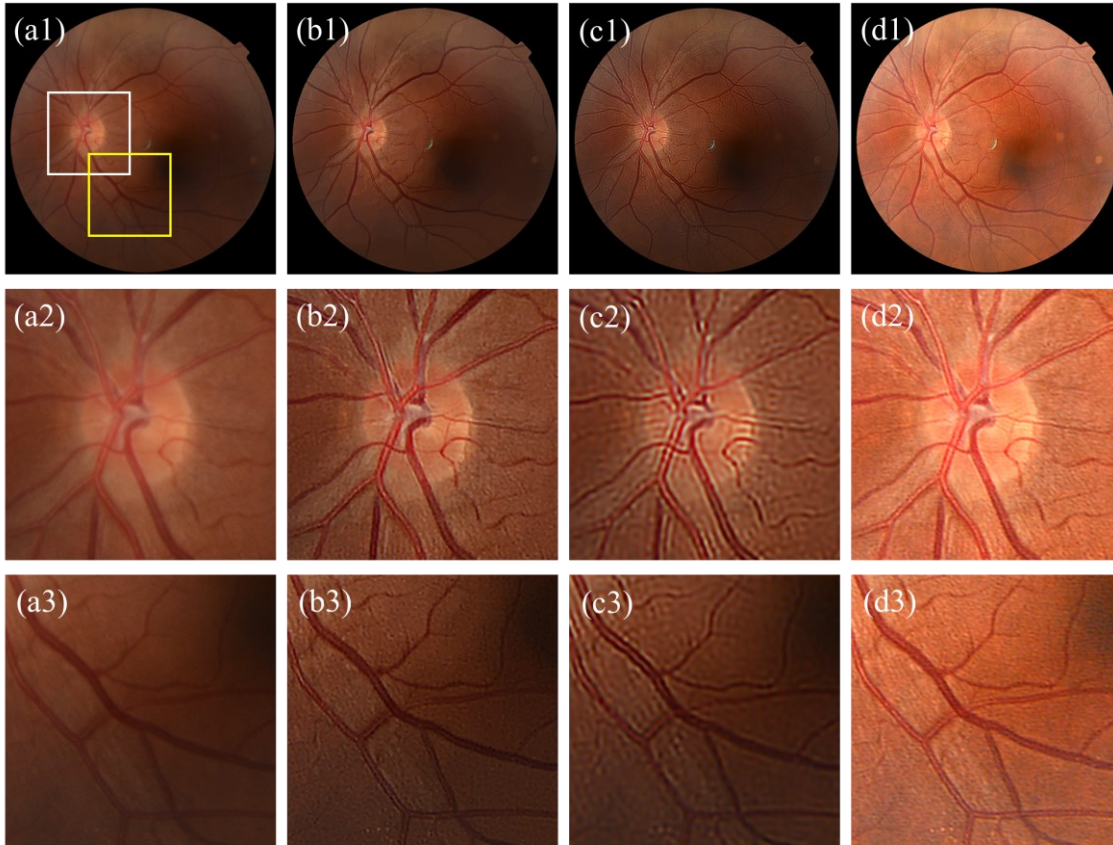


Fig. s7-12. Results of image in Fig. s7-9 (a2). (a1) to (d1) are raw image, results of PMP, SLC, and LRRL methods, respectively. (a2) to (d2) are zoomed-in parts for the white box. (a3) to (d3) are for the yellow box.

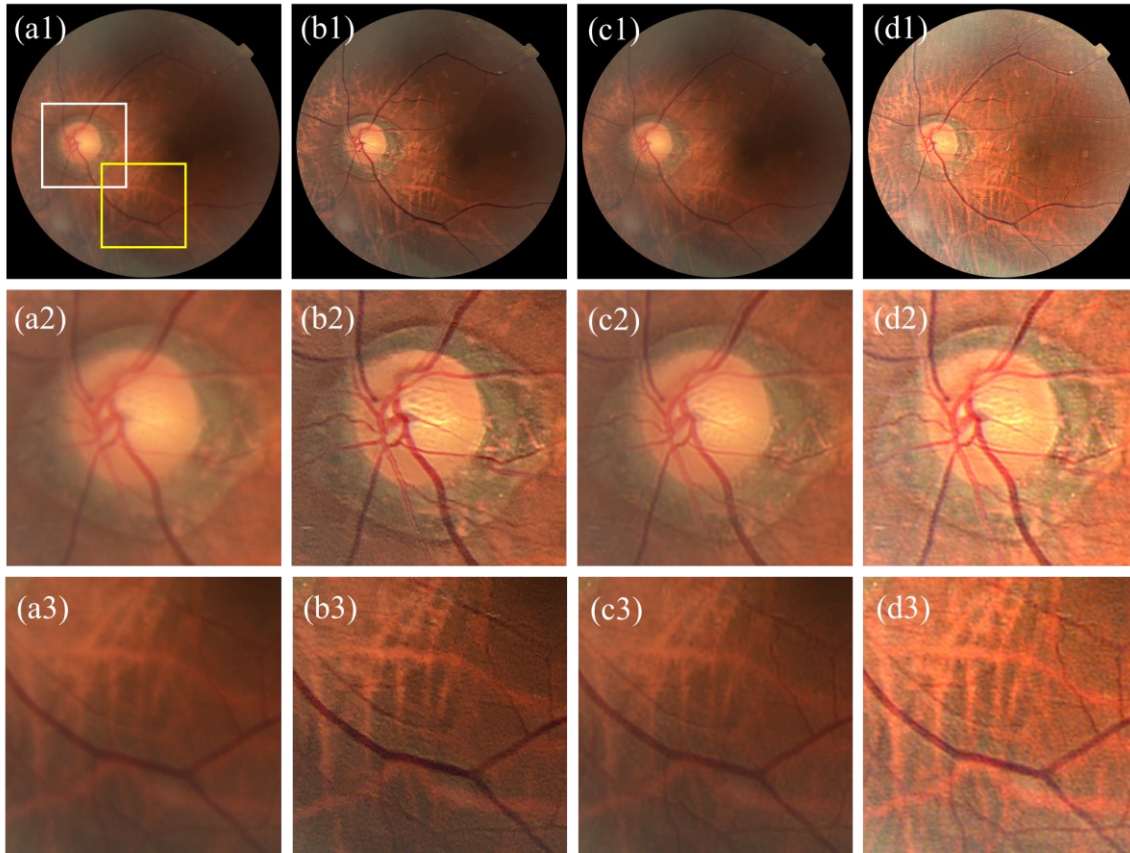


Fig. s7-13. Results of image in Fig. s7-9 (a3). (a1) to (d1) are raw image, results of PMP, SLC, and LRRL methods, respectively. (a2) to (d2) are zoomed-in parts for the white box. (a3) to (d3) are for the yellow box.

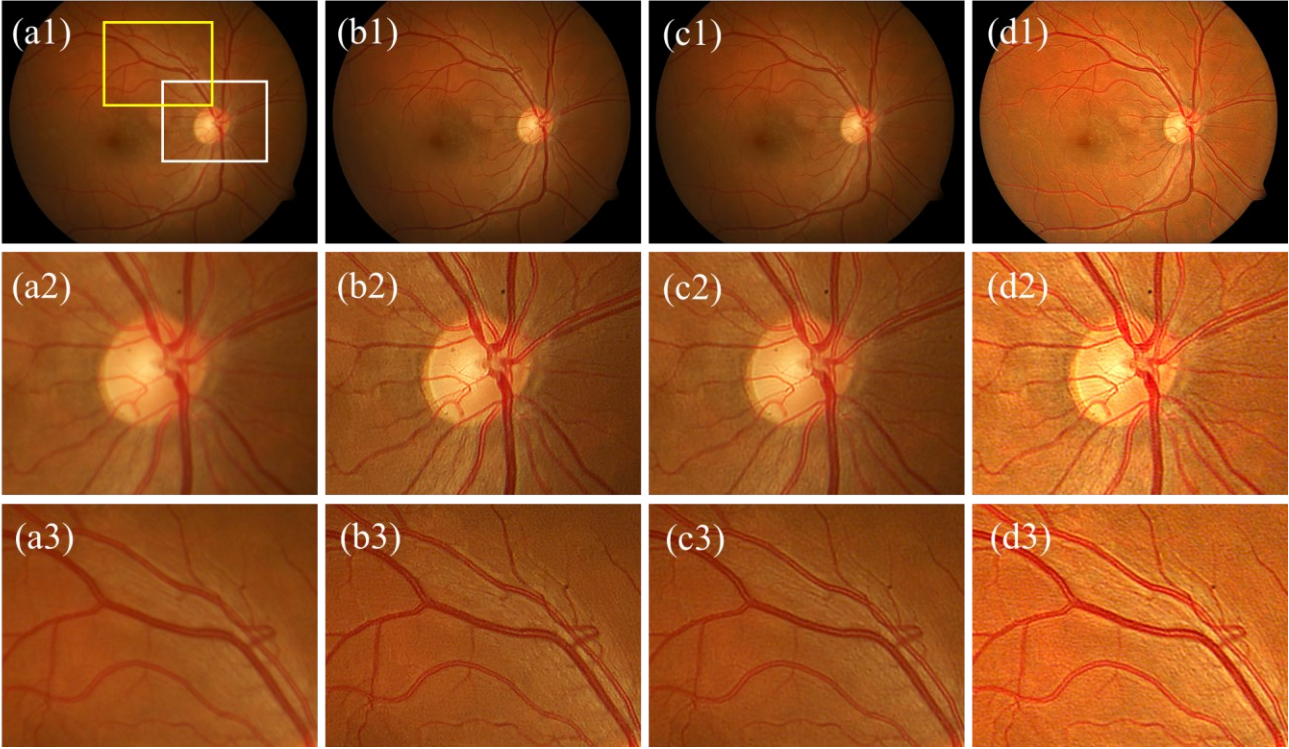


Fig. s 7-14. Results of image in Fig. s7-9 (a4). (a1) to (d1) are raw image, results of PMP, SLC, and LRRLL methods, respectively. (a2) to (d2) are zoomed-in parts for the white box. (a3) to (d3) are for the yellow box.

7-4.2 Objective assessments

The Tab. s7-2 lists the objective assessment of the four groups from Fig. s7-11 to Fig. s7-14.

Tab. s7-2. Quantitative assessment of the enhanced images among Fig. s7-9 (a1) to (d1).

Raw image	Matrixes	Raw	Methods		
			PMP	SLC	LRRL
Fig. s7-9 (a1)	DE	1.0449	2.8778	1.4811	4.0610
	UISM	1.8927	5.2896	3.6692	4.5101
	IE	5.7818	6.0520	5.8422	6.3562
	C _{RAMM}	1.1888	1.8050	1.3743	2.9894
Fig. s7-9 (b1)	DE	1.2653	2.4638	1.6520	4.2269
	UISM	1.6671	4.7171	3.7081	4.5685
	IE	6.2798	6.4050	6.3869	6.7220
	C _{RAMM}	1.3142	1.7159	1.6795	2.7714
Fig. s7-9 (c1)	DE	1.2868	3.0557	1.0147	3.6350
	UISM	1.3371	4.6375	3.3394	4.4732
	IE	6.4412	6.5684	6.4529	6.7740
	C _{RAMM}	1.5473	2.0436	1.6045	3.0164
Fig. s7-9 (d1)	DE	1.3747	2.2809	1.4154	4.7336
	UISM	2.3511	4.0714	2.4111	4.1135
	IE	6.2428	6.3501	6.2981	6.5461
	C _{RAMM}	1.0203	1.3097	1.1572	2.1570
Average	DE	1.2429	2.6696	1.3908	4.1641
	UISM	1.8120	4.6789	3.2820	4.4163
	IE	6.1864	6.3439	6.2450	6.5996
	C _{RAMM}	1.2677	1.7186	1.4539	2.7336

Supplementary Note 7-5: Richardson-Lucy deconvolution and its variation

We assume that the image formation model is described by a convolution process which is given as

$$s(x, y) = p(x, y) \otimes o(x, y), \quad (\text{s7-15})$$

where s is the observed image, o is the original image and p is the convolution kernel. x and y are the spatial coordinates. \otimes denotes the 2D convolution. Eq. (s7-15) ignores the noise signals.

7-5.1 Standard RL deconvolution: in the presence of the Poisson process

If the image is corrupted by Poisson noise, s meets the Poisson distribution. The deconvolution process where we wish to solve the o under the observation of s equals to maximize the probability $\mathcal{P}(o|s)$. Based on the Bayesian approach, this can be rewritten as [12]

$$\begin{aligned} o &= \arg \max \mathcal{P}(o|s) = \arg \max \mathcal{P}(s|o) \frac{\mathcal{P}(o)}{\mathcal{P}(s)} \\ &= \arg \max \frac{\mathcal{P}(o)}{\mathcal{P}(s)} \prod_{x,y} \frac{(p \otimes o)^s}{s!} \exp(-p \otimes o) \end{aligned}, \quad (\text{s7-16})$$

where $\mathcal{P}(s|o)$ is the likelihood probability, $\mathcal{P}(o|s)$ is the posterior probability. $\mathcal{P}(o)$ is the prior of image o . $\mathcal{P}(o)$ is assumed to be a constant for simplicity. Taking the logarithm to both sides of Eq. (s7-16) and ignoring the constant term yielding:

$$\begin{aligned} o &= \arg \min [-\log \mathcal{P}(o|s)] \propto \arg \min \mathcal{L}(o) \\ \text{where } \mathcal{L}(o) &= \int_x \int_y [p \otimes o - s \log(p \otimes o)] dx dy \end{aligned}, \quad (\text{s7-17})$$

$\mathcal{L}(o)$ is the cost function to be minimized. By setting the derivative of $\mathcal{L}(o)$ with respect to o to zeros, and using $\int_x \int_y p dx dy = 1$, we obtain:

$$\frac{\partial \mathcal{L}}{\partial o} = 1 - p^* \otimes \frac{s}{p \otimes o} = 0 \Rightarrow 1 = p^* \otimes \frac{s}{p \otimes o}, \quad (\text{s7-18})$$

where $p^* = p(-x, -y)$. By assuming that at convergence the ratio $o_{k+1} / o_k = 1$, Eq. (s7-18) can be rewritten as

$$o_{k+1} = o_k \left(p^* \otimes \frac{s}{p \otimes o_k} \right), \quad (\text{s7-19})$$

which is the standard form of multiplicative RL deconvolution. Eq. (s7-19) has three main drawbacks. The first one is that it amplifies the noise signal and may not converge to a suitable solution. The second one is that it converges slow. Finally, due to the divided form in Eq. (s7-19), it may result in problematic pixels where zero values are present to the denominator, and the convergence behavior of Eq. (s7-19) is thus unstable. The pixel value should be manually truncated to avoid the problem of dividing by zeros.

To avoid the problem, we can use the additive form of the algorithm. Following the conventional gradient descent method, we are able to use $o_{k+1} = o_k - \delta \cdot \partial_o \mathcal{L}$ to update the parameters yielding the additive-form RL deconvolution:

$$o_{k+1} = o_k - \delta \cdot \left(1 - p^* \otimes \frac{s}{p \otimes o} \right), \quad (\text{s7-20})$$

δ is the step-size. When $\delta = o_k$, Eq (s7-20) is equivalent to Eq. (s7-19). Since Eq. (s7-20) is standard gradient descent, we are able to extend it using more advanced optimizers such as momentum, or adaptive learning rate to update the parameters.

7-5.2 RL deconvolution in the presence of Gaussian process

When the noise signal is assumed to be Gaussian, Eq. (s7-15) is written as $s(x) = p(x) \otimes o(x) + \eta$ where η is the Gaussian noise. The likelihood follows Gaussian distribution, and Eq. (s7-17) now becomes:

$$o = \arg \min [-\log \mathcal{P}(o | s)] \propto \arg \min \mathcal{L}(o) \quad (s7-21)$$

where $\mathcal{L}(o) = \int_x \int_y (p \otimes o - s)^2 dx dy$.

The derivative of $\mathcal{L}(o)$ with respect to o is

$$\frac{\partial \mathcal{L}}{\partial o} = 2p^* \otimes (p \otimes o - s). \quad (s7-22)$$

Eq. (s7-22) is suitable for gradient descent-based methods.

7-5.3 RL deconvolution with Hessian regularization

In the above sections, we ignored the prior of the image o as $\mathcal{P}(o) = \text{constance}$. While in practice point of view, the $\mathcal{P}(o)$ highly depends on a certain type of input images and implementation purposes. For example, for denoising purposes, we can assume that the gradient of a noise-free o is less than a noise-corrupted one. Therefore, we incorporate this prior knowledge of images to the cost function as the penalty term to regularize the possible solution of o . For medical images, we use the Hessian gradient-the spatial second-order gradient, to penalize the cost function which is written as

$$-\log \mathcal{P}(o) = \|H \otimes o\| = \begin{cases} \int_x \int_y |H_{xx} \otimes o| + |H_{yy} \otimes o| + 2|H_{xy} \otimes o| dx dy & \text{for anisotropic} \\ \int_x \int_y \sqrt{(H_{xx} \otimes o)^2 + (H_{yy} \otimes o)^2 + 2(H_{xy} \otimes o)^2} dx dy & \text{for isotropic} \end{cases} \quad (s7-23)$$

where : $H = \nabla \nabla, H_{xx} = (1, -2, 1), H_{yy} = (1, -2, 1)^T, H_{xy} = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$

Taking the anisotropic Hessian regularization and Gaussian noise assumption as an example, submitting Eq. (s7-23) into Eq. (s7-21), yielding the Hessian regularized cost function:

$$\mathcal{L}(o) = \int_x \int_y (p \otimes o - s)^2 dx dy + \lambda \cdot \int_x \int_y |H_{xx} \otimes o| + |H_{yy} \otimes o| + 2|H_{xy} \otimes o| dx dy, \quad (s7-24)$$

where λ is the penalty parameter.

Then the derivative of $\mathcal{L}(o)$ with respect to o reads

$$\frac{\partial \mathcal{L}}{\partial o} = 2p^* \otimes (p \otimes o - s) + \lambda \left[H_{xx}^* \otimes \text{sign}(H_{xx} \otimes o) + H_{yy}^* \otimes \text{sign}(H_{yy} \otimes o) + 2H_{xy}^* \otimes \text{sign}(H_{xy} \otimes o) \right]. \quad (s7-25)$$

Here *sign* denotes the sign function. With the gradient, we can update parameter o based on the gradient descent method and its extensions.

7-5.4 RL deconvolution with illumination correction

We now introduce the illumination pattern i to Eq. (s7-15), and the new image formation model is written as

$$s(x, y) = p(x, y) \otimes [i(x, y) \cdot o(x, y)] + \eta, \quad (s7-26)$$

where we use the Gaussian process and η is the Gaussian noise. The cost function is given by

$$\mathcal{L}(o, i) = \int_x \int_y [p \otimes (i \cdot o) - s]^2 dx dy. \quad (s7-27)$$

The derivative of $\mathcal{L}(o)$ with respect to o can be calculated by chain rule which is

$$\frac{\partial \mathcal{L}(o, i)}{\partial o} = \frac{\partial \mathcal{L}(o, i)}{\partial (io)} \frac{\partial (io)}{\partial o} = 2i \cdot p^* \otimes [p \otimes (i \cdot o) - s], \quad (s7-28)$$

similarly, the derivative of $\mathcal{L}(o)$ with respect to i is given by

$$\frac{\partial \mathcal{L}(o, i)}{\partial i} = \frac{\partial \mathcal{L}(o, i)}{\partial(io)} \frac{\partial(io)}{\partial i} = 2o \cdot p^* \otimes [p \otimes (i \cdot o) - s]. \quad (\text{s7-29})$$

Eq. (s7-28) and Eq. (s7-29) form the first term for the Eq. (7-7) and Eq. (7-8) in the main text.

References

1. L. Xu, C. Lu, Y. Xu, and J. Jia, "Image smoothing via L_0 gradient minimization," in *Proceedings of the 2011 SIGGRAPH Asia conference*, 2011), 1-12.
2. J. Pan, Z. Hu, Z. Su, and M.-H. Yang, "Deblurring text images via L_0 -regularized intensity and gradient prior," in *CVPR*, (2014), pp. 2901–2908.
3. Y. Wang, J. Yang, W. Yin, and Y. Zhang, "A new alternating minimization algorithm for total variation image reconstruction," *SIAM Journal on Imaging Sciences* **1**, 248-272 (2008).
4. L. Huang, Y. Xia, and T. Ye, "Effective Blind Image Deblurring Using Matrix-Variable Optimization," *IEEE Trans. Image Process* **30**(2021).
5. A. Mitra, S. Roy, S. Roy, and S. K. Setua, "Enhancement and restoration of non-uniform illuminated Fundus Image of Retina obtained through thin layer of cataract," *Computer Methods and Programs in Biomedicine* **156**, 169-178 (2018).
6. K. Panetta, C. Gao, and S. Agaian, "Human-Visual-System-Inspired Underwater Image Quality Measures," *IEEE Journal of Oceanic Engineering* **41**, 1-11 (2016).
7. A. Rizzi, T. Algeri, G. Medeghini, and D. Marini, "A proposal for Contrast Measure in Digital Images," (2004).
8. W. Zhou, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing* **13**, 600-612 (2004).
9. Z. Kotevski and P. Mitrevski, "Experimental Comparison of PSNR and SSIM Metrics for Video Quality Estimation," in *ICT Innovations 2009*, (Springer Berlin Heidelberg, 2010), 357-366.
10. M. Sharif, L. Bauer, and M. K. Reiter, "On the Suitability of L_p -Norms for Creating and Preventing Adversarial Examples," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2018), 1686-16868.
11. "Ways of cheating on popular objective metrics: blurring, noise, super-resolution and others" (Video processing, compression and quality researc group, 2021), retrieved.
12. N. Dey, L. Blanc-Féraud, C. Zimmer, P. Roux, Z. Kam, J.-C. Olivo-Marin, and J. Zerubia, "3D Microscopy Deconvolution using Richardson-Lucy Algorithm with Total Variation Regularization," (INRIA, 2004).

Chapter 8

Geodesic Tracking of Retinal Vascular Trees with Optical and TV-Flow Enhancement in SE(2)

Nikcy van den Berg, Shuhe Zhang, Bart Smets, Tos T. J. M. Berendschot, and Remco Duits

International Conference on Scale Space and Variational Methods in Computer Vision (SSVM) 2023, DOI: 10.1007/978-3-031-31975-4_40

Abstract

Retinal images are often used to examine the vascular system in a non-invasive way. Studying the behavior of the vasculature on the retina allows for noninvasive diagnosis of several diseases as these vessels and their behavior are representative of the behavior of vessels throughout the human body. For early diagnosis and analysis of diseases, it is important to compare and analyze the complex vasculature in retinal images automatically. In previous work, PDE-based geometric tracking and PDE-based enhancements in the homogeneous space of positions and orientations have been studied and turned out to be useful when dealing with complex structures (crossing of blood vessels in particular). In this article, we propose a single new, more effective, Finsler function that integrates the strength of these two PDE-based approaches and additionally accounts for a number of optical effects (dehazing and illumination in particular). The results greatly improve both the previous left-invariant models and a recent data-driven model, when applied to real clinical and highly challenging images. Moreover, we show clear advantages of each module in our new single Finsler geometrical method.

Keywords

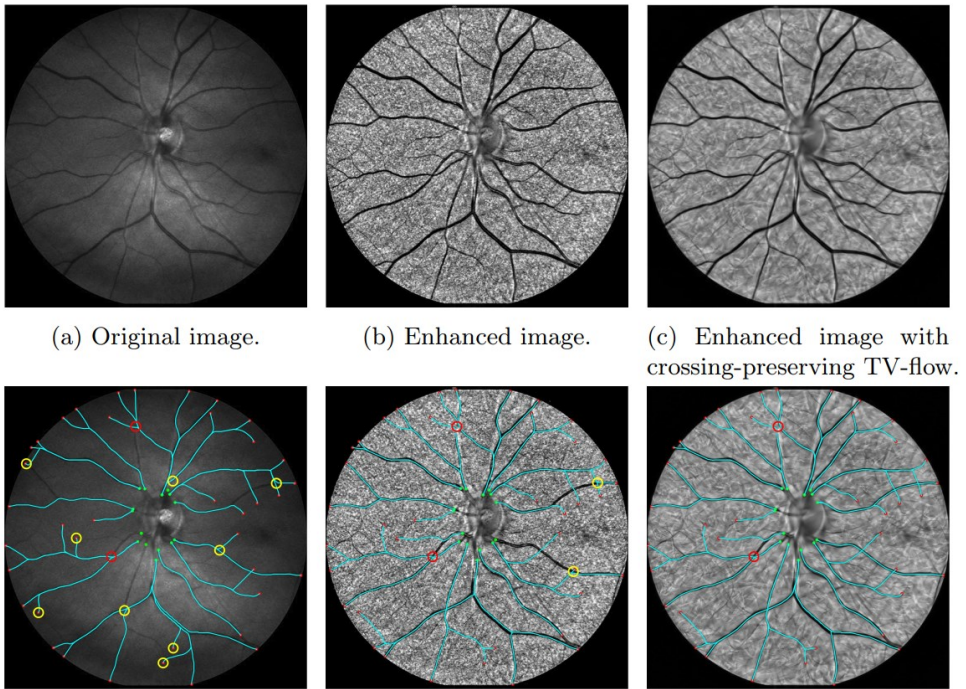
Geodesic Tracking; Optical Image Enhancement; TV-Flow Enhancement; Vascular Tree Tracking; Finsler Geometry.

1. Introduction

The retina allows for noninvasive examination of the vascular system since the vessels in the eye, and their corresponding behavior, are representative of the behavior of vessels throughout the rest of the body. Therefore, studying the behavior of the vasculature on the retina allows for noninvasive diagnosis of several diseases, like diabetes, hypertension, and Alzheimer's disease [1-3]. Automatic vessel tracking algorithms help the efficient diagnosis of these diseases. Here, we rely on geodesic tracking methods which calculate the shortest path connecting two points on the same blood vessel, following the biological structure. We will show that the single geodesic model will also allow for acceptable tracking of full vascular trees on realistic retinal images of limited quality.

Diseases such as cataract disease give rise to cloudy retinal images [4], while camera movements lead to motion artifacts [5] and uneven illumination [6]. This affects the clarity and visibility of the vasculature in the images we want to track. To cope with the limitations in the quality of ophthalmology images in practice, we must integrate both contrast enhancement from optical image processing [7] *and* crossing-preserving contextual TV-flows, in our correct geodesic tracking of the vasculature, as we will show.

Many approaches have been used when studying geodesic tracking methods on 2D images. However, in many methods, problems arise when dealing with difficult structures, like crossings and bifurcations, where standard geometric algorithms acting in the image domain \mathbb{R}^2 often take the wrong exit. Therefore, one lifts the image to the homogeneous space of positions and orientations \mathbb{M}_2 . In this lifted space, difficult structures are disentangled, cf. Fig. 8-2 (a). In previous works by various authors, it has been shown that PDE-based geometric tracking algorithms [8-11] perform well in \mathbb{M}_2 . Here, one first calculates a distance map which is based on the image data. After computing the distance map, the steepest descent algorithm is applied to find each shortest path from a tip (an endpoint) to the corresponding seed (a starting point).



(a) Original image.

(b) Enhanced image.

(c) Enhanced image with crossing-preserving TV-flow.

Fig. 8-1. Geodesic tracking on the original image, (contrast-)enhanced image, and enhanced image after which TV-flow enhancement is done (left to right). The seeds and tips are indicated in resp. green and red. Yellow (/red) circles indicate tracking mistakes that are (/are not) fixed in the tracking of another column.

In our tracking, we integrate PDE-enhancements, like crossing-preserving total variation flow (TV-flow) enhancement in \mathbb{M}_2 [12]. We will show this improves the results. Furthermore, optical enhancement [7] of limited-quality retinal images is required to keep equal contrast and intensity across the whole vasculature. This inevitably creates small noisy structures that are non-aligned with other structures in the data. Applying the TV-flow enhancement in \mathbb{M}_2 leads to crossing-preserving contextual denoising that preserves crossings, and line structures, and removes noisy non-aligned structures from the optical enhancement. Altogether the scheme results in a vascular tracking algorithm that provides better results as wavefronts follow the

complex branching vasculature better than the approach in [8], see Fig. 1. Even a single geodesic front propagation, where fronts follow the entire vasculature in one run produces good results, see Fig. 8-2.

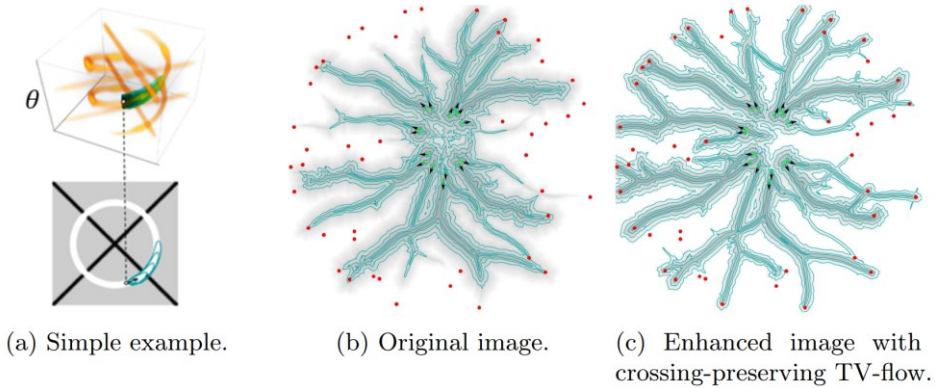


Fig. 8-2. Distance map built in \mathbb{M}_2 (2a; top) and iso-contours of the minimum projection back onto \mathbb{R}_2 (2a; bottom). The iso-contours of the minimum projection of the distance map are constructed based on the original image and enhanced image after which TV-flow enhancement is done (in Fig. 2b and 2c resp.).

The main contributions of this article are: 1) the development of a new asymmetric, data-driven left-invariant Finsler geometric model that includes contextual contrast enhancement via TV-flows on $SE(2)$; 2) experiments that show that application of this new Finsler geometric model reduces many tracking errors compared to previous left-invariant models [9, 13] and the recent data-driven model [8], 3) the new model performs very well on both realistic, unevenly illuminated retinal images and allows full vascular trees computations from a *single* distance map. The optical and TV-flow enhancements no longer require a 2-step approach as in [8].

2. Lifted space of positions and orientations \mathbb{M}_2

As explained in the introduction, it is beneficial to perform image processing (denoising, tracking) in the space of positions and orientations \mathbb{M}_2 . This 3D space will offer us sufficient space to separate difficult (crossing) structures.

Definition 1 (Space of positions and orientations \mathbb{M}_2). *The space of two-dimensional positions and orientations \mathbb{M}_2 is defined as a smooth manifold $\mathbb{M}_2 := \mathbb{R}^2 \rtimes S^1$, where $S^1 \equiv \mathbb{R}/(2\pi\mathbb{Z}) \equiv SO(2)$ using the identification*

$$\mathbf{n} = (\cos \theta, \sin \theta) \longleftrightarrow \theta \longleftrightarrow R_\theta \in SO(2), \quad (8-1)$$

where R_θ is the counter-clockwise planar rotation over angle θ . Elements in \mathbb{M}_2 are denoted by $(\mathbf{x}, \theta) \in \mathbb{R}^2 \times \mathbb{R}/(2\pi\mathbb{Z})$. To stress the semi-direct product of roto-translation group $SE(2) := \mathbb{R}^2 \rtimes SO(2)$ acting on \mathbb{M}_2 , we write $\mathbb{M}_2 := \mathbb{R}^2 \rtimes S^1$.

We lift the image from \mathbb{R}^2 to $\mathbb{M}_2 \equiv SE(2)$ in order to separate crossing structures, using the orientation score transform.

Definition 2 (Orientation Score). *The orientation score transform $W_\psi f : \mathbb{L}_2(\mathbb{R}^2) \rightarrow \mathbb{L}_2(\mathbb{M}_2)$ using anisotropic wavelet ψ maps an image $f \in \mathbb{L}_2(\mathbb{R}^2)$ to an orientation score $U = W_\psi f$ and is given by*

$$(W_\psi f)(\mathbf{x}, \theta) := \int_{\mathbb{R}^2} \overline{\psi[R_\theta^{-1}(\mathbf{y} - \mathbf{x})]} f(\mathbf{y}) d\mathbf{y}$$

In our experiments, we use cake wavelets [12, 14] as they do not temper data evidence.

In order to perform tracking in the lifted space of positions and orientations \mathbb{M}_2 , we need to introduce a metric that is used to describe distances. This metric needs to satisfy the property that a roto-translation of the input yields the same roto-translation on the output.

Definition 3 (Left-Invariant Metric). A metric tensor field \mathcal{G} on \mathbb{M}_2 is called left invariant if

$$\mathcal{G}_{g \cdot p} [(L_g)_* \dot{p}, (L_g)_* \dot{p}] = \mathcal{G}_p (\dot{p}, \dot{p}),$$

for all $p \in \mathbb{M}_2$, all $\dot{p} \in T_p(\mathbb{M}_2)$ and all $g \in SE(2)$, where $L_g(p) = g \cdot p = (y, R_\alpha) \cdot (x, \theta) = (y + R_\alpha x, \alpha + \theta)$ with push-forward $(L_g)_* \partial_x|_{p_0} U = \partial_x|_{p_0} (U \circ L_g)$.

More generally, a possibly asymmetric Finsler function defined on tangent-bundle $T(\mathbb{M}_2) = \{(p, \dot{p}) \mid \dot{p} \in T_p(\mathbb{M}_2)\}$ given by $\mathcal{F}: T(\mathbb{M}_2) \rightarrow \mathbb{R}^+$ is left invariant if $\mathcal{F}(p, \dot{p}) = \mathcal{F}[g \cdot p, (L_g)_* \dot{p}]$ for all $(p, \dot{p}) \in T(\mathbb{M}_2)$, $g \in SE(2)$.

Thereby the distance is left invariant:

$$d_{\mathcal{F}}(p_1, p_2) = \inf \left\{ \int_0^1 \mathcal{F}[\gamma(t), \dot{\gamma}(t)] dt \mid \gamma \in \Gamma_1, \gamma(0) = p_1, \gamma(1) = p_2 \right\} = d_{\mathcal{F}}(g \cdot p_1, g \cdot p_2),$$

for all $g \in SE(2)$, optimizing over the set Γ_1 of piecewise $C^1([0, 1], \mathbb{M}_2)$ -curves.

In our application, the asymmetric Finsler function will restrict backward movement as we will see in Sect. 3, and consequently cusps are avoided [9].

3. Existing Reeds-Shepp Car Models

Over the years, many geometric control problems have been proposed for geodesic tracking of blood vessels or vehicles. The ones closest to our model are the symmetric and asymmetric Reeds-Shepp car models. The symmetric Reeds-Shepp Car model, proposed in [15, 16], is given by

$$\mathcal{G}_p(\dot{p}, \dot{p}) = C(p)^2 \left(\xi^2 |\dot{x} \cdot n|^2 + \frac{\xi^2}{\zeta^2} \|\dot{x} \wedge n\|^2 + \|\dot{n}\|^2 \right), \tag{8-2}$$

for all $p = (x, n) \in \mathbb{M}_2$, $\dot{p} = (\dot{x}, \dot{n}) \in T_p(\mathbb{M}_2)$ with $\|\dot{x} \wedge n\|^2 := \|\dot{x}\|^2 - |\dot{x} \cdot n|^2$, where n is constructed using the identification in (1). The asymmetric Reeds-Shepp Car model, proposed in [9], is given by the asymmetric Finsler norm/function:

$$\|\mathcal{F}_p(\mathbf{p}, \dot{\mathbf{p}})\|^2 = \mathcal{G}_p(\dot{\mathbf{p}}, \dot{\mathbf{p}}) + C(\mathbf{p})^2 (\varepsilon^{-2} - 1) \xi^2 |(\dot{\mathbf{x}} \cdot \mathbf{n})_-|^2, \quad (8-3)$$

for all $\mathbf{p} = (\mathbf{x}, \mathbf{n}) \in \mathbb{M}_2$, $\dot{\mathbf{p}} = (\dot{\mathbf{x}}, \dot{\mathbf{n}}) \in T_p(\mathbb{M}_2)$ with $a_- := \min\{0, a\}$. The parameter ξ influences the flexibility of the tracking, weighing between spatial and angular movement. The anisotropy parameter ζ penalizes sideways movement. When $\zeta \downarrow 0$, the classical sub-Riemannian model appears.

The cost function $C: \mathbb{M}_2 \rightarrow [\delta, 1]$, $\delta > 0$ discourages movement outside the vascular structures via a crossing-preserving vesselness map $\mathcal{V}(\mathbf{p})$ [17],[[8], App.D].

Typically [[15], eq.5.1] one has $C(\mathbf{p}) = [1 + \lambda \mathcal{V}(\mathbf{p})^p]^{-1}$. The choice of the cost function is important and, in this article, (Sec. 4, 5, 6) we propose to include illumination enhancement and crossing preserving TV-flow (prior to the vesselness map computation) in the cost function as this will greatly improve tracking results.

In the asymmetric Reeds-Shepp car model, the parameter $\varepsilon \in (0, 1]$ describes how strongly the model has to adhere to the forward gear. When $\varepsilon = 1$, we are in the symmetric model. When $\varepsilon \downarrow 0$, backward motion becomes prohibited.

Computation of shortest paths (geodesics) connecting seeds and tips is done in 2 steps. First, the distances to all points in the domain are calculated, resulting in a distance map. Then, the shortest path is obtained by a steepest descent algorithm applied on this distance map. In all experiments we use the Anisotropic Fast Marching [15, 18] by J.-M. Mirebeau for distance map computations.

4. Methodology

4.1 Illumination Enhancement

Previous approaches in retinal vessel tracking typically consider the unprocessed picture S taken by the ophthalmologist, e.g. [9, 13, 17]. However, this may deviate from

the actual retinal image O which we aim to recover, due to possible cataract and uneven illumination.

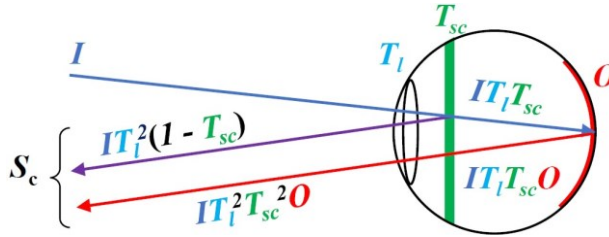


Fig. 8-3. Visualization of the physical model when imaging the retina; O is the actual image we would like to recover, S_c is the perceived image (sum of purple and red reflected light). I stands for input illumination, T_l and T_{sc} resp. denote the transmission ratio of the lens and the intraocular scattering (incl. cataract).

The physical model of the construction of the output image is visualized in Fig. 8-3. This yields the following standard optics formula:

$$S_c(\mathbf{x}) = L(\mathbf{x}) [T_{sc}^2(\mathbf{x})O_c(\mathbf{x}) + 1 - T_{sc}(\mathbf{x})] \quad \text{with } L = I \cdot T_l^2, \quad (8-4)$$

where c denotes the color channel in RGB or Y in YPbPr color space, and $L, T_{sc}: \Omega \rightarrow [0,1]$ denote the illumination from outside the eye and transmission of the intraocular scattering respectively on domain $\mathbf{x} \in \Omega \subset \mathbb{R}^2$. We apply an illumination correction, as done in [19]. After determining the illumination L , we re-express the Y channel (in YPbPr color space) of Eq. (8-4) to

$$O_Y = [L^{-1}(\mathbf{x})S_Y(\mathbf{x}) - 1 + T_{sc}(\mathbf{x})] / T_{sc}^2(\mathbf{x}), \quad O_Y \in [0,1] \quad \forall \mathbf{x} \in \Omega$$

$$\text{with } T_{sc} = 1 - A \left\{ L^{-1}(\cdot)S(\cdot) + \frac{1}{n} \sum_{l=1}^n \frac{1}{1 + \exp(-\varphi(\cdot))} G_{\sigma_l} \otimes [L^{-1}(\cdot)S(\cdot)] \right\}, \quad (8-5)$$

with Gaussian kernel standard deviations $\sigma_l = \text{pixelsize} \cdot 2^{l-1}$ of the retinal image at scale level $l \in \{1, \dots, n\}$ where we took $n=4$ and where the sigmoids on scale coefficients above are included to control the range in $[0,1]$ and to allow for stable optimization below. The Y channel of the actual image O is obtained by solving the Euler

Lagrange equation of the Tikhonov regularization problem via the Karush-Kuhn-Tucker conditions including the constraints $O_Y \in [0,1]$:

$$(\varphi^{\min}, A^{\min}) = \arg \min_{(\varphi, A)} \left\{ \|O_Y - \bar{O}_Y\|_{\mathbb{L}_2(\Omega)}^2 + \lambda \|\nabla O_Y\|_{\mathbb{L}_2(\Omega)}^2 \right\}, \quad (8-6)$$

w. r. t. $A > 0$ and $\varphi = (\varphi_i)_{i=1}^n \in \mathbb{R}^n$. Here, $O_Y \in \mathbb{R}$ is an estimation of the desired intensity level [19] in Sec. 3, 4.2, and λ regulates the smoothness of O_Y . After optimal non-constant $O_Y = O_Y(\varphi^{\min}, A^{\min})$ is retrieved by (8-5), image O follows by linear conversion of YPbPr- to RGB-colors, via updated Pb- and Pr-channels.

4.2 TV-flow Enhancement

TV-flow enhancement is a valuable technique to denoise surfaces, but at the same time preserve sharp edges. Recall that the metric intrinsic gradient is given by

$$\mathbf{p} \in \mathbb{M}_2, \mathbf{p} \mapsto \nabla_{\mathcal{G}} \varphi(\mathbf{p}) = (\mathcal{G}^{-1} d\varphi)(\mathbf{p}) \in T_{\mathbf{p}}(\mathbb{M}_2),$$

using \mathcal{G} in (2) with $\zeta = C\xi = 1$, $C = 10$. Then TV-flow $U \mapsto W_0(\cdot, t)$ is given by

$$\begin{cases} \partial_t W_\varepsilon(\mathbf{p}, t) = \operatorname{div} \left\{ \frac{\nabla_{\mathcal{G}} W_\varepsilon(\cdot, t)}{\varepsilon^2 + [\nabla_{\mathcal{G}} W_\varepsilon(\cdot, t)]^2} \right\}(\mathbf{p}), & \mathbf{p} \in \mathbb{M}_2, \quad t \geq 0 \\ W_\varepsilon(\mathbf{p}, 0) = U(\mathbf{p}) \end{cases}$$

And $W_0(\mathbf{p}, t) = \lim_{\varepsilon \downarrow 0} W_\varepsilon(\mathbf{p}, t)$. For proof of the \mathbb{L}_2 -convergence, see [12]. Training of the end-time t of the TV-flow is not needed as for all lifted optically enhanced (cf. Sec. 4) images U of the STAR-dataset [20], end-time $t = 0.5$ is nearly optimal for subsequent tracking, and $\Delta t = 0.1$ always remains in the stability region [12]. The same settings provided optimal PSNR-ratios in [12].

4.3 A Finsler Metric on \mathbb{M}_2 that Includes the Enhancements

Our goal was to track vasculature accurately. In order to do so, one needs a metric tensor field that describes distances on the manifold. In some cases, it is beneficial to construct a metric tensor field \mathcal{G}^U that depends explicitly on the underlying orientation score data U . This “data-driven” metric tensor field needs to be left invariant with respect to the roto-translation of the underlying data:

Definition 4 (Data-Driven Left-Invariant Metric (DDLIM)). *The metric tensor fields \mathcal{G}^U and \mathcal{F}^U on \mathbb{M}_2 are data-driven left invariant when they satisfy for all $(\mathbf{p}, \dot{\mathbf{p}}) \in T(\mathbb{M}_2)$ and all $g \in SE(2)$:*

$$\mathcal{G}_p^U(\dot{\mathbf{p}}, \dot{\mathbf{p}}) = \mathcal{G}_{g \cdot p}^{\mathcal{L}_g^U}[(L_g)_* \dot{\mathbf{p}}, (L_g)_* \dot{\mathbf{p}}], \text{ and } \mathcal{F}^U(\mathbf{p}, \dot{\mathbf{p}}) = \mathcal{F}^{\mathcal{L}_g^U}[g \cdot \mathbf{p}, (L_g)_* \dot{\mathbf{p}}], \quad (8-7)$$

where $\mathcal{L}_g U(\mathbf{h}) := U(L_{g^{-1}} \mathbf{h}) := U(g^{-1} \cdot \mathbf{h}), \forall \mathbf{h} \in \mathbb{M}_2$.

The considered data-driven left-invariant metric tensor fields are given by

$$\mathcal{G}_p^U(\dot{\mathbf{p}}, \dot{\mathbf{p}}) = \mathcal{G}_p(\dot{\mathbf{p}}, \dot{\mathbf{p}}) + \mu C^2(\mathbf{p}) \frac{\|HU|_p(\dot{\mathbf{p}}, \cdot)\|_*^2}{\max_{\|\dot{\mathbf{q}}\|=1} \|HU|_p(\dot{\mathbf{q}}, \cdot)\|_*^2}, \quad (8-8)$$

$$|\mathcal{F}^U(\mathbf{p}, \dot{\mathbf{p}})|^2 = |\mathcal{F}(\mathbf{p}, \dot{\mathbf{p}})|^2 + \mu C^2(\mathbf{p}) \frac{\|HU|_p(\dot{\mathbf{p}}, \cdot)\|_*^2}{\max_{\|\dot{\mathbf{q}}\|=1} \|HU|_p(\dot{\mathbf{q}}, \cdot)\|_*^2}, \quad (8-9)$$

for all $\mathbf{p} = (\mathbf{x}, \mathbf{n}), \dot{\mathbf{p}} = (\dot{\mathbf{x}}, \dot{\mathbf{n}})$, representing the symmetric and asymmetric metric tensor fields respectively. In Eq. (8-8) and Eq. (8-9), the metric tensor fields \mathcal{G} and \mathcal{F} were introduced in (2) and (3) respectively. The Hessian field HU is defined as $HU := \nabla(dU)$, w. r. t. plus Cartan connection $\nabla^{[+]}$ for computational details see [21], [[8], Rem.8], and $\|\cdot\|_*$ denotes the dual norm w. r. t. $\sqrt{\mathcal{G}_p(\dot{\mathbf{p}}, \dot{\mathbf{p}})}$, where $\zeta = \xi = C = 1$.

The parameter $\mu > 0$ regulates the inclusion of the new data-driven term, and $C(\mathbf{p})$ denotes the cost function described in [[8], App. D].

Remark 1. The construction of this cost now relies on the orientation score U of the optically enhanced image with TV-flow enhancement, whereas previously it relied on the orientation score of the unprocessed image. Akin to [4, App.C], one can show that the new Finsler/Riemannian metric tensor fields (8-8) are DDLIM.

The geodesics are calculated by steepest descent on distance maps using a metric that describes distances in \mathbb{M}_2 .

Definition 5 (Data-Driven Riemannian Distance). *The data-driven Riemannian distance $d_{\mathcal{G}^U}$ from a point $\mathbf{p} \in \mathbb{M}_2$ to a point $\mathbf{q} \in \mathbb{M}_2$ is given by*

$$d_{\mathcal{G}^U}(\mathbf{p}, \mathbf{q}) = \inf_{\substack{\gamma \in \Gamma_1, \\ \gamma(0) = \mathbf{p}, \\ \gamma(1) = \mathbf{q}}} \int_0^1 \sqrt{\mathcal{G}_{\gamma(t)}^U[\dot{\gamma}(t), \dot{\gamma}(t)]} dt, \quad (8-10)$$

where $\Gamma_1 := \{\gamma : [0,1] \rightarrow \mathbb{M}_2 \mid \gamma \in PC^1([0,1], \mathbb{M}_2)\}$ with PC^1 the space of piecewise continuously differentiable curves in \mathbb{M}_2 , and $\dot{\gamma}(t) := d\gamma(t)/dt$. The quasi-distance that belongs to the asymmetric Finslerian model (8-9) is given by

$$d_{\mathcal{F}^U}(\mathbf{p}, \mathbf{q}) = \inf_{\substack{\gamma \in \Gamma_1, \\ \gamma(0) = \mathbf{p}, \\ \gamma(1) = \mathbf{q}}} \mathfrak{L}_{\mathcal{F}^U}(\gamma) = \inf_{\substack{\gamma \in \Gamma_1, \\ \gamma(0) = \mathbf{p}, \\ \gamma(1) = \mathbf{q}}} \int_0^1 \mathcal{F}^U[\gamma(t), \dot{\gamma}(t)] dt. \quad (8-11)$$

Lemma 1. If \mathcal{F}^U is DDLIM (Def. 4) then distance $d_{\mathcal{F}^U}$ satisfies:

$$\forall_{g \in SE(2)} \forall_{\mathbf{p}_1, \mathbf{p}_2 \in \mathbb{M}_2} : d_{\mathcal{F}^U \circ g}(\mathbf{g} \cdot \mathbf{p}_1, \mathbf{g} \cdot \mathbf{p}_2) = d_{\mathcal{F}^U}(\mathbf{p}_1, \mathbf{p}_2). \quad (12)$$

Proof. One has $d_{\mathcal{F}^U \circ g}(\mathbf{g} \cdot \mathbf{p}_1, \mathbf{g} \cdot \mathbf{p}_2) = \inf_{\substack{\gamma \in \Gamma_1, \\ \gamma(0) = \mathbf{g} \cdot \mathbf{p}_1, \\ \gamma(1) = \mathbf{g} \cdot \mathbf{p}_2}} \mathfrak{L}_{\mathcal{F}^U \circ g}(\gamma) \stackrel{(7)}{=} \inf_{\substack{g^{-1} \cdot \gamma \in \Gamma_1, \\ g^{-1} \cdot \gamma(0) = \mathbf{p}_1, \\ g^{-1} \cdot \gamma(1) = \mathbf{p}_2}} \mathfrak{L}_{\mathcal{F}^U}(g^{-1} \cdot \gamma) =$

$d_{\mathcal{F}^U}(\mathbf{p}_1, \mathbf{p}_2)$, where $g^{-1} \cdot \gamma \in \Gamma_1 \Leftrightarrow \gamma \in \Gamma_1$, from which (12) follows.

The shortest curves are computed using the steepest descent on the distance map, departing from tip $\mathbf{p} \in \mathbb{M}_2$ towards seed $\mathbf{p}_0 \in \mathbb{M}_2$ as described in **Theorem 1**.

Theorem 1. *The shortest curve $\gamma : [0, 1] \rightarrow \mathbb{M}_2$ with $\gamma(0) = \mathbf{p}$ and $\gamma(1) = \mathbf{p}_0$ can be computed by steepest descent tracking on distance map $W(\mathbf{p}) = d_{\mathcal{F}^U}(\mathbf{p}, \mathbf{p}_0)$*

$$\gamma(t) := \gamma_{\mathbf{p}, \mathbf{p}_0}^U(t) = \text{Exp}_{\mathbf{p}}[t\nu(W)], \quad t \in [0, 1], \quad (8-13)$$

where Exp integrates the following vector field on $\mathbb{M}_2 : \nu(W) := -W(\mathbf{p})\nabla_{\mathcal{F}^U}W$ and where W is the viscosity solution of the eikonal PDE system

$$\begin{cases} \mathcal{F}_U^*[\mathbf{p}, dW(\mathbf{p})] = 1 & \mathbf{p} \in \mathbb{M}_2, \\ W(\mathbf{p}_0) = 0, \end{cases} \quad (8-14)$$

assuming \mathbf{p} is neither a 1st Maxwell-point nor a conjugate point, with dual Finsler function $\mathcal{F}_U^*(\mathbf{p}, \hat{\mathbf{p}}) := \max\{\langle \hat{\mathbf{p}}, \dot{\mathbf{p}} \rangle \mid \dot{\mathbf{p}} \in T_{\mathbf{p}}(\mathbb{M}_2), \text{ with } \mathcal{F}^U(\mathbf{p}, \dot{\mathbf{p}}) \leq 1\}$. As $\nu(W)$ is data-driven left invariant, the geodesics carry the symmetry

$$\gamma_{g \cdot \mathbf{p}, g \cdot \mathbf{p}_0}^{\mathcal{F}_U^*}(t) = g \gamma_{\mathbf{p}, \mathbf{p}_0}^U(t) \text{ for all } g \in SE(2), \mathbf{p}, \mathbf{p}_0 \in \mathbb{M}_2, t \in [0, 1]. \quad (8-15)$$

Proof. This is a special case of [[8], Thm.1] with Lie group $SE(2) \cong \mathbb{M}_2$. Then this yields the symmetric case $\|\nabla_{\mathcal{G}^U}W(\mathbf{p})\| = 1$ in the usual eikonal PDE form. Inclusion of the asymmetric front propagation (relying on asymmetric Finsler metric \mathcal{F}^U) requires a replacement of $\|\nabla_{\mathcal{G}^U}W(\mathbf{p})\| = 1$ with a dual norm expression $\mathcal{F}_U^*[\mathbf{p}, dW(\mathbf{p})] = 1$, where one takes the positive part of the spatial momentum component in direction $\cos\theta dx + \cos\theta dy \in T^*(\mathbb{M}_2)$. This is similar to the technique in [[9], Thm.4] but due to the data-driven behavior \mathcal{F}^U this is subtle [[8], Eq.43, Lem. 3] and also directly applies

to our model (8-16) using cost function C (incl. optical & TF-flow enhancement) as explained in Remark 1. Also, the backtracking requires a subtle adaptation: instead of ordinary intrinsic gradient descent in direction $\nabla_{\mathcal{G}^U} W = (\mathcal{G}^U)^{-1} dW$ it now becomes more general descent in direction $\nabla_{\mathcal{F}^U} W(\cdot) := d\mathcal{F}_U^*[\cdot, dW(\cdot)]$ as explained in [[9], prop.4].

In the experimental section, we rely on the mixed metric tensor field, which is needed to avoid wrong exits at complex structures, see [8], and is given by:

$$\begin{aligned} \mathcal{G}_p^M(\dot{p}, \dot{p}) &= \kappa(x) \mathcal{G}_p(\dot{p}, \dot{p}) + [1 - \kappa(x)] \mathcal{G}_p^U(\dot{p}, \dot{p}), \quad p = (x, n) \in \mathbb{M}_2 \\ \mathcal{F}^M(p, \dot{p})^2 &= \kappa(x) \mathcal{F}(p, \dot{p})^2 + [1 - \kappa(x)] \mathcal{F}^U(p, \dot{p})^2, \end{aligned} \tag{8-16}$$

with $\kappa = \mathbb{1}_A \otimes G_\sigma$ and A the crossing structure locations and Gaussian $G_{\sigma=1 \text{ pix}}$.

5. Experimental Results

We rely on the asymmetric metric tensor field (16) to calculate the geodesics of the 3 different models. These models are constructed based on a) the picture taken by the ophthalmologist (original image), b) the original image with illumination enhancement as explained in Sec. 4.1 (enhanced image), and c) the enhanced image with crossing-preserving TV-flow enhancement discussed in Sec. 4.2. In this section, we illustrate the results for a specific image and refer to Tab. 8-1 for an overview of the performance of the different models on the STAR dataset [1,19]. These results are consistent with the discussed example, and for reproducible code and all processed images see github.com/anonymous. In all experiments we set standard parameter settings [[15], eq.5.1], [[8], eq.65] for cost-function $C(p=3, \lambda=1000)$, for the metrics ($\xi = \zeta = 0.1$) in (2), for TV-flow ($t=0.5$).

In prior research, the “original image” directly entered the metric tensor field on M_2 when calculating the geodesics. These images are of varying quality, depending on the patient’s condition and the used equipment. Applying a tracking algorithm, like

Anisotropic Fast Marching [8], on the metric tensor field based on the original image, often results in tracking mistakes due to uneven illumination, both along vascular structures and on the background, making it hard to distinguish different structures. In Fig. 8-1, one sees the tracking results on an original, unevenly illuminated, non-enhanced image, where all vessels were tracked in one single run. At a lot of locations (13), the tracking connects the seeds p_0 and tips p incorrectly.

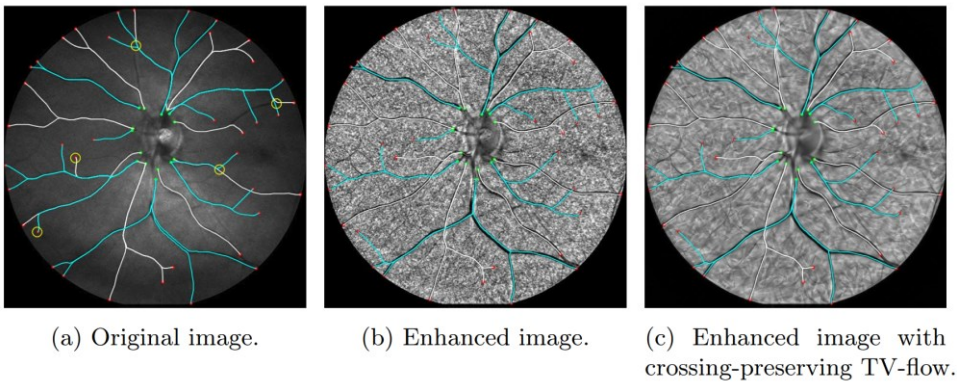


Fig. 8-4. Tracking of Vascular Tree per Vessel Type: Tracking with the mixed model (M_2, \mathcal{F}^M) proposed in Eq. (7-16) with $\mu = 15$. Prior classification of vascular trees by type (artery/vein resp. white/cyan) only results in perfect tracking of the vessel tree on the enhanced images. Yellow circles indicate tracking mistakes.

The optical enhancement explained in Sec. 4.1 corrects for uneven illumination. Calculating the geodesics using the metric tensor field relying upon the enhanced image, reduces the number of tracking mistakes significantly (to 5), cf. Fig. 8-1 (b). Due to pointwise optimization in the optical enhancement, noise is generated. The crossing-preserving total variation flow enhancement suppresses this noise and indeed results in even fewer tracking mistakes (3), cf. Fig. 8-1 (c).

Calculating the tracking results in Fig. 8-1 uses no knowledge about the vasculature, apart from seed and tip locations. One might incorporate prior knowledge A) on vessel types (artery/vein), or B) on the connectivity of tips and seeds.

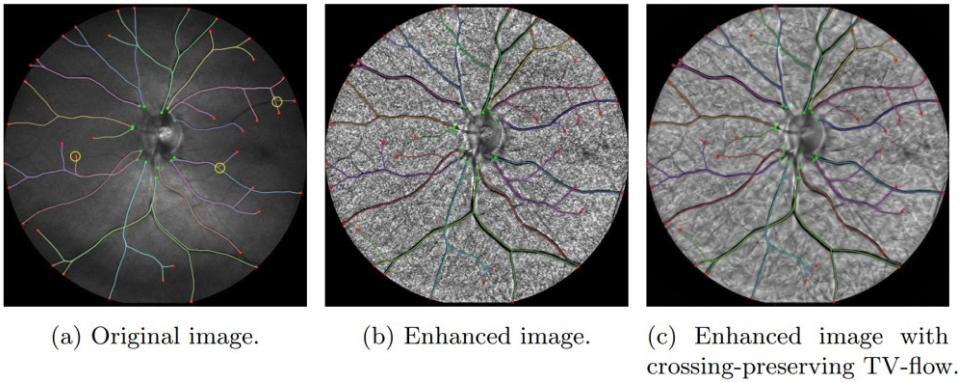


Fig. 8-5. Tracking of Vascular Tree per Seed on the Optic Disk: Tracking with the mixed model (M_2, \mathcal{F}^M) proposed in Eq. (16) with $\mu = 15$. Prior grouping of tips (in red) and seeds (in green) only results in perfect tracking of the vessel tree on the enhanced images. Tracking mistakes are indicated by yellow circles.

We start by investigating prior knowledge on A), where we first connect all tips on arteries to the seeds on arteries, and similarly for the tips and seeds on veins. Fig. 8-4 shows that the tracking results improve significantly for all cases; the number of tracking mistakes at crossings reduces from (13, 5, 3) to (5, 0, 0) for resp. the original, enhanced and enhanced with TV-flow image.

Second, we investigate the prior knowledge on B). In Fig. 8-5, the tracking results connecting the tips to their corresponding seed are presented. The number of correct tracks has improved once again, to only (3, 0, 0) mistakes for resp. the original, enhanced and enhanced with TV-flow image.

We report the tracking results for the three different approaches on images from STAR [20, 22], in particular the example in Fig. 8-1, Fig. 8-4, and Fig. 8-5. We observe the same trend in performance for other images which we summarize in Tab. 7-1. We calculate the (weighted) percentage of incorrectly calculated geodesics by means of:

$$\varepsilon := \frac{1}{|T|} \sum_{\mathbf{y} \in T} \left(1 - \frac{\|\mathbf{y} - \mathbf{x}_0(\mathbf{y})\|}{\sqrt{N_x^2 + N_y^2}} \right) C_0[\mathbf{x}_0(\mathbf{y}), \mathbf{y}] \geq 0, \tag{7-17}$$

Here S, T denote the sets of resp. seeds (near the optic disk) and tips. The image size is $N_x \times N_y$. The ground truth seed and calculated seed (first arriving front in the distance map, cf. Fig. 8-2) corresponding to the tip \mathbf{y} are resp. denoted by $\mathbf{x}_0(\mathbf{y}) \in S$ and $\mathbf{x}(\mathbf{y}) \in S$. Function $C_0 : S \times T \rightarrow \{0,1\}$ is given by $C_0(\mathbf{x}, \mathbf{y}) = 0$ if the tracking between \mathbf{x} and \mathbf{y} is correct and $C_0(\mathbf{x}, \mathbf{y}) = 1$ otherwise.

Tab. 7-1. Error measure ε of each tracking applied to STAR images in [23], calculated by Eq. (7-17). Green: best results per tracking.

	Original image	Enhanced image	Enhanced image with crossing-preserving TV-flow
Single Run	0.34	0.23	0.20
Per Type (A/V)	0.25	0.12	0.10
Per Seed	0.23	0.09	0.09

We evaluate with a harsh error measure Eq. (7-17): one crossing mistake (indicated by a circle) often causes more errors, when the vessel bifurcates after the crossing.

The illumination enhancement and TV-flow regularization applied on the original images, result in more accurate geodesics compared to those calculated directly on the original images, as can be seen in Tab. 7-1. The more prior information we use, the more accurate the geodesics follow the vasculature. Remarkably, tracking requiring artery-vein classification of seeds and tips performs similarly and is easier to automate than tracking with knowledge of seed-tip connectivity.

6. Conclusion

We developed a new asymmetric, data-driven left-invariant Finsler geometric model that includes contextual contrast enhancement via TV-flows on $SE(2)$. Experiments reveal that application of this new Finsler geometric model has benefits over previous left-invariant models [9, 13] and the recent data-driven model [8]. The new model reduces many errors and performs very well on both realistic and challenging low-quality retinal images where full vascular trees are computed from a *single* asymmetric Finslerian distance map. Although we have shown that both the contrast enhancement and the TV-flow on $SE(2)$ in the new Finslerian model are highly beneficial, there are still exceptional cases where vessel tracts take the wrong exit. This happens at places where both a crossing and a bifurcation occur [cf. the red circles in Fig. 8-1 (c)]. Therefore, in future work, we aim to tackle these cases by automatic artery vein classification via PDE-G-CNNs [24], as our experiments show this allows us to obtain the same good practical results as with the ‘tracking per seed’ (that requires too costly user-knowledge).

Acknowledgements. We gratefully acknowledge the Dutch Foundation of Science NWO for its financial support by Talent Programme VICI 2020 Exact Sciences (Duits, Geometric learning for Image Analysis, VI.C 202-031).

References

1. P. Colligris, M. J. Perez de Lara, B. Colligris, and J. Pintor, "Ocular manifestations of Alzheimer's and other neurodegenerative diseases: the prospect of the eye as a tool for the early diagnosis of Alzheimer's disease," *Journal of ophthalmology* **2018** (2018).
2. M. B. Sasongko, T. Y. Wong, T. T. Nguyen, C. Y. Cheung, J. E. Shaw, R. Kawasaki, E. L. Lamoureux, and J. J. Wang, "Retinal vessel tortuosity and its relation to traditional and novel vascular risk markers in persons with diabetes," *Current eye research* **41**, 551-557 (2016).
3. D. L. Weiler, C. B. Engelke, A. L. Moore, and W. W. Harrison, "Arteriole tortuosity associated with diabetic retinopathy and cholesterol," *Optometry and Vision Science* **92**, 384-391 (2015).
4. L. Xiong, H. Li, and L. Xu, "An approach to evaluate blurriness in retinal images with vitreous opacity for cataract diagnosis," *Journal of healthcare engineering* **2017** (2017).
5. A. Mitra, S. Roy, S. Roy, and S. K. Setua, "Enhancement and restoration of non-uniform illuminated fundus image of retina obtained through thin layer of cataract," *Computer methods and programs in biomedicine* **156**, 169-178 (2018).
6. M. Zhou, K. Jin, S. Wang, J. Ye, and D. Qian, "Color retinal image enhancement based on luminosity and contrast adjustment," *IEEE Transactions on Biomedical engineering* **65**, 521-527 (2017).
7. S. Zhang, C. A. Webers, and T. T. Berendschot, "A double-pass fundus reflection model for efficient single retinal image enhancement," *Signal Processing* **192**, 108400 (2022).
8. N. v. d. Berg, "Data-driven left-invariant tracking on optically enhanced images with total variation flow in Mathematica," https://github.com/NickyvdBerg/tracking_enhancement_total_variation_flow.
9. R. Duits, B. Smets, A. Wemmenhove, J. Portegies, and E. Bekkers, "Recent Geometric Flows in Multi-orientation Image Processing via a Cartan Connection," *Handbook of Mathematical Models and Algorithms in Computer Vision and Imaging: Mathematical Imaging and Vision*, 1-60 (2021).
10. B. Franceschiello, A. Mashtakov, G. Citti, and A. Sarti, "Geometrical optical illusion via sub-Riemannian geodesics in the roto-translation group," *Differential Geometry and its Applications* **65**, 55-77 (2019).
11. L. Liu, M. Wang, S. Zhou, M. Shu, L. D. Cohen, and D. Chen, "Curvilinear Structure Tracking Based on Dynamic Curvature-penalized Geodesics," *Pattern Recognition* **134**, 109079 (2023).
12. B. M. Smets, J. Portegies, E. St-Onge, and R. Duits, "Total variation and mean curvature PDEs on the homogeneous space of positions and orientations," *Journal of Mathematical Imaging and Vision* **63**, 237-262 (2021).
13. E. Bekkers, R. Duits, T. Berendschot, and B. ter Haar Romeny, "A multi-orientation analysis approach to retinal vessel tracking," *Journal of Mathematical Imaging and Vision* **49**, 583-610 (2014).
14. R. Duits, S. P. Meesters, J.-M. Mirebeau, and J. M. Portegies, "Optimal paths for variants of the 2D and 3D Reeds–Shepp car with applications in image analysis," *Journal of Mathematical Imaging and Vision* **60**, 816-848 (2018).
15. E. J. Bekkers, R. Duits, A. Mashtakov, and G. R. Sanguinetti, "A PDE approach to data-driven sub-Riemannian geodesics in SE (2)," *SIAM Journal on Imaging Sciences* **8**, 2740-2770 (2015).
16. J. Reeds, and L. Shepp, "Optimal paths for a car that goes both forwards and backwards," *Pacific journal of mathematics* **145**, 367-393 (1990).
17. J. Hannink, R. Duits, and E. Bekkers, "Crossing-preserving multi-scale vesselness," in *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2014: 17th International Conference, Boston, MA, USA, September 14-18, 2014, Proceedings, Part II 17*(Springer2014), pp. 603-610.

18. J.-M. Mirebeau, "Fast-marching methods for curvature penalized shortest paths," *Journal of Mathematical Imaging and Vision* **60**, 784-815 (2018).
19. S. Zhang, C. A. Webers, and T. T. Berendschot, "Luminosity rectified blind Richardson-Lucy deconvolution for single retinal image restoration," *Computer Methods and Programs in Biomedicine* **229**, 107297 (2023).
20. J. Zhang, B. Dashtbozorg, E. Bekkers, J. P. Pluim, R. Duits, and B. M. ter Haar Romeny, "Robust retinal vessel segmentation via locally adaptive derivative frames in orientation scores," *IEEE transactions on medical imaging* **35**, 2631-2644 (2016).
21. R. Duits, "Perceptual organization in image analysis : a mathematical approach based on scale, orientation and curvature," (Technische Universiteit Eindhoven, 2005).
22. S. Abbasi-Sureshjani, I. Smit-Ockeloen, J. Zhang, and B. Ter Haar Romeny, "Biologically-Inspired Supervised Vasculature Segmentation in SLO Retinal Fundus Images," in *Image Analysis and Recognition*, M. Kamel, and A. Campilho, eds. (Springer International Publishing, Cham, 2015), pp. 325-334.
23. N. v. d. Berg, B. Smets, G. Pai, J.-M. Mirebeau, and R. Duits, "Geodesic Tracking via New Data-driven Connections of Cartan Type for Vascular Tree Tracking," *arXiv preprint arXiv:2208.11004* (2022).
24. B. M. Smets, J. Portegies, E. J. Bekkers, and R. Duits, "PDE-based group equivariant convolutional neural networks," *Journal of Mathematical Imaging and Vision* **65**, 209-239 (2023).

Chapter 9

General discussions and concluding remarks

1. General discussions

Retinal images contain rich information such as retinal structure information (blood vessels, ocular disk, nerve bundles), and tissue color information, which are deeply associated with eye-related diseases and systemic diseases [1]. Image quality is related to diagnosis accuracy and impacts applications such as computer-aided diagnosis. Increasing retinal image quality using high-end fundus cameras is an option, however, it lacks generalization, and not all eye clinics can afford such expensive equipment. Therefore, digital retinal image restoration has attracted interest as it may provide a simple and cheap way to increase the image's quality.

In this thesis, we propose to use non-deep-learning-based methods to achieve single retinal image restoration tasks. These tasks are categorized into (1) retinal image illumination correction; (2) cataractous retinal image dehazing; (3) retinal image blind-deconvolution. We designed comprehensible algorithms to bring the degraded retinal image back to high-visual quality in which each step can be understood by mathematicians and ophthalmologists. Therefore, the safety of these algorithms can be evaluated not only by their final outputs but also by their intermediate products, while the deep-learning method can't.

Since we use convolution in our algorithms, we first, in Chapter 2, developed a reflective background padding method to pad the black background in every retinal image. This background padding will reduce and even eliminate the boundary effects of the convolution process. Then in **Chapter 3**, we developed a double-pass fundus reflection (DPFR) model based on the optical pathways in a fundus camera, to describe how a retinal image may be degraded step-by-step and proposed a simple restoration algorithm to validate our proposed DPFR model. Further in **Chapter 4** and **Chapter 5**, we present two improvements on the DPFR model, one for cataract retinal images, and the other one for color-naturalness preservation retinal image restoration. In **Chapter 6**, and **Chapter 7** we investigated the possibility of combining image illumination

correction and image blind deconvolution. Finally, **Chapter 8** explores how retinal image restoration can benefit blood vessel segmentation.

1.1 On Chapter 3

The main focus of **Chapter 3** is establishing a mathematical model to describe the formation of retinal images. The image formation model (IFM) based retinal image restorations share a similar idea of computational imaging, i.e. a physical model built to describe the optical process of forming an image that is affected by optical aberrations, unstable vibrations, or a limited optical resolution. By directly or indirectly measuring the optical properties of these degeneration agents, one can compensate for the degeneration agents by digitally mimicking the propagation of the optical wave and modifying the wavefront of light [2-5].

We noticed that in previous publications the IFMs were not designed nor optimized for retinal image restorations. We, therefore, developed the DPFR model, which reveals the specific double pass fundus reflection feature that was hitherto neglected in modeling the light propagation of fundus imaging. The DPFR model considers how the illumination light enters the ocular system, interacts with different layers (say intraocular scattering layers (cataracts)), and is reflected by the retina. We modeled the transmission of light through a certain layer as the product between the incident light matrix and a transmission matrix. To validate our DPFR, we used off-the-shelf image enhancement methods, viz Retinex, and dark-channel prior to achieve retinal image contrast enhancement. Our experimental results on retinal images show the DPFR's promising ability and potentiality on single retinal image restoration!

Although we were quite happy with the outcomes of the DPFR model, we liked to further specialize the DPFR model on cataractous retinal image dehazing, motivated by the fact that some patients with cataracts can also have macular degeneration. In the latter case, their visual ability cannot recover to a normal level even if cataract surgery is performed. If ophthalmologists were able to “see through” the cataract layers to find out whether cataract patients have macular degeneration, they could make better and

more accurate decisions on the necessity of cataract surgery. This idea contributed to our development of MUTE, a multilevel-stimulated denoising strategy for single cataractous retinal image dehazing in Chapter 4.

1.2 On Chapter 4

Chapter 4 discusses the MUTE algorithm, developed to enhance the contrast and reveal the hidden structures in cataractous retinal images. We studied the duality between cataractous retinal image dehazing and image denoising and found that the dehazing task for cataractous retinal images can be achieved with a combination of image denoising and a sigmoid function. To do so, we introduced the double-pass fundus reflection model in the YPbPr color space and developed a multilevel stimulated denoising strategy termed MUTE.

The transmission matrix of the cataract layer was expressed as the superposition of denoised raw images of different levels weighted by pixel-wise sigmoid functions. We further designed an intensity-based cost function that can guide the updating of the model parameters. They are updated by gradient descent with adaptive momentum estimation, which gives us the final refined transmission matrix of the cataract layer.

We tested our methods on cataract retinal images from both public and proprietary databases and compared the performance of our method with other state-of-the-art enhancement methods including deep-learning-based methods. Both visual assessments and objective assessments show the superiority of our method. We further demonstrated three potential applications that may benefit from our methods, including blood vessel segmentation, retinal image registrations, and diagnosing with enhanced images.

To further validate the safety of MUTE, we collaborated with a vitreo-retinal surgeon who collected 100 retinal images from patients before and after cataract surgery. The MUTE indeed eliminated the haze effect, without introducing any unexpected artifacts. In conclusion, the MUTE algorithm significantly improves the visual quality and the contrast of cataractous retinal images and reveals retinal

structures that might not be detected due to insufficient contrast in the raw images. It enables and facilitates the detection of features like drusen, hard exudates, bleeding hemorrhages, or microaneurysms, as these tasks may use more color or local detail information. Early detection of these findings through cataractous retinal images may improve the outcome of treatment of retinal diseases, and influence the decision-making on surgery.

1.3 On Chapter 5

Although the MUTE algorithm significantly increases the contrast of the image, it distorts the color naturalness of the retinal image. To solve this problem, we propose in **Chapter 5**, the intensity quenching (InQue) scheme for color-restored retinal image restoration/enhancement. In InQue, we present a semi-learning-based algorithm to achieve fully unsupervised single retinal image enhancement. It does not need any synthetic images in any form, the only input is a single retinal image.

InQue consists of four modules namely color implantation, luma heater, luma cooler, and luma rectifier. In the color implantation module, the background color of the input retinal image is replaced by a desired color, sampled from a normal retinal image. The image is then converted into the YCbCr color space, and the luma heater, a one-layer network, increases the intensity of the Y-channel of the color-implanted image to a certain level while correcting uneven or insufficient illumination. Next, the luma cooler, also a one-layer network, decreases the intensity of the Y-channel, increasing the contrast of the image. The cooled-down Y channel is further rectified by the luma rectifier including denoising and unifying the intensity distribution. Finally, the output Y channel, together with the original Cb and Cr channels, is then converted back into the RGB color space, yielding the final enhanced retinal image. We tested InQue on cataract retinal images from public databases and compared its performance with other state-of-the-art deep-learning-based methods using both visual and objective assessments, where it outperformed the others. We also conducted ablation studies on

each module and demonstrated potential applications of InQue in enhancing ultra-wide-field retinal images, retinal image vessel tracking, and diagnosing.

Compared with methods in **Chapter 3** and **Chapter 4**, the InQue has the following two main contributions:

(1) A color implantation module that standardizes the color distribution of the input images while retaining detailed information of the retinal images. This stabilizes the output and preserves the naturalness of the images.

(2) **Unify the retinal image enhancement process into two steps:** illumination correction and dehazing. We propose an intensity-based cost function based on the observation that illumination correction increases the average intensity of the image, while dehazing decreases it. This allows InQue to learn model parameters based on single images and tune the output performance adaptively.

The InQue outputs a naturalness-preserved retinal image with contrast enhancement. It can normalize the intensity and color distribution of input images even from different datasets. We believe that InQue can assist ophthalmologists in accurately diagnosing and treating various retinal diseases, such as diabetic retinopathy, macular degeneration, and glaucoma. Furthermore, the ability of InQue to enhance ultra-wide-field retinal images could prove valuable in detecting peripheral pathologies that may be missed with traditional imaging techniques.

1.4 On Chapters 6 and 7

Chapter 6 and **Chapter 7** focus on a different category of retinal image restoration, namely deconvolution, modeled by a convolutional process of the clear retinal image and a blur kernel. Since retinal images may also suffer from problematic illumination, we embedded the illumination correction in the deconvolution task. We first proposed a separate version, which first corrects the illumination, and then performs the deconvolution. This method was validated on microscopic images in detail in **Chapter 6**. Since the illumination pattern and convolution kernel are unknown, they should be simultaneously estimated during the image restoration. Motivated by this, we

investigated the LR²L method which embedded illumination correction on retinal image blind deconvolution under the framework of Richardson-Lucy deconvolution in **Chapter 7**.

Since the single image blind illumination correction and deconvolution task are severely ill-posed, as there exist many different solution groups of illumination pattern, blurry kernel, and clear image that give rise to the same observation, we use the L_0 -norm penalty on the spatial gradient of the clear latent image to promote its sparsity and L_1 -norm penalty on the spatial gradient of illumination pattern to promote its smoothness. As the L_0 -norm is non-differentiable, we further developed a differentiable approximation of the L_0 -norm. The illumination rectified-blind deconvolution is described by a non-convex optimization problem. To solve this, we derived the closed-form expression of the gradients and used gradient descent with Nesterov-accelerated adaptive momentum estimation to accelerate the optimization, which is more efficient than the traditional half quadratic splitting method.

The LR²L was tested on 1719 images from three public databases. We used four image quality matrixes including image definition, image sharpness, image entropy, and image multiscale contrast for objective assessment, to compare the LRRL against state-of-the-art retinal image blind deconvolution methods. In conclusion, our LRRL corrects the problematic illumination and improves the clarity of the retinal image simultaneously, showing its superiority in terms of restoration quality and implementation efficiency. As the LR²L restores image clarity and corrects for uneven illuminations, the restored images can benefit clinical applications like retinopathy diagnosis.

1.5 On Chapter 8

Finally, in **Chapter 8**, we investigated how retinal image enhancement can benefit retinal blood vessel tracking. We propose a single new, more effective, Finsler function that integrates the strength of these two PDE-based approaches and additionally accounts for a number of retinal image enhancements (dehazing and illumination in

particular). This greatly improved both the previous models and a recent data-driven model, when applied to real clinical and highly challenging images. Moreover, we show clear advantages of each module in our new single Finsler geometrical method.

2. Concluding remarks

Now we are able to answer the six questions proposed in **Chapter 1**.

(7) What is the appropriate model, in both mathematical and optical aspects, which can best describe the image formation process in a fundus image?

The double-pass fundus reflection model is the most physically self-consistent model according to iteration. While it can be regarded as the combination of a haze-formation and illumination model from an optical point-of-view, it also can be simplified into an image structure model from an information point-of-view. Although, the DPFR model also simplified the optical process of retinal image formation, algorithms developed based on this model generate promising results for both visual and objective assessment.

(8) How image illumination correction and image dehazing are combined for single cataract retinal image enhancement/restoration?

The Retinex theory is a bridge to illumination correction and dehazing. As such, illumination correction can be achieved using dehazing, and vice versa. Furthermore, assuming that illumination patterns and haze layers are spatially smoothly varying, the dehazing and illumination tasks can be achieved by image denoising.

(9) To what degrees can a cataract retinal image be enhanced using a dehazing algorithm?

This is answered in Chapter 4, where the original hazy retinal image is significantly enhanced. Hidden structures of the retinal image can be clearly observed after enhancement.

(10) Can the proposed model be further optimized? How?

We proposed two ways to optimize the DPFR model, one for dehazing and one for color-naturalness preservation. The DPFR model may be further optimized by a Physics

Informed Neural Network as the DPFR contains the optical knowledge for retinal image formation. However, there is still a long way to go before the method is safe for clinical applications.

(11) How illumination correction and blind-deconvolution are combined for retinal image deblur?

The illumination pattern can be regarded as an additional model parameter to be optimized, with the same size as the input image. However, the optimization is ill-posed since the number of unknown parameters is way more than the known parameters. To make it well-posed, a penalty function should be introduced to regularize the solution of the optimization. By choosing a proper regularization function, the optimization problem can be solved using gradient descent, and related methods.

(12) How can the proposed models benefit the community of ophthalmology?

We demonstrated how restoring retinal images benefits clinical applications like blood vessel segmentation, retinal image registration, and diagnosis. It enables and facilitates the detection of features like drusen, hard exudates, bleeding hemorrhages, or microaneurysms, as these tasks may use more color or local detail information. The correction of illumination restored the visual quality and structures hidden by the shadow can be observed after enhancement. The dehazing method further increases the image's contrast, especially for cataractous retinal images. Early detection of findings through cataractous retinal images may improve the outcome of treatment of retinal diseases, and influence the decision-making on surgery. For example, patients with severe cataracts might also have macular degeneration. If macular degeneration can be observed through enhanced cataractous retinal images, ophthalmologists are able to choose a better treatment plan to minimize risks and stress for patients.

In conclusion, in this thesis, we developed non-deep learning-based methods for retinal image restoration. The results are very promising, and we expect that after further and massive clinical validation of the safety, reliability, and efficiency of the retinal image restoration methods, they will come from theory to practice., which will

result in better diagnosis and prognosis and lead to more diverse therapeutic strategies and personalized medicine for health care in ophthalmology.

References

1. A. London, I. Benhar, and M. Schwartz, "The retina as a window to the brain—from eye research to CNS disorders," *Nature Reviews Neurology* **9**, 44-53 (2013).
2. N. D. Shemonski, Fredrick A. South, Y.-Z. Liu, S. G. Adie, P. S. Carney, and S. A. Boppart, "Computational high-resolution optical imaging of the living human retina," *Nat. Photonics* **9**, 1-5 (2015).
3. J. Chung, G. W. Martinez, K. C. Lencioni, S. R. Sadda, and C. C. Yang, "Computational aberration compensation by coded aperture-based correction of aberration obtained from optical Fourier coding and blur estimation," *Optica* **6**, 647-661 (2019).
4. A. Arias, and P. Artal, "Wavefront-shaping-based correction of optically simulated cataracts," *Optica* **7**, 22-27 (2020).
5. R. Dutta, S. Manzanera, A. Gambín-Regadera, E. Irlés, E. Tajahuerce, J. Lancis, and P. Artal, "Single-pixel imaging of the retina through scattering media," *Biomedical Optics Express* **10**, 4159-4167 (2019).

Appendix

Impact

1. Scope of this thesis

In this thesis, we propose to use non-deep-learning-based methods to achieve single retinal image restorations tasks. These tasks are categorized into (1) retinal image illumination correction; (2) cataractous retinal image dehazing; (3) retinal image blind-deconvolution. We designed comprehensible algorithms to bring the degraded retinal image back to high-visual quality. The scientific and social impacts are highlighted here.

2. Scientific Impact

The chapters have been published in peer-reviewed scientific journals and presented at (inter)national conferences. They contribute to research in the field of ophthalmology, mathematics, and the ophthalmic photography community.

(1) **Better physical(optical) model.** The first scientific impact is our double-pass fundus reflection model, which enables modeling image formation in retinal photography physically and mathematically correct.

(2) **State-of-the-art enhancement ability.** Following (1), the second scientific impact is the ability of our model to tackle cataractous retinal images. Our algorithm significantly improves the visual quality of hazy retinal images which has great potential in clinical applications. Our research got noticed by the Ophthalmic Photographers' Society and the author was invited to present his works in an oral presentation in ICOP 2023.

(3) **Reveals deep relationship between different image processing tasks.** Our contribution to using image denoising to achieve image dehazing/illumination correctly unifies three classical image processing tasks into a single framework, which can further bring new scientific research topics to the image processing community.

(4) **New retinal image deconvolution algorithm.** We first demonstrated the embedded illumination correction for retinal image blind deconvolution, which outperformed other state-of-the-art methods.

(5) **User-friendly designing.** Most of our models adaptively adjust the model parameters according to the input images. This allows user-friendly designing of software for people who are not familiar with image processing.

(6) **Safe and reliable.** We are not using a deep-learning model to tackle the retinal image restoration, but take an analytical approach, which allows each step as well as intermediate outcomes of the algorithms can be understood by users. As such, we believe that our models are safer, more interpretable, and more reliable than deep-learning methods, that act as black boxes between input and output.

3. Social Impact

A fundus camera is a basic device that is not only used in every eye clinic, but also more generally, such as in general practice and by optometrists where it is used in areas including but not limited to education, research, and diagnosis. Digital retinal image restoration provides a simple and cheap way to increase the image quality of every fundus camera and as such our restoration models have the following social impacts:

Early detection of eye diseases

The retinal image enhancement algorithm can improve the quality of images taken of the eye, making it easier for doctors to identify early signs of eye diseases such as glaucoma, macular degeneration, and diabetic retinopathy. Early detection of findings through cataractous retinal images may improve the outcome of treatment of retinal diseases, and influence the decision-making on surgery that might in some cases even be not necessary. For example, patients with severe cataracts might also have macular degeneration. In these cases, if the macular degeneration can be observed through enhanced cataractous retinal image, ophthalmologists are able to choose a better treatment plan to minimize risks and stress for patients.

Access to healthcare

The high cost of equipment and lack of trained professionals can limit access to techniques of retinal imaging, particularly in low-income and rural areas. By developing

novel image processing algorithms that can enhance retinal images, this research can potentially reduce the need for expensive equipment and trained professionals, making retinal imaging more accessible and affordable for patients in these areas.

Additionally, our research can contribute to the development of telemedicine applications for retinal imaging. Telemedicine enables remote consultation and diagnosis of patients by healthcare professionals, which can be particularly beneficial for patients in rural or remote areas who may not have access to specialized medical care. By enhancing the quality of retinal images, this research can improve the accuracy of telemedicine diagnosis, reducing the need for patients to travel long distances for medical appointments.

Furthermore, our research can also contribute to the development of mobile health (mHealth) applications for retinal imaging. These applications can enable patients to take retinal images using their smartphones or other mobile devices, which can be sent to healthcare professionals for diagnosis. By developing image processing algorithms that can enhance the quality of these images, our research can potentially improve the accuracy of diagnosis through mHealth applications, making healthcare more accessible and affordable for patients.

Better surgical outcomes

The retinal image enhancement algorithm can help ophthalmologists obtain clearer images of the eye during surgery, which can help improve surgical outcomes and reduce the risk of complications.

Further, during retinal surgery, obtaining a clear and detailed image of the eye is essential for accurate diagnosis and treatment. Our image restoration algorithm can enhance the clarity and detail of retinal images, which can help ophthalmologists visualize the structures of the eye more clearly during surgery. This, in turn, can improve surgical outcomes and reduce the risk of complications.

Improved visualization can also help ophthalmologists identify and avoid critical structures, such as blood vessels, during surgery, which can further reduce the risk of complications and improve patient safety.

Collaboration research

During the development of the algorithms in this thesis, the author collaborated with researchers from the Vitreo-Retinal Surgeon in India, the Technology University of Eindhoven, and OiVi from Norway. He has established a network of experts with diverse backgrounds, skills, and perspectives, which enriched our research and led to new ideas and innovations including investigation of new retinal image processing methods, solving practical clinical problems, and commercialization of algorithms.

The author presented research at high-level conferences in the fields of optics, biomedical science, and ophthalmology, including ARVO (The Association for Research in Vision and Ophthalmology), SPIE Photonics Europe, and ICOP (International Conference on Ophthalmic Photography). These presentations generated significant interest and discussion around the research presented in this thesis, contributing to the broader dissemination of knowledge and potential collaborations with other researchers and institutions.

Economic

As mentioned in **Access to Healthcare**, the research in this thesis has the potential to reduce healthcare costs. Besides this, it can contribute to the development of new medical technologies and products. For example, our image processing algorithm can be integrated into existing retinal imaging equipment, creating a new product that can be sold to healthcare providers. This can generate revenue for companies that manufacture and sell this equipment and create job opportunities in the medical device industry.

Our research can also lead to the development of new telemedicine and mHealth applications, creating new business opportunities for healthcare providers and companies. For example, telemedicine companies can incorporate our image-

processing algorithm into their platforms, improving the accuracy of diagnosis and treatment and attracting more customers.

4. Target audience

The first target audiences for this research are academics and researchers in the field of medical image processing, ophthalmology, and computer science. They may be interested in learning about our novel image-processing algorithm and its potential applications in the diagnosis and treatment of eye diseases.

Second, ophthalmologists and healthcare providers who work with patients with retinal diseases may also be interested in this research, since providing clearer and more detailed images of the eye, can help to make more accurate diagnoses and tailor treatments more precisely to each patient's needs. This can lead to better patient outcomes, including improved visual acuity and quality of life.

Last, the medical device industry, including companies that manufacture and sell retinal imaging equipment, may be interested in my research as well. Our image processing algorithm can be integrated into existing equipment or used to develop new products that can improve patient care and generate revenue.

Thank words

致谢

Appendix 2

Over the past four years, my Ph.D. journey has been a profound and transformative experience, akin to a captivating odyssey through the realms of knowledge. It has been a journey illuminated by the pursuit of truth, guided by the compass of resilience, and marked by the sheer joy of intellectual exploration. In the quiet depths of research, I unearthed the most resounding echoes of inspiration, and amidst the challenges and late nights, I discovered the boundless beauty of academic growth. Each experiment conducted, each thesis page written, and each hurdle overcome served as brushstrokes on the canvas of this remarkable expedition, painting a portrait of unwavering determination and unbridled passion. As I prepare to stand on this threshold of graduation, I am reminded that this journey was not solely about acquiring expertise; it was about forging connections, nurturing relationships, and celebrating the triumph of human potential. These four years have been an enduring testimony to the limitless capacity of my mind to explore, discover, and innovate, and I am profoundly grateful for every moment that has shaped this remarkable chapter of my life. I would like to thank so many people and institutions for accompanying me and helping me along the way. Dear all, I am grateful to have you in my life. You lighten my world and warm my heart.

First, I sincerely appreciate my promotor, Prof. Dr. Carroll A. B. Webers, for offering to accept me as a Ph.D. candidate under this supervision. I sincerely thank the Chinese Scholarship Council for granting me the scholarship to support me in pursuing a Ph.D. degree at Maastricht University, the Netherlands. I also thank my daily supervisor, Dr. Tos T. J. M. Berendschot, for supervising me, guiding me, and helping me to realize my academic dream. From the very beginning, you have been an inspiring source of knowledge and wisdom. Your dedication to research, your passion for your field, and your unwavering commitment to the success of your students have made a lasting impact on me. I feel incredibly fortunate to have had the opportunity to work under your supervision. Your guidance has not only helped shape my research but has also contributed significantly to my personal and professional growth. I appreciate your patience during the challenging moments and your encouragement during the times when I needed it the most. Your mentorship has not only equipped me with the necessary skills but has also instilled in me a deep passion for research and learning. As I prepare to embark on the next phase of my academic and professional journey, I want you to know that your influence will continue to guide me throughout my career.

Throughout these four years of rigorous PhD studies, I've been fortunate to have Christian as one of my best friends in the Netherlands and a steadfast companion on this academic journey. Your warm friendship extended far beyond the classroom and research labs. Your selfless efforts in introducing me to the campus and helping me navigate the academic landscape were invaluable.

I would also like to thank Dr Ashwin Mohan, and my colleagues, Luigi, Magali, Pascal, Maartje, Lindsay, Joukje, Yu Yu, Lotte, and Wenting for providing me the great support during my stay at the University.

接下来，我要衷心感谢来自祖国的同胞和亲友们在这四年来的关心，照顾和帮助。是你们的存在和出现丰富温暖了我在异国他乡的学习和生活。衷心祝愿大家生活顺利，事事开心。

我想趁着这个特殊的时刻向我的爷爷张景桥先生表达我的感激之情。您是我生命中一位不可或缺的导师和启发者，您的陪伴和教导对我而言是无价之宝。自从我还是个孩子的时候，您就以丰富的知识和温和的智慧为我讲述科普小故事。每当晚上睡前，那些故事都如一颗颗珍贵的种子，落在我心灵深处，激发了我的好奇心。这些瞬间不仅丰富了我的童年，也在我心中埋下了追求知识、探索未知的激情。正是您，让我明白了知识的珍贵和力量。您是一位早年的大学生，拥有丰富的学识和经验，并从不吝啬地将这一切分享给我。您的耐心教导和无私支持，为我塑造了坚韧和勇气，这些品质在我攻读研究生学位的道路上发挥了至关重要的作用。如今，我站在即将获得博士研究生学位的门槛上，我要感谢您的教导，这个成就与您的关怀和支持分不开。您一直是我前进路上的灯塔，您的智慧和善意激励着我继续追求知识，继续探索未知的领域。亲爱的爷爷，我深知没有您，我不会成为今天的自己。感恩您一直以来的教导和关爱。我永远珍惜这份特殊的亲情和知识的馈赠。

亲爱的父亲张洪涛先生、母亲许宇漫女士以及孙荣叔叔，我希望能借此机会表达我深深的感激之情，感谢你们为我创造了温暖、健康的家庭，以及给予了我无限的理解与支持。在我追求自己热爱事物的旅程中，你们的陪伴和鼓励一直是我最大的动力。父亲，您一直是我动手和逻辑思维能力的榜样。您拥有一双灵巧的双手，创造出令人惊叹的手工模型，这一直是我学习计算机编程并开发问题解决算法的灵感来源。您的耐心指导和技巧传授一直都是我成长过程中的宝贵财富。母亲，您是我自律和抗压能力的楷模。您一直展现出女强人的形象，

不断强调自我约束和抵抗压力的重要性。这使我能够在海外生活中保持坚韧，时刻充满开朗和自信。孙荣叔叔，感谢您一直以来对我和母亲的支持与帮助。您的陪伴让我在海外学习期间感到安心和宽慰。您的关心和支持为我们提供了坚实的后盾。你们的爱和支持是我成为今天的自己的基石。感谢你们无私的奉献和慷慨的爱，使我能够追求我的梦想并不断前进。我将永远珍视你们为我创造的幸福家庭和美好回忆。

尊敬的周金华老师、梁振老师、龚雷老师和任宇轩老师，感谢您们在我学术生涯中的关怀、栽培和指导。周老师和梁老师，您们是我硕士学业中的伟大导师。在您们的悉心指导下，我不仅提高了解决问题和论文写作的技能，还培养了对科研的深刻兴趣。您们的包容和耐心使我能够不断成长，为将来的学术道路打下了坚实的基础。龚老师和任老师，感谢您们在我博士阶段的指点和帮助。您们的专业知识和慷慨分享为我的博士研究提供了宝贵的支持。您们的教导使我能够更快地融入新的科研领域，提高了我的学术能力。在您们的指导下，我不仅在学术方面取得了进展，还在人生道路上受益匪浅。您们的教诲和支持塑造了我成为一个更好的学者和更有价值的人。再次感谢您们为我所做的一切。我将永远珍惜您们的指导和帮助，对您们的感激之情无法言表。

我要向我的爱人庄堰林女士表达我深深的感激之情。在我即将获得博士学位的时刻，我不仅仅想要庆祝这个成就，还想要感谢你为我所做的一切。整个博士学位的过程是漫长而充满挑战的，但有你一直在我身边，支持我，鼓励我，陪伴我度过了这段旅程。你的理解、耐心和无私的支持一直是我不断前进的动力。你在我生活中的存在让我倍感幸运。我还记得你在我最困难的时候为我所做的一切：夜晚的陪伴，为我做饭，倾听我的烦恼，鼓励我不放弃。这一切都深深地烙在我的心中，我永远不会忘记。现在，我获得了博士学位，这是我们共同的成就，因为没有你的陪伴和支持，我可能不会走到今天。未来充满了希望和机会，我知道有你在我身边，我们将一起迎接这一切挑战。

我要感谢我的好朋友陈军华。你的友情和帮助在我博士研究中起到了至关重要的作用，特别是在探索机器学习领域方面。正是你将我引入了机器视觉这个令人兴奋世界，为我提供了宝贵的指导和建议。你分享的见解和思考方式不仅激发了我的好奇心，还启发了我对医学图像处理的热情。正是与你的探讨给了我博士第一篇论文的思路，这是我学术生涯中的一个重要里程碑。你的支持和友情一直是我前进路上的明灯。我感到非常幸运能够拥有你这样一个出色的朋友，一个总是愿意分享知识和启发思考的朋友。

张硕学长，我想借此机会表达我的感激之情。尽管我们共事的时间并不长，但您在我的就业指导 and 心态调整方面发挥了关键作用。我永远不会忘记您为我所做的一切。每当我在就

业方面遇到困难时，您总是一位不知疲倦的导师和朋友。您耐心地倾听我的问题，细心指导我，给予我宝贵的建议。更重要的是，您在我心理辅导方面的帮助，帮我度过了许多焦虑和犹豫的时刻。您的支持和关心让我感到无比温暖和鼓舞。感谢您不仅为我的职业发展提供了

帮助，还为我树立了一个榜样，展示了如何成为一位善良、慷慨的人。我由衷感谢您的耐心和善意。祝愿您的事业蒸蒸日上，愿您的爱人和孩子们在您的陪伴下幸福美满。您的友情和支持对我意义深远，我将永远感激不已。后面一起多发论文！

亲爱的赵沅蒂、孙思雨、杨金洲、李因因、汪婷、王甜甜、曾海燕、凡行、吴敏、于雨，我想趁此机会向你们表达我的感激之情。四年前我们的生活并没有交集，但是从那以后在我生活中的各个阶段，你们都是我宝贵的朋友，伴随我走过了许多旅程。你们的陪伴和帮助对我而言意义非凡。每一次我们在一起度过的时光都是宝贵的回忆，充满欢笑和互相支持。无论是分享生活琐事，还是探讨人生哲理，你们的存在使我的生活更加丰富多彩。虽然我可能即将回国，但我深信我们的友情将一直保持。我期待着将来有机会再一起享受美味的火锅、辣糊糊、胡辣汤、螺蛳粉，深入探讨人生的各种可能性。你们是我生活中不可或缺的一部分，我感到非常幸运能够拥有你们这样的朋友。再次感谢你们的陪伴和帮助。愿我们的友情继续在未来绽放，共同创造更多美好的回忆。

我要感谢我的挚友们张镨、秦天健、彭康豪，我们一起度过的那些课余时间，组成的四人团队，是我宝贵的回忆，也是我学生生涯中的一大幸运。在我们的团队中，我们共同努力，合作无间，克服了一个个的难关。康豪兄总是无论是面临学术挑战还是娱乐乐趣，我们总是相互支持，共同前行。你们的友情和团队精神让我感到无比幸运。在这段时间里，我们不仅仅是玩伴，更是彼此的知己和伙伴。我们一起成长，一起分享成功和挫折（笑，这些经历将永远珍藏在我的心中。再次感谢你们的陪伴和合作。愿我们的友情继续长久，一直支持彼此的成长和追求。期待着未来还有更多的冒险和欢笑。

最后，我深知，我所取得的一切成就都离不开祖国的培养和关怀。祖国提供了良好的教育资源和科研环境，为我提供了广阔的发展空间。她鼓励了我的求知欲望，培养了我的才华，让我有机会在国际舞台上展示自己的价值。在遥远的异国他乡，我时刻感到祖国的鼓励和温暖，这让我充满信心去追求更高的目标。祖国的培养之恩永远铭刻在我的心中。无论我身在何地，我都将继续传承祖国的文化，积极参与国际交流，为祖国的发展贡献一份微薄的力量。

Curriculum Vitae

Shuhe Zhang was born in Wuhan, 13-August 1994. He has a 10-year education background and experience in the field of optics. He is proficient in Geometrical optics, Physical optics, as well as Fourier Optics in both experimental and simulation studies. He masters the professional software and hardware skills required for optical imaging instruments from theoretical analysis, and simulation to design and testing. He obtained his bachelor's degree in July 2016 at Anhui Medical University and continued to finish his master's degree in Biomedical Engineering in July 2019. In the same year, he was supported by the China Scholarship Council to pursue his Ph.D. degree at Maastricht University, the Netherlands. He has rich experience in in-depth cooperation with English-speaking international teams and has led a small scientific research team to develop a new quantitative phase imaging system.

Research projects and experiences during the Ph. D's study 2019 to 2023

- Image processing algorithm for fundus imaging

Shuhe Zhang takes fully responsible for the collection of pre- and post-operative retinal images for cataract patients, model design, and algorithm research for cataracts retinal image restoration. In cooperation with Vasan Eye Center in India, He develops fundus image enhancement algorithms that are beneficial to clinical diagnosis. He is now collaborating with the Eindhoven University of Technology, and developing a fundus blood vessel segmentation algorithm. Related works are published in *Signal Processing*, as well as ARVO 2021 and ARVO 2022. Other related works are published on *Medical Image Analysis* and *Computer methods and programs in biomedicine*.

- Computational imaging: New reconstruction algorithm for quantitative differential phase imaging

Shuhe Zhang participated in the development of a quantitative phase microscopy platform and algorithm based on differential phase contrast (DPC) imaging in cooperation with Anhui Medical University. He was fully responsible for the development of new phase reconstruction algorithms to enhance the robustness of quantitative DPC microscopy to system background errors and noise signals. Compared with the traditional algorithm, the algorithm proposed by his team can stably reconstruct high-fidelity phase information under 10 times the noise intensity, which greatly improves the practicability and stability of quantitative DPC microscopy. The related work was submitted in *Computer methods and programs in biomedicine*. This work is funded by the University Fund of the Province of Limburg (SWOL 2022), The Netherlands.

- Computational imaging: New reconstruction algorithm for Fourier ptychographic microscopy

He developed a new reconstruction algorithm for Fourier ptychographic microscopy to improve the robustness of imaging quality to systematic errors. Compared with the traditional algorithm, the new reconstruction algorithm proposed by his team enhances the stability of the system to 100 times the intensity of Gaussian noise, 1000 times the intensity of salt and pepper noise, and 10 times the intensity of Poisson noise. And the reconstruction quality is no longer affected by the LED lighting position deviation within 2mm. This work was published on *Signal Processing*.

- Developing hyperspectral fundus camera

Shuhe Zhang participated in the research and development of hyperspectral fundus cameras. In cooperation with the internal scientific research of Maastricht University in the Netherlands, we have explored the system design and data analysis of multispectral fundus cameras.

List of publications

Appendix 4

1. **S. Zhang**, C. A. Webers, and T. T. Berendschot, "Inque: An Intensity-Quenching Scheme for Fully Unsupervised Large-Scale Retinal Image Enhancement," Available at SSRN 4539119 (2023).
2. **S. Zhang**, C. A. Webers, and T. T. Berendschot, "Luminosity rectified blind Richardson-Lucy deconvolution for single retinal image restoration," *Computer Methods and Programs in Biomedicine* 229, 107297 (2023).
3. **S. Zhang**, T. Peng, Z. Ke, H. Yang, T. T. Berendschot, and J. Zhou, "Retinex-qDPC: Automatic background-rectified quantitative differential phase contrast imaging," *Computer Methods and Programs in Biomedicine* 230, 107327 (2023).
4. **S. Zhang**, T. Peng, Z. Ke, M. Shao, T. T. Berendschot, and J. Zhou, "High-fidelity quantitative differential phase contrast deconvolution using dark-field sparse prior," in Sixteenth International Conference on Photonics and Imaging in Biology and Medicine (PIBM 2023)(SPIE2023), pp. 234-242.
5. **S. Zhang**, A. Mohan, C. A. Webers, and T. T. Berendschot, "MUTE: A multilevel-stimulated denoising strategy for single cataractous retinal image dehazing," *Medical Image Analysis* 88, 102848 (2023).
6. **S. Zhang**, T. T. Berendschot, and J. Zhou, "ELFPIE: an error-laxity Fourier ptychographic iterative engine," *Signal Processing* 210, 109088 (2023).
7. N. J. van den Berg, **S. Zhang**, B. M. Smets, T. T. Berendschot, and R. Duits, "Geodesic Tracking of Retinal Vascular Trees with Optical and TV-Flow Enhancement in SE (2)," in International Conference on Scale Space and Variational Methods in Computer Vision(Springer International Publishing Cham2023), pp. 525-537.
8. T. Peng, Z. Ke, **S. Zhang**, J. He, P. Wang, F. Wang, Z. Zhong, S. Fang, H. Shi, and R. Lu, "Quantitative differential phase contrast phase reconstruction for sparse samples," *Optics and Lasers in Engineering* 163, 107478 (2023).
9. **S. Zhang**, C. A. Webers, and T. T. Berendschot, "A double-pass fundus reflection model for efficient single retinal image enhancement," *Signal Processing* 192, 108400 (2022).
10. **S. Zhang**, A. Mohan, C. A. Webers, and T. T. Berendschot, "MUTED: Reveal the hidden structures of cataractous retinal images," *Investigative Ophthalmology & Visual Science* 63, 222–F0069-0222–F0069 (2022).
11. **S. Zhang**, T. T. Berendschot, J. Zhou, and M. Shao, "BDIC: boosting the performance of optical microscopy using blind deconvolution and illumination correction," in Optics, Photonics and Digital Technologies for Imaging Applications VII(SPIE2022), pp. 92-99.
12. M. Shao, K. Wang, Z. Wang, T. Peng, **S. Zhang**, J. Zhang, S. Fang, F. Wang, S. Zhang, and M. C. Zhong, "An apparatus for qualitative assessment of the shading ratio of oblique illumination and real-time high-contrast imaging," *Journal of Biophotonics* 15, e202200122 (2022).
13. T. Peng, Z. Ke, **S. Zhang**, M. Shao, H. Yang, X. Liu, H. Zou, Z. Zhong, M. Zhong, and H. Shi, "A Compact Real-time Quantitative Phase Imaging System," (2022).
14. T. Berendschot, **S. Zhang**, A. Mohan, and C. Webers, "Revealing the information of cataractous retinal images using image processing," in ACTA OPHTHALMOLOGICA(WILEY 111 RIVER ST, HOBOKEN 07030-5774, NJ USA2022), pp. 16-17.

15. S. Zhang, **S. Zhang**, and T. T. Berendschot, "Harnessing abruptly auto-defocusing beam to enhance the Raman signal in aqueous humor: A simulation analysis," *Optics Communications* 496, 127125 (2021).
16. **S. Zhang**, W. Carroll, and T. T. Berendschot, "A double-pass fundus reflection model for efficient single retinal image enhancement," *Investigative Ophthalmology & Visual Science* 62, 1777-1777 (2021).
17. Q. Zhao, H.-W. Wang, P.-P. Yu, **S.-H. Zhang**, J.-H. Zhou, Y.-M. Li, and L. Gong, "Trapping and manipulation of single cells in crowded environments," *Frontiers in bioengineering and biotechnology* 8, 422 (2020).
18. **S. Zhang**, M. Shao, X. Yang, and J. Zhou, "Calculating optical forces with skew line ray model for Gaussian beam," *Optik* 203, 163935 (2020).
19. **S. Zhang**, J. Zhou, M.-C. Zhong, and L. Gong, "Nonparaxial structured vectorial abruptly autofocusing beam," *Optics Letters* 44, 2843-2846 (2019).
20. **S. Zhang**, J. Zhou, and Y.-X. Ren, "Ray optics analysis of optical forces on a microsphere in a $(2+ 1)$ D Airy beam," *OSA Continuum* 2, 378-388 (2019).
22. M. Shao, **S. Zhang**, J. Zhou, and Y.-X. Ren, "Calculation of optical forces for arbitrary light beams using the Fourier ray method," *Optics Express* 27, 27459-27476 (2019).
23. 张书赫, 邵梦, 张盛昭, and 周金华, "傅里叶域中的光线," *物理学报* 68, 214202 (2019).
24. 张书赫, 邵梦, 王奕, 段宇平, and 周金华, "光波的光线表征," *Ray Characterization of Optical Waves* 56, 230003 (2019).
25. **S. Zhang**, J. Zhou, and L. Gong, "Skew line ray model of nonparaxial Gaussian beam," *Optics express* 26, 3381-3393 (2018).
26. 张书赫, 邵梦, and 周金华, "光线庞加莱球法构建的结构光场及其传输特性研究," *物理学报* 67, 204201 (2018).
27. 张书赫, 梁振, and 周金华, "运用四元数分析椭球微粒所受的光阱力," *物理学报* 64, 098702 (2015).

