Mouse Brain Reconstruction and Analysis with Applications to Epilepsy Study

---

A Dissertation

Presented to

the Faculty of the School of Engineering and Applied Science

University of Virginia

In Partial Fulfillment

of the requirements for the Degree

Doctor of Philosophy (Electrical Engineering)

by

Haoyi Liang

May 2019

Approval Sheet

This dissertation is submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy (Electrical Engineering)

Haoyi Liang

This dissertation has been read and approved by the Examining Committee:

Daniel S. Weller, Adviser

Jaideep Kapur, Committee Member

Tingting Zhang, Committee Member

Gustavo Rohde, Committee Member

Scott Acton, Committee Chair

Accepted for the School of Engineering and Applied Science:

Dean, School of Engineering and Applied Science

May 2019

Abstract

Biomedical images play an important role in biomedical research and diagnostics. However, the raw data acquired by imaging equipment sometimes are not suitable for direct observation and analysis. For example, the confocal microscope enables the observation of tissue in detail and reflects structures on the scale of a single cell or even finer. However, the raw data acquired by a confocal microscope usually contain multiple artifacts. These distortions include low SNR, irrelevant tissue clutter and geometric distortions. Image restoration and reconstruction algorithms, such as denoising, stitching and registration, are necessary before further analysis on these data. Another kind of widely used biomedical image is Magnetic resonance imaging (MRI). As an important non-invasive imaging technique, MRI facilitates the diagnosis of and research in diseases like brain cancer, Alzheimer's and Parkinson's. Reconstruction algorithms that turn the data in the frequency domain into the spatial domain are necessary after MRI scanning. Selecting proper parameters for these reconstruction algorithms is crucial to get high quality MRI.

In this thesis, we focus on the image processing requirements on mouse brains for status epilepticus (SE) research. Epilepsy is a group of neurological disorders characterized by epileptic seizures. The rate of adverse outcomes of SE correlates with the duration of seizures, and thus early termination of SE is important. High-resolution 3D mouse brains provide details about SE development at multiple scales from cells, circuits, systems, to the whole brain level. Figuring out the pathways of SE development helps neuroscientists better understand the mechanism of SE and develop new drugs to terminate SE at an early stage.

Currently, the main way to investigate brain activity during SE at single neuron resolution is microscopy imaging. However, the penetration depth of some immunohistochemical neuron stains is limited to about 200 microns, and this requires mouse brains to be sliced before imaging. To better visualize and understand brain activity during SE, this thesis comprises three parts: 3D mouse brain reconstruction with microscopy data, auxiliary modality imaging to aid multi-brain analysis, and analysis of microscopy data. First, to recover the high resolution 3D mouse brain volumes, we propose tissue flattening and structure-based intensity propagation for 3D mouse brain reconstruction. Experiments are conducted on 367 multilayer sections from 20 mouse brains. The average reconstruction quality measured by the structure consistency index increases by 29% with the proposed structure-based intensity propagation. In order to better conduct multi-brain comparison and registration, an auxiliary imaging technique, MRI, is investigated in the second part. MRI is able to provide a complete 3D mouse brain volume before slicing. With the proposed parameter selection method, high quality synthetic MRI are reconstructed from measured data in the frequency domain. Finally, automatic cell detection enables neuroscientists to obtain cell activation information on the whole brain scale. To improve the detection accuracy in regions with densely-packed granule cells, we design a new center coding scheme for convolutional neural networks (CNN). With 3D mouse brain reconstruction and automatic cell detection, the 3D topology of cell activation is acquired, and this facilitates neuroscientists' investigations of the mechanism of SE at multiple scales.

Acknowledgments

I would like to thank my advisor, Dr. Daniel Weller. It is really lucky to work with the guidance from Dr. Daniel Weller. Besides his deep and wide knowledge on signal processing, I learned a lot from the ways that he organizes works, collaborates with colleagues and mentors students. I hope I could be as thoughtful, energetic, and diligent, as Dr. Daniel Weller in my future work. I am also grateful to my committee members, Dr. Gustavo Rohde, Dr. Tingting Zhang, Dr. Jaideep Kapur and Dr. Scott Acton for their mentoring and guidance on the work presented in my thesis.

I wish to thank Dr. Scott Acton and friends from the VIVA Lab. The times with VIVA friends help me get familiar with the life at UVA in the initial months. I will always treasure the memories at the Walnut Creek Park, and Dr. Acton's, Dr. Mukherjee's, and Dr. Vaccari's places. I wish to thank Dr. Jaideep Kapur and many members from Dr. Kapur's group. The opportunities to work on the exiting neuroscience project with Natalia Dabrowska and Aijaz Naik are precious.

I would like to thank the friends who I met at the conferences of Asilomar, ICIP and ISBI. Discussions with them broadened my knowledge on the image processing, and helped me to better learn the trends in both academia and industry research.

Also I wish to thank the friends I met at Charlottesville. It is great to spend time with you in the Shenandoah National Park, the North Ground Gym and weekend gatherings. With you being side, the hard time is easier.

Finally, I wish to thank my father and mother. My gratitude towards your educating and supports through all the time is beyond words.

Contents

Contents **g**
List of Tables . . . i
List of Figures . . . j

1 Introduction **1**
1.1 Background of the thesis . . . 1
1.2 Objectives and contribution . . . 2
1.3 Dissertation outline . . . 4

2 3D Mouse Brain Reconstruction **6**
2.1 Reconstruction backgrounds . . . 9
2.1.1 Existing brain reconstruction works . . . 9
2.1.2 Alignment pipeline in this work . . . 11
2.2 Proposed structure correction methods . . . 12
2.2.1 Flattening . . . 12
2.2.2 Structure-based intensity propagation . . . 14
2.2.3 Implementations of structure-based intensity propagation . . . 20
2.3 Experiments . . . 23
2.3.1 Data acquisition . . . 25
2.3.2 Evaluation criteria . . . 26
2.3.3 Section-to-section alignment . . . 29
2.3.4 Flip detection . . . 30
2.3.5 Whole brain reconstruction . . . 31
2.4 Discussion . . . 34

3 Auxiliary Imaging Modality **36**
3.1 Existing NR-IQA methods . . . 38
3.1.1 Approaches with statistical features . . . 39
3.1.2 Approaches with structure indexes . . . 39
3.1.3 MetricQ . . . 40
3.2 Comparison-based image quality assessment . . . 41
3.2.1 Framework of C-IQA . . . 42
3.2.2 Justification of CQ . . . 45
3.2.3 Distortion sensitivity weighting . . . 47

3.2.4 Comparison between CDQ and SSIM . . . 49
3.3 Parameter selection . . . 50
3.3.1 Image reconstruction . . . 51
3.3.2 Parameter trimming . . . 52
3.4 Experiments . . . 54
3.4.1 Minimum resolution of C-IQA . . . 55
3.4.2 Experiment verification for parameter selection . . . 56
3.4.3 Application in parameter trimming . . . 63
3.5 Discussion . . . 64

4 Cell Detection for Reconstructed Brains 66
4.1 Related works . . . 69
4.1.1 Object Detection in Natural Scenes and Cell Analysis . . . 69
4.1.2 Counting vs. Detection . . . 71
4.2 Proposed coding scheme . . . 71
4.2.1 Criteria . . . 72
4.2.2 Repel Coding . . . 74
4.2.3 Relation to existing works . . . 75
4.3 Experiments . . . 76
4.3.1 Datasets . . . 76
4.3.2 Experimental settings . . . 78
4.3.3 Results and analysis . . . 80
4.4 Conclusion . . . 82

5 Conclusions and Future work 85
5.1 Discussions and summary of the proposed works . . . 85
5.2 Future works . . . 87
5.3 Publication list of this work . . . 88
5.3.1 Journals . . . 88
5.3.2 Conferences . . . 89
5.3.3 Patents . . . 90

Bibliography 91

List of Tables

2.1 Run time comparison of two SIP implementations 20
2.2 Structure consistence index of section-to-section alignment 28
2.3 Incorrect flip detection number 31
2.4 Structure consistence index of whole brain reconstruction 34

3.1 Accuracy of parameter selection for image reconstruction 59
3.2 Accuracy of parameter selection for bilateral filter 60
3.3 Accuracy of parameter selection for BM3D 63
3.4 Computation saved by parameter trimming 64

4.1 Entropy and Reversibility of Different Coding Schemes 71
4.2 F1 score of the FCNN-A 81
4.3 F1 score of the FCNN-A with res. blocks 82
4.4 Detection results with FCNN with plane network 82
4.5 Detection results with FCNN with res-block 83

List of Figures

2.1 The top row is the reconstruction pipeline adopted in this work. Section flattening and structure propagation are the two techniques proposed to improve the reconstruction quality. The three figures at the bottom row are the stacked raw sections, the reconstructed brain without the proposed techniques, and the reconstructed brain with the proposed techniques. The arrows in the three bottom figures indicate the ventral-dorsal direction. 7

2.2 Intermediate results of tissue flattening. (a) is the maximum intensity projection of one section sliced on the horizontal plane. The dashed line in (a) indicates the positions of side views. (b) is the side view of the raw section. (c) shows the detected surfaces before hole fixing. If the tissue is projected onto the bottom surfaces in (c), the structure around the ventricle is changed as shown in (e). (d) shows the detected surface after hole fixing. However, with incorrect projection direction, structures at the boundaries are altered as shown in (f). With correctly detected surfaces and the projection direction, (g) shows the flattened tissue from (b). 12

2.3 First and second rows are the cross-section views of a section before and after flattening. In (a), tissue boundaries are not accurately reflected on the bottom layer because of the upward tilt. In (b), the size of the hippocampus grows from the top to the bottom, but this structural trend is not correctly reflected on the bottom layer. 13

2.4 (a) A rendered multilayer section of size $9569 \times 9569 \times 200\ \mu m^3$ in 3D. This section is sliced on the horizontal plane. From the top to the bottom (on z-direction), the brain and the hippocampus expand. (b) Estimated structure map on the direction of top-bottom in the imaging-plane. (c) Estimated structure map the direction of left-right in the imaging-plane. MIP, the bottom surface of flattened section and the propagated bottom surface are (d), (e) and (f) respectively. . . 16

2.5 An illustration of the structure-based intensity propagation on the 2D case. (a) is a cross-section view of the raw data. The solid arrow indicates the structure direction, and the dashed arrow corresponds to the normal direction. (b) and (c) are the stacks after transforming each lines to the surface positions, equivalent to I^{top} and I^{bottom} in Alg. 1. (d) plots the gradients in (a) and illustrates the SVD operation. The propagated surface lines reflect the positions of boundaries more accurately. 19

2.6 Data collection procedure. The red fluorescent markers are expressed during seizure spread while the mice are alive. After sacrificing animals, immunofluorescence staining for NeuN is performed. 25
2.7 (a) NeuN. (b) tdTomato. (c)Aligned two channels. All three images, (a), (b) and (c), are sections on the coronal plane. 25
2.8 (a) is a cross-section view of two aligned sections. (b) is the cross-section view of extracted edges. (c) is the edge prominence map voted by edge points in (b). The structures in the top section extend one layer downward, and the structures in the bottom section extend one layer upward. Therefore, the edge prominence in the top section and the bottom section has two layers overlapped in (c). (d) plots the edge prominence in overlapped layers in (c). 27
2.9 Structure consistence index of registered section pairs. (a) Sections imaged on the horizontal plane. (b) Sections imaged on the coronal plane. 28
2.10 Examples of section-to-section registration with four different representatives. (a) MIP. (b) Surface. (c) BRS [1]. (d) Proposed SIP. 29
2.11 The surface layer after tissue flattening in (a) shows less signal intensity and structure contrast compared with the propagated surface in (b). Compared with the maximum intensity projection in (c), the propagated surface in (d) removes the structures that only stay in certain layers and preserves the major structures. (a) and (b) are sections on the horizontal plane. (c) and (d) are sections on the coronal plane. 31
2.12 Reconstructed Brain 1, sectioned on the horizontal plane. The arrows indicate the ventral-dorsal direction. The frames indicate the positions of the cross-section views. The cross-section views are on the coronal plane. (a) MIP: 0.2746. (b) Surface: 0.2820. (c) BRS: 0.2560. (d) SIP: 0.4419. 32
2.13 Reconstructed Brain 11, sectioned on the coronal plane. The arrows indicate the posterior-anterior direction. The frames indicate the positions of the cross-section views. The cross-section views are on the horizontal plane. (a) MIP: 0.3991. (b) Surface: 0.3419. (c) BRS: 0.4079. (d) SIP: 0.5240. 33

3.1 Flow Chart of the Comparison-based IQA: P_1 and P_2 are local patches from input images, I_1 and I_2, at the same location respectively. The Content Detection module determines whether there is a meaningful structure in the difference patch; the Contribution module calculates which patch mainly contributes to the difference patch; the Distortion Sensitivity Weighting module compensates the distortion sensitivity difference of patches with various texture complexities. The output, comparison-based index, indicates the relative quality of P_1 based on P_2. 42
3.2 Patch samples with various texture complexity are selected from LIVE [2] and CSIQ [3] to verify the difference of distortion sensitivity in CQ. The assumption is that patches with complex texture are more robust to noise, while patches with flat texture are more robust to blurring. 49
3.3 Relations between the output of Contribution module, $ctri$, and texture complexity, T_ind. Each circle in the figure represents a sample patch. All the sample patches are degraded by the same amount of distortion for blurring and noise. . 50

3.4 (a): original Brain image [4]; (b): reconstructed result with $\beta = 1.22 \times 10^{-6}$; (c): reconstruction result with $\beta = 4.46 \times 10^{-1}$; (d): reconstructed result with $\beta = 10$. 52

3.5 (a): Each line corresponds to an algorithm instance with a different regularization parameter. (b): Qualities of reconstructed results with different regularization parameters after 160 iterations. 53

3.6 Images in the gray scale [2] for the illustration of minimum resolution. 54

3.7 Minimum resolution of Comparison-based IQA algorithm: (a) CDQ scores of denoised images compared with their previous images (series1) and the one before previous images (series2); (b) SSIM scores and cumulated CDQ scores of "caps" in (a); (c) SSIM scores and cumulated CDQ scores of "coinsinfountain" in (a). 54

3.8 One example of image reconstruction parameter selection. The best regularization parameter for this image, $\beta^* = 2.81 \times 10^{-2}$. The area highlighted by the red box is enlarged in Fig. 3.9. 57

3.9 Patches from the highlighted areas in Fig.3.8. The regularization parameter of the total variation term is β. RR-IQA [5] selects β_1, DIIVINE and BRISQUE select β_2, Anisotropy selects β_3, CQ and CDQ select β^*, and MetricQ select β_4. As β increases, noise is suppressed as shown in blue rectangles, while subtle structures is blurred as shown in red rectangles. 58

3.10 CQ index of Fig. 3.9 (b) (β_1) based on Fig. 3.9 (e) (β^*). (a) is the difference between the image with β_1 and β^*. The white in (b) stands for structured areas in (a) and the black stands for noisy area. (c) shows the contribution difference between β_1 and β^*. On this local patch, $CQ(P_{\beta_1}, P_{\beta^*}) = -1.83 \times 10^{-3}$. 59

3.11 CQ index of Fig. 3.9 (f) (β_4) based on Fig. 3.9 (e) (β^*). (a) is the patch difference between the image with β_4 and β^*. The white areas in (b) corresponds to structured area. In (c), the negative values means that the contribution comes from β^*. On this local patch, $CQ(P_{\beta_4}, P_{\beta^*}) = -2.50 \times 10^{-4}$. 60

3.12 The SSIM differences of the best images chosen by different NR, RR, and Comparison-based IQA methods for image reconstruction parameter selection. 62

3.13 The SSIM differences of the best images chosen by different NR, RR, and Comparison-based IQA methods for the bilateral filter parameter selection. . . 63

3.14 Outliers of parameter selection 64

3.15 The SSIM differences of the best images chosen by different NR, RR, and Comparison-based IQA methods for the BM3D parameter selection. 65

3.16 Comparison between convergence with and without parameter trimming on "buildings" 65

4.1 Cell detection with the proposed repel coding for the granule cells in mouse brain tissues. (a) Input image. (b) Output of the network. (c) Zoomed-in patches. Green dots are the labeled cell centers, and red dots are the detected cell centers. 67

4.2 Illustrations of different coding schemes. (a) An example image with cell centers labeled with red dots. The rectangle in (a) indicates the zoomed-in patches in (b)-(e). (b) shows the coding with the rectangle kernel [6]. (c) shows the coding with the Gaussian kernel [6, 7]. (d) shows the coding with proximity coding [8, 9]. (e) is the proposed repel coding. The proposed center coding scheme is superior to proximity coding in two ways. First, the repel coding provides a slower response decay away from the cell center, which is helpful to the stability during training. Secondly, the repel coding suppresses the responses between two cell centers. This is important to recover the raw dot labels during prediction. Coding values in (b) to (e) are normalized to 0-1 for comparison. . 72
4.3 Sample images from the four datasets. (a) Synthetic Vgg cells. (b) Adipocyte cells. (c) Human bone marrow cells. (d) Granule cells in the mouse dentate gyrus. 76
4.4 The architecture of the CNN in this experiment is based on a U-net with eight layers. If all layers are implemented with plane convolutional blocks, the network is the same as the FCNN-A [8]. 79
4.5 (a) Detection results with proximity coding on an example from the DG dataset. (b) Detection results with the proposed repel coding. (c) The proximity coding output. (d) The repel coding output. (e) The Gaussian kernel coding output. (f) The rectangle kernel coding output. The green dots indicate labeled cell centers, and the red dots represent detected cell centers. 83
4.6 Different coding schemes on an example from the Vgg dataset. The green dots indicate labeled cell centers, and the red dots represent detected cell centers. The proposed method distinguishes adjacent cells in most cases. (a) Dot labels. (b) The proximity coding. (c) The repel coding. (d) The Gaussian kernel coding for counting. (e) The rectangle kernel for counting. 84
4.7 (a) and (b) are proximity and repel codings of an example from the Adip dataset. (c) and (d) are proximity and repel codings of an example from the HBM dataset. The green dots indicate labeled cell centers, and the red dots represent detected cell centers. 84

Chapter 1

Introduction

1.1 Background of the thesis

This thesis mainly documents the imaging processing techniques developed for a cross-disciplinary study on status epilepticus (SE) with mouse brains. In collaboration with researchers with different backgrounds, we study a series of intriguing brain activities, such as the spreading process of epilepsy, and how epilepsy interferes with memory formation in the brain. The next two paragraphs provide an overview of the biology and imaging techniques used in this work.

Status epilepticus SE is a significant neurological emergency that afflicts 120,000-160,000 Americans each year with a high overall mortality (17%). The neuronal synaptic plasticity during SE and the underlying molecular mechanisms have been explored extensively. However, neuronal circuits that sustain prolonged seizures have only been explored with low spatial resolution and sparse sampling of brains. Previous studies [10, 11] on neuronal circuits that sustain seizures were carried out using Carbon-14 deoxy-glucose autoradiography. Restricted by the point spread function of the radiation, the resolution limit of the previous methods is 200 microns. A high-resolution activation map that reflects seizure circuits at multiple scales from cells to the whole brain level will help to define pathways of seizure spread during SE. Now with tissue clearing techniques and advanced microscopy, whole mouse brains in SE are

imaged with an in-plane resolution of 2.5 microns in this work. With this unprecedented high resolution of whole brain imaging, we provide neuroscientists with new tools to visualize and understand better the pathways of seizure spreading with SE. These findings help understand the neurobiological basis for the use of drugs for terminating SE.

Imaging techniques As a comprehensive study of brain activity, different imaging techniques and image processing topics are involved in brain reconstruction and analysis. First, confocal microscopy can acquire high resolution images of brain tissues, where single cells are distinguishable. Another great advantage of confocal microscopy is the flexible choice of fluorophores. In our experiments, three fluorophores, tdTomato, Dapi, and NeuN, are used for different purposes. Because some of these fluorophores can only penetrate about 200 microns into the tissue, mouse brains are sliced before imaging. This motivates the image processing task of whole brain reconstruction. Although with the methods proposed in this thesis, 3D mouse brains reconstructed from microscopy data largely preserve the original brain shape, the ground truth information is not available after slicing the brains. As a widely-used non-invasive imaging technique, magnetic resonance imaging (MRI) is able to provide the contours of major structures in these brains before slicing. The complete 3D brain data before slicing also would benefit cross brain registration. In order to acquire higher quality MRI, choosing proper parameters for MRI reconstruction algorithms is important. In this thesis, we discuss a new parameter tuning technique to acquire high quality MRI images.

1.2 Objectives and contribution

The goal of this thesis is designing image processing algorithms to recover the cell activation topology in the mouse brain for studying epilepsy. To achieve this, multiple contributions are made, and these novel works can be applied to other image processing tasks as well. Next, three main contributions of this thesis are summarized.

Contribution 1: Structure-based intensity propagation for 3D mouse brain reconstruction With the high resolution microscopy images of brain tissues, the most straightforward approach to recover complete brain volumes is registering each section of the brain to its previous one. This thesis describes a new 3D registration algorithm for non-overlapping sections that significantly improves structural consistence of the resulting volume. According to the number of layers in each section, the brain reconstruction can be classified into the reconstruction with single-layer sections and the reconstruction with multi-layer sections. In this work, multi-layer sections are imaged with confocal microscopy. How to select the 2D representative of a multi-layer section for registration is a key step in the multi-layer reconstruction. Existing choices of the 2D representatives include surface layers, maximum intensity projections and best selected references [1]. However, these methods either lead to unstable registration results or structure inconsistence in the reconstructed brain. In this thesis, a structure-based intensity propagation method is designed for the robust representation of multilayer sections. Different from existing methods, the proposed structure-based intensity propagation preserves the structure trend in each multi-layer section during the whole brain reconstruction. Experiments are conducted on 367 multilayer sections from 20 mouse brains. Since our work aims at registering multilayer sections to each other, the edge-based tensor voting [12] is used as the structure consistence measurement for evaluation. The average reconstruction quality measured by the structure consistence index increases by 45% with the tissue flattening method, and 29% further with the structure-based intensity propagation.

Contribution 2: Comparison-Based image quality assessment for selecting MRI reconstruction parameters MRI is able to image the complete mouse brain volumes without slicing, but at a much lower resolution compared with microscopy. High quality MRI is helpful to guide the reconstruction with the microscopy data and align different mouse brains into the same 3D coordinate [13]. To reconstruct high quality MRI in the spatial domain from the frequency domain signals, MRI reconstruction algorithms rely on the proper choice of parameter sets. To achieve this, we propose a comparison-based image quality assessment

(IQA) method. This method is a new way to assess image quality that enables optimal tuning of denoising and reconstruction algorithms. The new comparison-based framework parallels full reference image quality assessment (FR-IQA) by requiring two input images and resembles no reference IQA (NR-IQA) by not using the original image. As a result, the new comparison-based approach has more application scenarios than FR-IQA does, and takes greater advantage of the accessible information than the traditional single-input NR-IQA does. Experimental results show that the proposed method outperforms other NR-IQA methods for parameter selection by comparing with a widely-used FR-IQA method, structural similarity index method [14].

Contribution 3: Automatic cell counting in the dentate gyrus in the mouse brain With genetically engineered mice, the activated granule cells in the dentate gryus are highlighted in the near-red channel under confocal microscopy. We propose a new coding scheme of raw center labels to accurately and automatically detect all the activated cells using deep convolutional neural networks (CNNs). We demonstrate that trained with the proposed coding scheme, the CNN prediction is more accurate. The most compelling advantage of the proposed coding scheme is the ability to distinguish neighboring cells in crowded regions. Cell counting and detection experiments are conducted for five coding schemes on four types of cells and two network architectures. The proposed coding scheme improves the counting accuracy with the widely-used Gaussian and rectangle kernels up to 12%, and also improves the F1 score of [15] detection with the common proximity coding up to 14%.

1.3 Dissertation outline

Chapter 2 first reviews the reconstruction pipeline and then introduces the proposed tissue flattening and structure-based intensity propagation for mouse brain reconstruction. Experiments on both section-to-section registration and whole brain reconstruction are conducted to test the performance of the proposed methods. The auxiliary imaging modality, MRI, is discussed in Chapter 3. The parameter selection approach is proposed to reconstruct high quality MRI.

With the reconstructed 3D mouse brains, Chapter 4 documents the cell detection framework and introduces the proposed coding scheme for cell detection. At last, Chapter 5 reviews the work in this thesis and discusses the future works can be extended from this thesis.

Chapter 2

3D Mouse Brain Reconstruction [1]

As one of the most sophisticated organs, brains consist of intricate structures and diversified cells. Although other imaging techniques, such as magnetic resonance imaging (MRI) and computed tomography (CT), are valuable noninvasive imaging approaches, microscopy remains irreplaceable for the SE study because of its high resolution [13, 16–19] and the flexible choice of stains [20–23]. However, a whole brain usually is sectioned into slices for high resolution microscopy imaging. Reconstructed 3D virtual brains benefit brain-related research in three aspects: visualization, anatomical labeling, and 3D measurement.

One of the main goals of the Allen Brain Atlas (ABA) project is allowing researchers to make comprehensive queries about gene expression patterns in 3D with reconstructed brains [20]. Visualizing the neuron morphology in reconstructed brains relaxes the view angle limit caused by physical sectioning, and is crucial to understand the communication of neural signals [24, 25]. Atlas-matching provides critical anatomical information of the imaged brain activities. However, the deviation between the actual cutting plane and the standard cutting planes would compromise the registration accuracy. With 3D reconstruction of brain volumes, a 2D cross-section view from any angle is possible and thus leads to more accurate anatomy labeling [13, 26, 27]. Because of the importance of 3D brain volumes, many brain atlases provide

[1] ©2018 IEEE. Reprinted with permission from Haoyi Liang, Natalia Dabrowska, Jaideep Kapur and Daniel Weller, "Structure-based Intensity Propagation for 3D Brain Reconstruction with Multilayer Section Microscopy", IEEE transactions on medical imaging, October, 2018

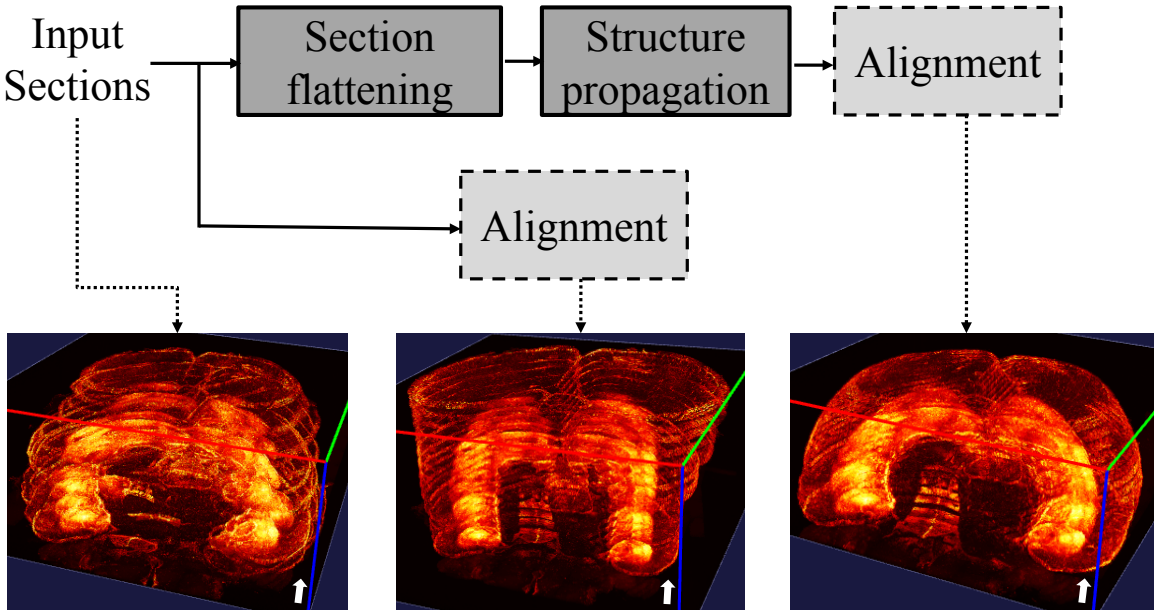


Figure 2.1: The top row is the reconstruction pipeline adopted in this work. Section flattening and structure propagation are the two techniques proposed to improve the reconstruction quality. The three figures at the bottom row are the stacked raw sections, the reconstructed brain without the proposed techniques, and the reconstructed brain with the proposed techniques. The arrows in the three bottom figures indicate the ventral-dorsal direction.

3D brain reference and gene expression patterns, such as Allen brain atlas [20], Hof's brain atlas [28] and Paxinos & Franklin's brain atlas [29]. Accurate 3D measurements for diagnosis and pathology study are also facilitated by the reconstructed brain volumes [24, 30, 31]. The activation pattern of grid cells in the hippocampus plays an important role in memory formation and environment recognition [32–34]. Extending this pattern analysis from 2D to 3D with reconstructed brain volumes is helpful to further understand the underlying mechanism [32]. For this work, brain tissues collected from mice expressing fluorescent tdTomato and co-labeled with the neuronal marker NeuN, are imaged. The reconstructed brain volumes provide details about seizure spreading at multiple scales: from circuits, systems, to the whole brain. The seizure spreading pathways are helpful for neuroscientists to find the neurobiological basis for the use of drugs to terminate status epilepticus at an early stage.

Existing brain reconstruction works can be classified into two approaches by the way that specimens are sliced and imaged [23]: reconstruction with single-layer sections [1, 35–38]

and reconstruction with multilayer sections [23, 24, 39–42]. Imaging and reconstruction with multilayer sections are getting more attention recently because of the less labor required for the tissue preparation [23, 39] and the greater robustness to distortions such as warping and tearing.

For this work, multilayer sections are the raw data for whole brain reconstruction. Our previous work in [43] describes an automatic tissue flattening method to remove the warping artifacts along the cutting axis in multilayer sections. The reconstructed brains are more compact with flattened sections, but the structure inconsistence and the cylindrical artifact [39] still exist as shown in Fig. 2.1.

Numerous methods are proposed to improve the structure consistence and to eliminate the cylindrical artifact for brain reconstruction. Extra equipment, such as high-resolution MRI scanners [35, 36], and algorithms imposing smoothness constraints [1, 37, 38], are two common approaches. Among these methods, the best reference selection (BRS) [1] proposes an interesting idea: registering a section to its nearest key section, rather than directly to the previous one. The key sections are selected by criteria such as contrast, entropy and intensity [1]. A similar idea is also explored in [36].

Inspired by the previous works [1, 36], structure-based intensity propagation (SIP) is proposed to improve the structure consistence of brain reconstruction. Different from BRS [1], the references used for registration are created by propagation, rather than selected as in BRS. Fig. 2.1 illustrates the overall work flow. This work obtains robust representatives of multilayer sections with modules of tissue flattening and structure propagation. By feeding the alignment module with better representatives of multilayer sections, SIP improves the reconstruction quality in two aspects. First, the structure transition among sections is smoother because the structural trend within each 3D section is accurately reflected in propagated surfaces. Secondly, the flip detection is more reliable. Since brain structures are highly symmetric on coronal and horizontal planes, automatic flip detection is challenging. The proposed structure-based intensity propagation facilitates flip detection by enhancing the signal intensity and structure contrast.

The rest of this chapter is organized as follows. Section 2.1 summarizes existing brain reconstruction works, and then introduces the brain registration pipeline adopted in this work. Section 2.2 reviews the tissue flattening method in [43], and then elaborates on the proposed structure-based intensity propagation. In Section 2.3, experiments demonstrate how the proposed methods improve the quality of the brain reconstruction. At last, Section 4.4 reviews the novelty and experimental verification of the proposed methods, and discusses the further work of brain reconstruction.

2.1 Reconstruction backgrounds

2.1.1 Existing brain reconstruction works

As mentioned in Section 4, brain reconstruction can be classified into two approaches according to the number of layers in each section. Although the specimen preparation procedures are different for single-layer sections and multilayer sections, the following reconstruction shares similar frameworks and faces some common challenges.

Compared with multilayer sections, each single-layer section is easier to prepare and image. The typical thickness of single-layer sections is between 10 to 100 microns [1, 35–38], and each physical section is imaged with a single 2D image. However, the preparation of a high-quality single-layer section is challenging and time-consuming [1, 36, 38, 39, 42]. In [36], exhaustive pre-processing, such as visual quality inspection and mask detection, is required before registration. After imaging these sections, the major challenge is maintaining the smoothness of the reconstructed brains. Accurate section-to-section registrations cannot avoid the cylindrical artifacts due to the aperture problem [39, 44]: minor registration errors accumulate along the z-direction and compromise the overall shape of the reconstructed volume. In [13, 35, 36], MRI scanners image the overall shapes of brains before slicing. Some algorithms are designed by optimizing the registration order [1, 36] or applying smoothness constraints [37, 38].

The brain reconstruction with multilayer sections [23,24,38,40] is getting additional attention because of two techniques: tissue clearing and high resolution confocal microscopy. Lipids are removed with tissue clearing [19, 24, 45–48], and thus the light emitted by fluorescent markers undergoes less scattering. Advanced confocal microscopy enables specimens to be imaged at different depth without the compromise of resolution. However, reconstruction with multilayer sections has its own unique challenges. As one of the important steps during specimen preparation, the tissue clearing turns specimens both transparent and warped. How to select proper 2D representatives of multilayer sections for registration is difficult. For single-layer sections, one physical section is directly used for registration, while the surface layers in a raw multilayer section are not good choices for registration. Due to the distortions during tissue preparation and the partial volume effect during imaging, the surface layers do not precisely reflect the intra-section structural trends. To solve this challenge, existing reconstruction methods with multilayer sections either rely on manually labeled information or highly customized equipment. Manually selected key points on tissue surfaces are used as 2D representatives of multilayer sections for registration [23, 24, 41, 42, 49]. Traced neuron features in an interactive GUI are used for 3D volume reconstruction in [24, 41, 49]. However, this laborious approach is not scalable. Alternatively, customized equipment is assembled to facilitate the reconstruction process [24, 39]. For example, staining is done before slicing in [24], and thus many organic immunohistochemical stains cannot be used due to the limits of the penetration depth.

At last, although light-sheet microscopy together with tissue clearing techniques is technically possible to image the whole brain without any slicing, such whole brain imaging is constrained by the choice of stains and the resolution [31]. Many antibodies and stains do not penetrate through the whole intact brain. Also, the resolution of a typical light-sheet microscope is not enough to resolve individual cells, while in-plane resolution under 1 μm is common for point scanning microscopes such as two-photon or confocal microscopy.

2.1.2 Alignment pipeline in this work

As stated in Section 4, the brain reconstruction pipelines for single-layer and multilayer sections are all based on section-to-section registration. In this work, a three-step registration pipeline is implemented with rough alignment, affine transformation and non-rigid registration. Our proposed structure correction methods do not rely on particular registration methods. The works [50–52] are selected because of their robust performance and public implementations. The first step, rough alignment, only takes translation and rotation into consideration, and both non-flipped and flipped versions of the section to be registered are evaluated. The parameter sets that achieve highest correlation scores [50] in the non-flipped and flipped versions are saved. The second step, affine registration, maximizes the mutual information [51] of the outputs from rough alignment. The flip status is decided after affine registration: the status that achieves higher mutual information index is selected. Unlike single-layer sections, only one combination of the four flip statuses of two adjacent multilayer sections achieves the highest mutual information. For single-layer sections, if two adjacent sections both are incorrectly flipped, the registration cost is the same as the correct flip situation. However, for multilayer sections, because the top surface and the bottom surface are different, four different flip combinations of two adjacent sections lead to four different registration costs. The last step is non-rigid registration that minimizes the residual complexity [52] between two input images. The resolutions also gradually increase from rough alignment, affine registration to non-rigid registration. Such hierarchical registration approaches are common in brain reconstruction works for the purpose of computation time and registration accuracy [30, 36]. Readers are referred to [43] for details about the implementation.

To apply the registration pipeline for brain reconstruction with multilayer sections, a representative has to be selected. In the next section, the proposed tissue flattening and structure-based intensity propagation provide accurate representatives for the multilayer section registration.

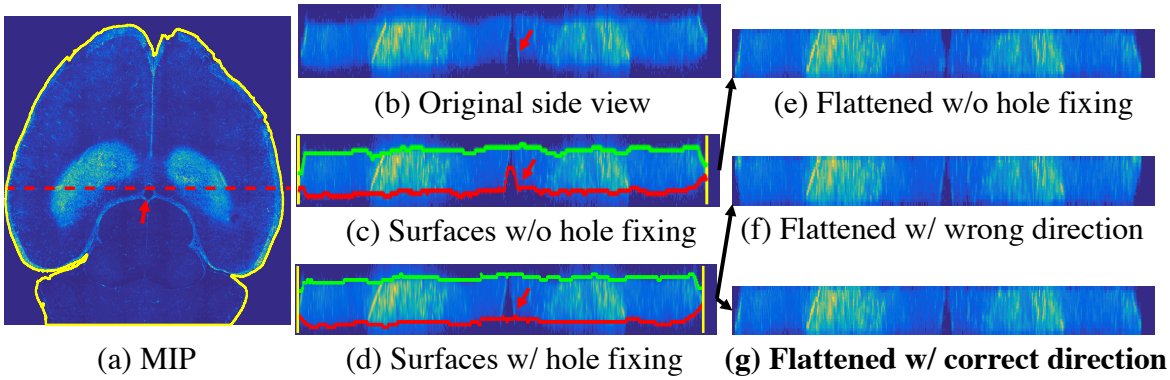


Figure 2.2: Intermediate results of tissue flattening. (a) is the maximum intensity projection of one section sliced on the horizontal plane. The dashed line in (a) indicates the positions of side views. (b) is the side view of the raw section. (c) shows the detected surfaces before hole fixing. If the tissue is projected onto the bottom surfaces in (c), the structure around the ventricle is changed as shown in (e). (d) shows the detected surface after hole fixing. However, with incorrect projection direction, structures at the boundaries are altered as shown in (f). With correctly detected surfaces and the projection direction, (g) shows the flattened tissue from (b).

2.2 Proposed structure correction methods

The proposed structure correction for brain reconstruction contains two parts: tissue flattening [43] and structure-based intensity propagation. Before tissue flattening, the structures in most layers of a multilayer section are distorted by the unevenness on z-direction. After tissue flattening, the warping artifacts on the z-direction are removed, and the surface layers show the general contours and major structures. However, the structures presented on the surface layers after flattening are not in accord with the intra-section structural trend, and the signal intensity is usually weak. Structure-based intensity propagation is designed to overcome these limitations in surface layers for accurate registration and flip detection.

2.2.1 Flattening

The tissue clearing process not only removes the lipids from the specimens, but also slightly warps the specimens. In order to process large numbers of tissues, an automatic tissue flattening method [43] is proposed. Fig. 2.2 illustrates the key intermediate results of tissue flattening. The warping distortion exists in the raw section as shown in Fig.2.2 (b). Fig. 2.2 (c) shows the

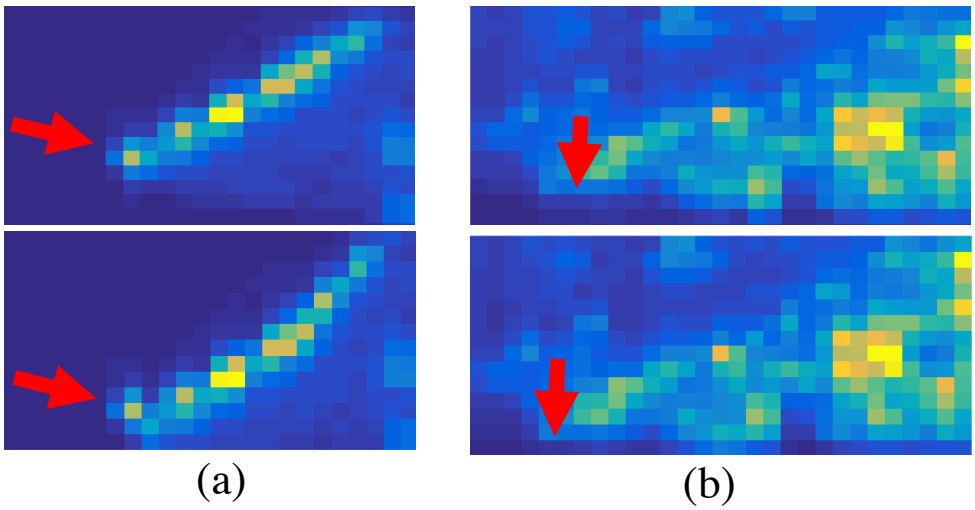


Figure 2.3: First and second rows are the cross-section views of a section before and after flattening. In (a), tissue boundaries are not accurately reflected on the bottom layer because of the upward tilt. In (b), the size of the hippocampus grows from the top to the bottom, but this structural trend is not correctly reflected on the bottom layer.

detected surface layers with adaptive thresholds. By assuming that the distance between the top and the bottom layers is constant, the hump on the bottom surface is removed in Fig. 2.2 (d) after hole fixing. At last, the projection direction is decided by the total variation along the surface layer rims. The surface with flatter rim is selected as the layer onto which we project the rest of the section. Details about the tissue flattening can be found in [43].

Tissue flattening improves the quality of reconstructed brains in two aspects. First, surface layers, rather than maximum intensity propagation (MIP), of the flattened tissues can be used for registration. The overall shape of reconstructed brains is more natural [43]. Secondly, the wide gaps among sections are removed. However, minor structure inconsistence, such as the zig-zag contour on the brain outline, still exists [43]. Another limitation of tissue flattening is that structure contrast and signal intensity are weak on the surface layers. To overcome these drawbacks, structure-based intensity propagation (SIP) is proposed in the next section.

2.2.2 Structure-based intensity propagation

Once obtaining the flattened sections, a straightforward reconstruction scheme registers the surfaces from neighboring sections to each other [23, 24, 39–42]. Fig. 2.3 illustrates why this approach is not suitable for automatic reconstruction. In Fig. 2.3 (a), the flattened patch on the bottom preserves the upwards tilt at the end. Surfaces with such tilts cannot accurately delineate the boundaries, and are the source of the zig-zag artifacts in the reconstructed brains [43]. In Fig. 2.3 (b), the signal strength on the surface layer is weak. One reason for this phenomenon is partial volume imaging on the surfaces. As representatives of a multilayer section, these surfaces with weak signal strength and contrast lead to unstable registration or even flip errors due to mirror symmetry [43]. In order to improve the signal intensity and contrast with intra-section information, one straightforward idea is applying the median or mean filter along the z-direction. The defect of this scheme is that the structure changing along the z-axis is neglected. However, if these filters are applied along the structure directions in the sections, both the aims of signal enhancement and structure preservation are achieved. The proposed structure-based intensity propagation implements the idea of the structure-oriented median filter.

Alg. 1 summarizes the structure-based intensity propagation. The inputs of Alg. 1 are a 3D stack of size $M \times M \times D$ with isotropic resolution and a patch size, N. The width and the height (the first two dimensions) of the input 3D image stack I do not have to be equal. A square 3D stack is used for the simplicity of notation in Alg. 1. If the raw data acquired by confocal microscopy is of anisotropic resolution, the raw data should be scaled so that the 3D stack is of isotropic resolution before fed to Alg. 1. In our implementation, the scaling is done with cubic interpolation [53]. The outputs of Alg. 1 are two 2D images of size $M \times M$, L^{top} and L^{bottom}. These two 2D images are the propagated surfaces to align two adjacent multilayer sections. The steps in Alg. 1 are classified into two parts. The first part, steps 1-6, is structure estimation. The second part, steps 7-10, is surface layer estimation. Fig. 2.4 illustrates key intermediate results in Alg. 1. Fig. 2.4 (a) is a rendered multilayer section in 3D. The thickness

Algorithm 1 Structure-based Intensity Propagation

Inputs:

I: 3D image stack of size $M \times M \times D$

N: patch size

Outputs:

L^{top}: Propagated top surface of size $M \times M$

L^{bottom}: Propagated bottom surface of size $M \times M$

Structure Estimation:

for every p on the imaging plane of size $M \times M$ **do**

1. $G(p) = [g_x(p)\ g_y(p)\ g_z(p)]$ ▷ gradient matrix
2. $U(p)S(p)V(p)^T = SVD(G(p))$ ▷ SVD
3. $[v_1(p)\ v_2(p)\ v_3(p)] = V(p)$
4. $h(p) = \underset{h(p)\perp v_1(p), \|h(p)\|_2^2=1}{\operatorname{argmin}} h_z(p)$ ▷ structure vector
5. $F_x(p) = \frac{h_x(p)}{h_z(p)}$
6. $F_y(p) = \frac{h_y(p)}{h_z(p)}$

end for

Surface Layer Estimation:

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

7. $I_i^{top} = I_i(p_x + F_x \cdot (i-1), p_y + F_y \cdot (i-1))$
8. $I_i^{bottom} = I_i(p_x - F_x \cdot (D-i), p_y - F_y \cdot (D-i))$

end for

9. $L^{top} = median_z(I^{top})$
10. $L^{bottom} = median_z(I^{bottom})$

of all the multilayer sections in our experiments is 200 μm, which is also the typical penetration depth of many immunohistochemical stains. The in-plane range of the multilayer section in Fig. 2.4 (a) is 9569×9669 μm^2. Fig. 2.4 (b) and (c) are the outputs of the first module in Alg. 1, F_x and F_y. The estimated structure maps, F_x and F_y, are two 2D maps of size $M \times M$. The expansion direction and magnitude for each in-plane position are characterized in Fig. 2.4 (b) and (c). The MIP of the raw section in Fig. 2.4 (d) cannot differentiate the top and the bottom surfaces, and thus leads to the cylindrical artifact in the reconstructed brain. The bottom surface of the flattened section in Fig. 2.4 (e) has weak signal contrast and loses some asymmetric information that is key to flip detection. The propagated bottom surface in Fig. 2.4 (f) retains both the signal strength and structures. The detailed computational

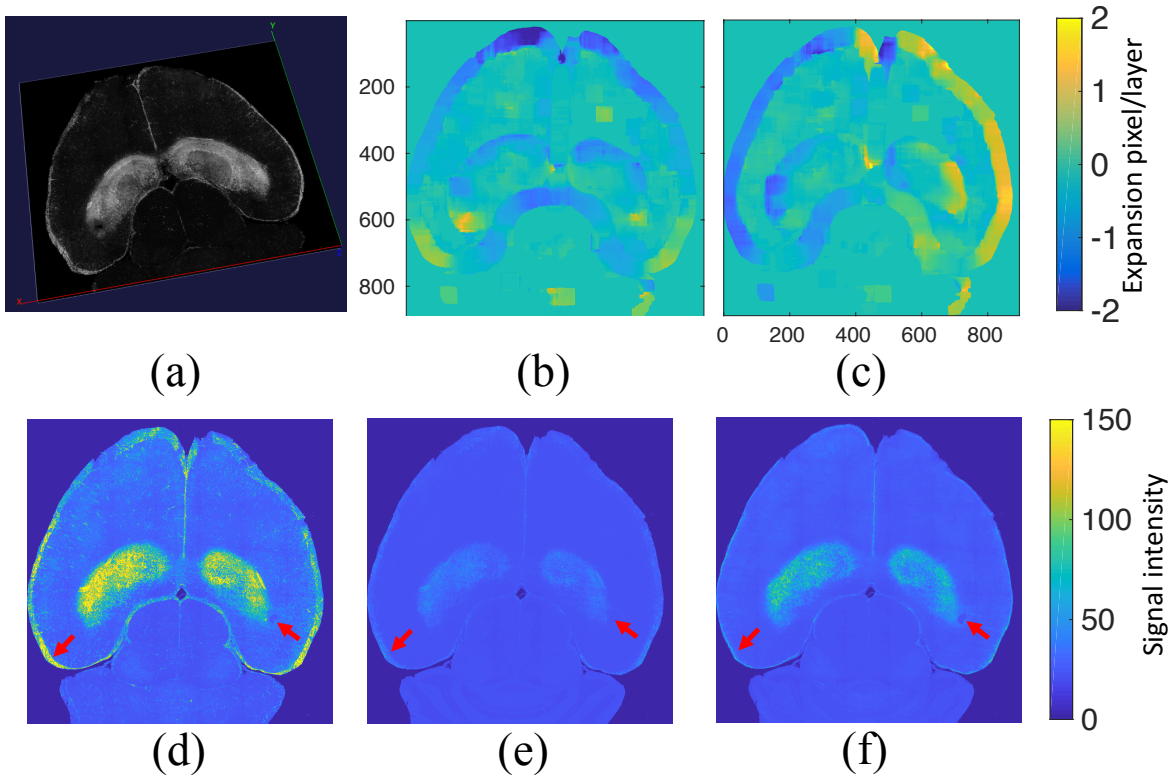


Figure 2.4: (a) A rendered multilayer section of size $9569 \times 9569 \times 200$ μm^3 in 3D. This section is sliced on the horizontal plane. From the top to the bottom (on z-direction), the brain and the hippocampus expand. (b) Estimated structure map on the direction of top-bottom in the imaging-plane. (c) Estimated structure map the direction of left-right in the imaging-plane. MIP, the bottom surface of flattened section and the propagated bottom surface are (d), (e) and (f) respectively.

complexity analysis and an optimized implementation of Alg. 1 are provided in the next section. The structure-based intensity propagation is introduced with the implementation in Alg. 1 for its clarity, and the computational complexity of Alg. 1 is $O(M^6D^2)$. The computational complexity of the optimized implementation in the supplementary material is $O(M^2Dlog(MD))$. The following two paragraphs explain the two modules, structure estimation and surface layer estimation, in detail.

The outputs of structure estimation are two 2D maps showing the structure changes on two in-plane directions. One assumption of the structure estimation is that the main structures in cross-section views are linear because each multilayer section is thin, as shown in Fig. 2.4 (a).

The typical dimension of a multilayer section in our experiment is $10000 \times 10000 \times 200\ \mu m^3$, and thus the depth-to-length ratio is about 1/50. The first step in Alg. 1 is constructing the gradient matrix, $G(p)$, for each position, p, on the imaging plane. The gradient matrix, $G(p)$, of size $N^2D \times 3$ is composed of three vectors that correspond to gradients along three directions. In our implementation, the gradients at each pixel are computed by a central difference on a $3 \times 3 \times 3$ neighborhood. The singular value decomposition (SVD) of the gradient matrix, $G(p)$, is defined as

$$G = USV^T = U \begin{bmatrix} s_1 & 0 & 0 \\ 0 & s_2 & 0 \\ 0 & 0 & s_3 \end{bmatrix} \begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix}^T, \tag{2.1}$$

where U and V are both orthonormal matrices. Vector v_1 is of size 3×1 and corresponds to the dominant direction of the local gradient; v_2 and v_3 are orthogonal to v_1. The three singular values, s_1, s_2 and s_3, represent the amount of gradient variation on the three corresponding singular vectors v_1, v_2 and v_3, and $s_1 > s_2 > s_3$. With the definition of SVD, the dominant intensity changing direction, v_1, is always perpendicular to the main structure direction [54]. Next, the structure direction, h, is estimated by selecting the unit vector that is perpendicular to v_1 and has the largest z-direction descent. This process is reflected in Step 4 in Alg. 1. The property of being perpendicular to v_1 guarantees h lies on the structure surface. Among these vectors, the one that has the largest downward z-direction component is selected as h. The three components in h are h_x, h_y and h_z. The minimization operation in Step 4 of Alg.1 specifies the downward direction of the unit vectors have negative signs. The property of having the largest z-direction component can be interpreted as having the smallest in-plane component because of the first property of being a unit vector. Therefore, when propagating an intra-section voxel to the positions of surface layers, the propagation path with direction h stays on the structure, and has the smallest in-plane displacement. After obtaining the structure direction h, the structure change on xy-planes is calculated in Step 5 and 6 in Alg. 1. The values in $F_x(p)$ and $F_y(p)$ indicate the displacement of a pixel at p traveling to its next layer. In other words, the units of $F_x(p)$ and $F_y(p)$ are pixel/layer. Fig. 2.4 (b) and (c) show the structure change maps.

The second part of Alg. 1 estimates the propagated surfaces with the structure change maps. In Step 7 and Step 8 in Alg. 1, every layer is transformed to positions of the top layer and the bottom layer. The deformation fields, $(p_x + F_x \cdot (i-1), p_y + F_y \cdot (i-1))$ and $(p_x - F_x \cdot (D-i), p_y - F_y \cdot (D-i))$, are calculated with structure changes and the distances to surface layers. Two 3D stacks, I^{top} and I^{bottom}, are created with reference to the top and bottom layers respectively. Ideally, only vertical structures should be presented in the side views of I^{top} and I^{bottom}. At last, median values in the z-direction are taken from I^{top} and I^{bottom} as the propagated surface layers. Fig. 2.4 (f) shows the propagated bottom surface of one multilayer section.

To better interpret how the proposed structure-based intensity propagation works, a 2D example is shown in Fig. 2.5. The method described in Alg. 1 can be easily applied to 2D cases. The difference is that the surface layers in Alg. 1 are 1D lines in 2D cases. Fig. 2.5 (a) is a 2D patch from the cross-section view of the raw data. The most significant structure in Fig. 2.5 (a) is the boundary of a mouse brain. Two defects in the raw data are presented. First, the position of the boundary on the top line is not accurate. Second, the intensity of the bottom line is weak. Fig. 2.5 (d) plots the gradients in Fig. 2.5 (a) and the fitted ellipse with SVD. The orientation in Fig. 2.5 (d) reflects the direction of the significant structure. The dashed arrow in Fig. 2.5 (a) is equivalent to the normal vector, v_1, in Alg. 1, and the solid arrow is equivalent to the unit structure vector, h, in Alg. 1. Computing the solid arrow in Fig. 2.5 (a) fulfills the structure estimation in Alg. 1, and the next step is the surface layer estimation. Applying the displacement fields to each layer in Alg. 1 is simply shifting each line in the 2D cases. Fig. 2.5 (b) and (c) are the propagated stacks whose counterparts in Alg. 1 are I^{top} and I^{bottom}. At last, applying a median filter to Fig. 2.5 (b) and (c), the propagated top line and the propagated bottom line are obtained. Fig. 2.5 (e) compares the original surface lines and the propagated surface lines. The boundary position is more precisely reflected in the propagated top line, and is more distinct in the propagated bottom line.

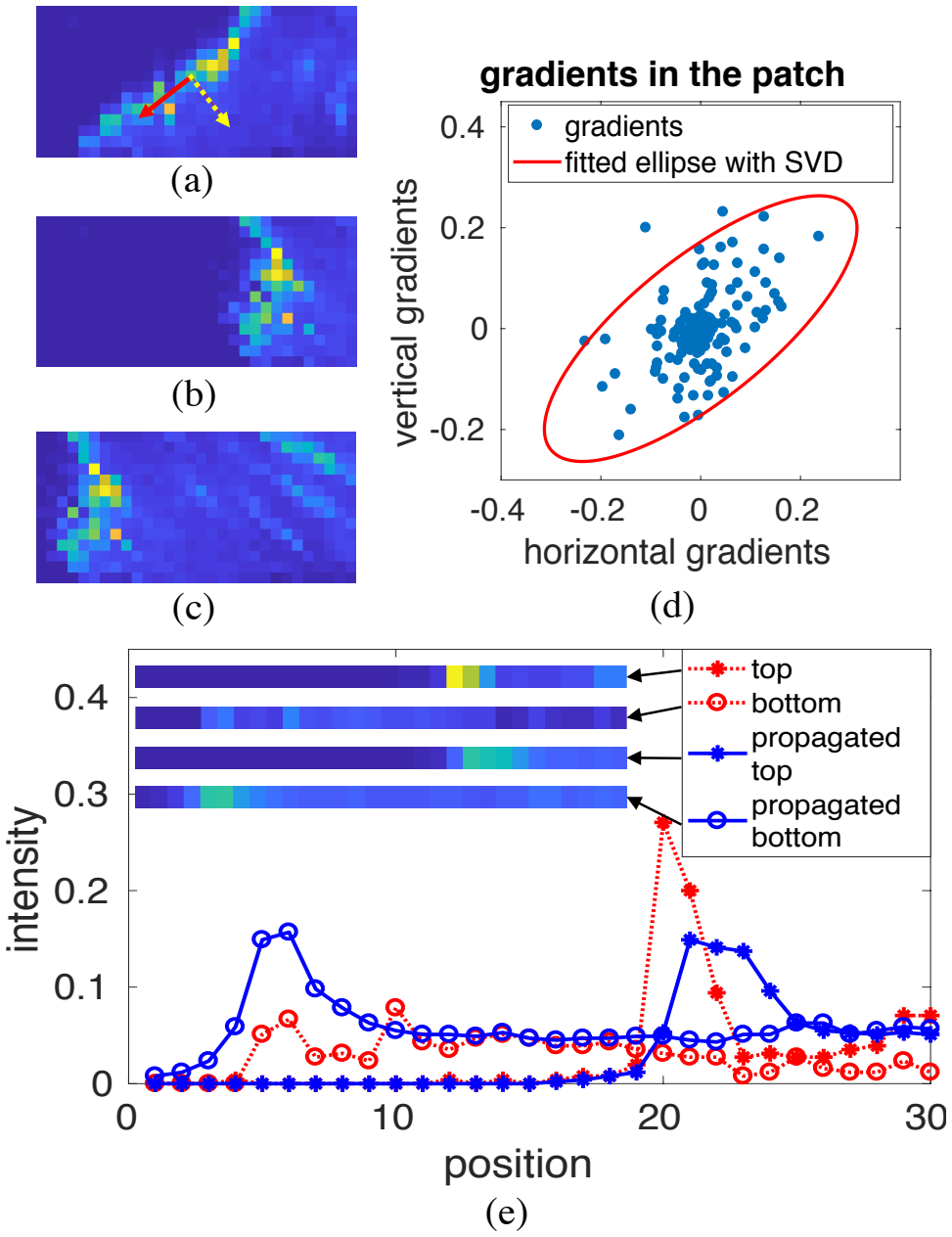


Figure 2.5: An illustration of the structure-based intensity propagation on the 2D case. (a) is a cross-section view of the raw data. The solid arrow indicates the structure direction, and the dashed arrow corresponds to the normal direction. (b) and (c) are the stacks after transforming each lines to the surface positions, equivalent to I^{top} and I^{bottom} in Alg. 1. (d) plots the gradients in (a) and illustrates the SVD operation. The propagated surface lines reflect the positions of boundaries more accurately.

Table 2.1: Run time comparison of two SIP implementations

Image stack size ($M \times M \times D$)	Computational complexity	$256 \times 256 \times 5$	$512 \times 512 \times 10$	$1024 \times 1024 \times 20$
Patch-based implementation (s)	$O(M^6D^2)$	7.64	167.71	128813.72
Optimized implementation (s)	$O(M^2Dlog(MD))$	0.21	1.05	10.90
Speed up ratio	–	36.38	159.72	11817.77

2.2.3 Implementations of structure-based intensity propagation

Structure-based intensity propagation (SIP) is the major contribution of our work. In the main draft, SIP is introduced with a patch-based implementation for the purpose of clarity. However, the patch-based implementation of SIP is slow in practice because considerable computation is redundant when patches overlap. An optimized implementation of SIP is necessary for brain reconstruction with high resolution.

In this section, we first review the patch-based implementation of the proposed SIP in Section 2.2.3. The optimized implementation is introduced in Section 2.2.3. At last, an experiment is conducted to compare run times in Section 2.3.

Structure-based Intensity Propagation: Patch-based implementation

Alg. 1 summarizes steps in SIP with the patch-based implementation. The inputs of Alg. 1 are two: a multilayer stack of $M \times M \times D$ voxels with isotropic resolution, and a patch size, N. The multilayer stack has D layers, and each layer contains $M \times M$ pixels. If the raw data acquired by confocal microscopy is of anisotropic resolution, the raw data should be scaled so that the in-plane resolution and the cross-plane resolution are the same. The outputs of Alg. 1 are two 2D images both of size $M \times M$ pixels, the propagated top surface L^{top} and the propagated bottom surface L^{bottom}. The steps in Alg. 1 can be generally classified into two parts. The first part, steps $1-6$, is structure estimation. The second part, steps $7-10$, is surface layer estimation.

Steps $1-6$ are executed M^2 times to acquire the structure map F_x and F_y for all the positions on the imaging plane. Both F_x and F_y are of size $M \times M$. Step 1 constructs the gradient matrix $G(p)$ of size $N^2D \times 3$. The vectors $g_x(p)$, $g_y(p)$ and $g_z(p)$ are the gradients of voxels in

the cubic patch centered at p. The computational complexity of Step 1 is therefore $O(N^2D)$. Step 2 calculates the singular vectors of the gradient matrix, $G(p)$. With the definition of SVD, the dominant intensity changing direction, v_1, is always perpendicular to the main structure direction [55]. The computational complexity of Step 2 is decided by the first dimension size of $G(p)$, and is $O(N^4D^2)$. Step 3 simply extracts the three single vectors from $V(p)$. Next, Step 4 determines the structure direction, h, by selecting a unit vector that is perpendicular to v_1 and has the largest z-direction descent. After obtaining the structure direction h, the structure change is calculated in Step 5 and Step 6. The computational complexities of steps $3-6$ are $O(1)$. Considering the M^2 loops for steps $1-6$, the computational complexity for the structure estimation is $O(M^2N^4D^2)$.

Steps $7-10$ estimate the propagated surfaces, L^{top} and L^{bottom}, with the structure change maps. In Step 7 and Step 8, every layer is transformed to positions of the top layer and the bottom layer. If nearest interpolation is adopted, the computational complexity is $O(M^2)$ for Step 7 and Step 8. Considering the D loops, the overall computational complexity of Step 7 and Step 8 is $O(M^2D)$. At last, median values on the z-direction are taken from I^{top} and I^{bottom}. The computational complexity of a median-finding algorithm is linear, and the median-finding operation has to be performed for each position on the image plane. Therefore, the computational complexities of Step 9 and Step 10 are both $O(M^2D)$. The overall computational complexity for surface layer estimation, steps $7-10$, is $O(M^2D)$. Considering the filter size N is proportional to the image size M, **the overall computational complexity of Alg. 1 is $\mathbf{O(M^6D^2)}$**.

During the computational complexity analysis of Alg. 1, we can see that the computation cost is dominated by Step 2 in the loop of structure estimation. In the next section, an optimized implementation is introduced to avoid this loop in the part of structure estimation.

Structure-based Intensity Propagation: Optimized implementation

In Alg. 1, the structure changes $F_x(p)$ and $F_y(p)$ are estimated with the SVD in a loop of M^2 times. In order to bypass the SVD, a matrix $Q(p)$ is constructed as

$$\begin{aligned} Q(p) &= G^T(p)G(p) \\ &= V(p)S^T(p)U^T(p)U(p)S(p)V^T(p) \\ &= V(p)S^T(p)S(p)V^T(p) \end{aligned} \tag{2.2}$$

The matrix $Q(p)$ is of sized $3 * 3$, and the eigenvector of $Q(p)$ with the largest eigenvalue is the same as $v_1(p)$ in Alg. 1. The matrix $Q(p)$ can be also written as

$$\begin{aligned} Q(p) &= G^T(p)G(p) \\ &= \begin{bmatrix} g_x^T(p) \\ g_y^T(p) \\ g_z^T(p) \end{bmatrix} [g_x(p)g_y(p)g_z(p)] \\ &= \begin{bmatrix} g_x^T(p)g_x(p) & g_x^T(p)g_y(p) & g_x^T(p)g_z(p) \\ g_y^T(p)g_x(p) & g_y^T(p)g_y(p) & g_y^T(p)g_z(p) \\ g_z^T(p)g_x(p) & g_z^T(p)g_y(p) & g_z^T(p)g_z(p) \end{bmatrix} \\ &= \begin{bmatrix} E_{xx}(p) & E_{xy}(p) & E_{xz}(p) \\ E_{xy}(p) & E_{yy}(p) & E_{yz}(p) \\ E_{xz}(p) & E_{yz}(p) & E_{zz}(p) \end{bmatrix} \end{aligned} \tag{2.3}$$

The symmetric matrix $Q(p)$ is of size 3×3, and defined by six distinct numbers: $E_{xx}(p)$, $E_{yy}(p)$, $E_{zz}(p)$, $E_{xy}(p)$, $E_{xz}(p)$, $E_{yz}(p)$. It should be noticed that matrices, E_{xx}, E_{yy}, E_{zz}, E_{xy}, E_{xz} and E_{yz}, are all of size $M \times M$, and Q is a tensor of size $M \times M \times 3 \times 3$. The QR algorithm with Householder reduction [56] is adopted to solve $v_1(p)$ from the $Q(p)$.

The optimized implementation of SIP, Alg. 2, also contains two parts: structure estimation and surface layer estimation. The changes in Alg. 2 are reflected in steps 1-4.

In Alg. 2, Step 1 calculates the gradient matrices, G_x, G_y and G_z, of the whole input image stack I. These three gradient matrices are of size $M \times M \times D$. Step 2 and Step 3 calculate the six matrices that define Q. Again, it should be noticed that E_{xx}, E_{yy}, E_{zz}, E_{xy}, E_{xz} and E_{yz} are 2D matrices of size $M \times M$. Step 2 creates a 3D average filter of size $N \times N \times D$. The $\cdot$ in Step 3 stands for the element-wise production, and $*$ stands for the convolution. In practice, the convolution is done in the Fourier domain with the fast Fourier transformation. Step 2 and Step 3 are the implementation of Eqn. 2.3. Step 4 corresponds to the decomposition in Eqn. 2.2. The overall computational complexity of Alg. 2 is $O(M^2 D \log MD)$ determined by Step 3. **The overall computational complexity of Alg. 2 is $\mathbf{O(M^2 D \log(MD))}$.**

2.3 Experiments

In this section, the data acquisition procedure and the evaluation criteria are first reviewed. Three experiments, section-to-section registration, flip detection and whole brain reconstruction, are conducted to illustrate how the proposed methods improve upon conventional reconstruction methods.

Four reconstruction approaches with different representatives of multilayer sections are evaluated. The registration pipeline introduced in Section 2.1.2 is adopted. Without tissue flattening and SIP, the registration pipeline with the MIP of each section serves as the baseline method (MIP). After tissue flattening, the top and the bottom layers are used as the representatives for registration in the second version (Surface). As the motivation of the proposed SIP, BRS [1] is evaluated by selecting the best representatives from flattened multilayer sections (BRS [1]). At last, the fourth version includes both tissue flattening and SIP. The propagated top and bottom surfaces are used for registration in the fourth version (SIP). The patch size, N, in Alg. 1 is set as 51 for all the experiments. The raw data is acquired with the resolution of $2.77 \times 2.77 \times 10\ \mu m^3$, and is scaled to $10 \times 10 \times 10\ \mu m^3$ with cubic interpolation [53] before reconstruction.

Algorithm 2 Optimized Implementation

Inputs:

I: 3D image stack of size $M \times M \times D$

N: patch size

Outputs:

L^{top}: Propagated top surface of size $M \times M$

L^{bottom}: Propagated bottom surface of size $M \times M$

Structure Estimation:

1. acquire G_x, G_y and G_z ▷ $O(M^2D)$
2. create average filter: $Filt$ ▷ $O(N^2D)$
3. calculate E_{xx}, E_{yy}, E_{zz}, E_{xy}, E_{xz}, E_{yz} as,
   $E_{xx} = (G_x \cdot G_x) * Filt$,
   $E_{xy} = (G_x \cdot G_y) * Filt$,
   ... ▷ $O(M^2D \log MD)$

for every p on the imaging plane of size $M \times M$ **do**

4. $v_1(p) = QR(E_{xx}(p), E_{yy}(p), \dots)$ ▷ $O(1)$
5. $h(p) = \underset{h(p)\perp v_1(p), \|h(p)\|_2^2=1}{\operatorname{argmin}} h_z(p)$ ▷ $O(1)$
6. $F_x(p) = \frac{h_x(p)}{h_z(p)}$ ▷ $O(1)$
7. $F_y(p) = \frac{h_y(p)}{h_z(p)}$ ▷ $O(1)$

end for

Surface Layer Estimation:

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

8. $I_i^{top} = I_i(p_x + F_x \cdot (i-1), p_y + F_y \cdot (i-1))$ ▷ $O(M^2)$
9. $I_i^{bottom} = I_i(p_x - F_x \cdot (D-i), p_y - F_y \cdot (D-i))$ ▷ $O(M^2)$

end for

10. $L^{top} = median_z(I^{top})$ ▷ $O(M^2D)$
11. $L^{bottom} = median_z(I^{bottom})$ ▷ $O(M^2D)$

Comparison between two implementations

In this section, the run times of two implementations are evaluated with three image stacks of different sizes. The computer configuration has a Intel Core i7-4770 quad core CPU of 3.40GHz and 32 GB RAM. Both implementations are written in MATLAB. From Table 2.1, it is clear that the optimized implementation is much faster on all scales, and the speed advantage is more obvious when the images size is larger.

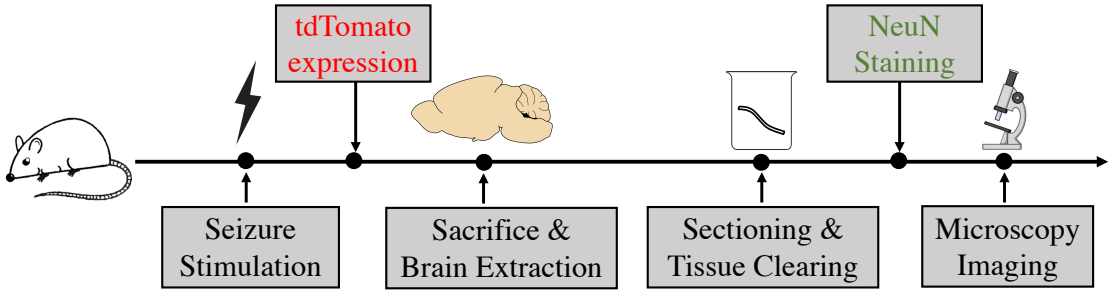


Figure 2.6: Data collection procedure. The red fluorescent markers are expressed during seizure spread while the mice are alive. After sacrificing animals, immunofluorescence staining for NeuN is performed.

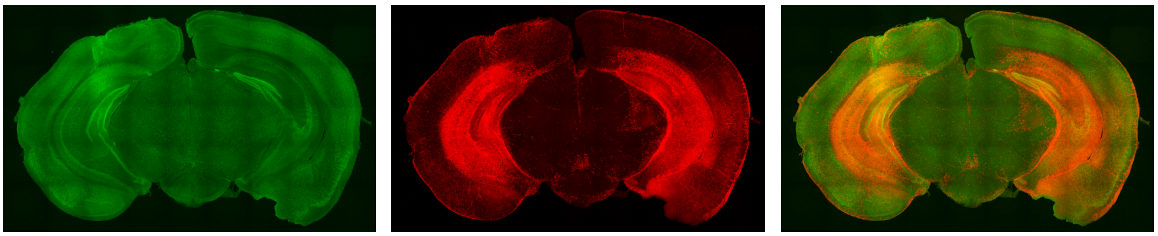


Figure 2.7: (a) NeuN. (b) tdTomato. (c)Aligned two channels. All three images, (a), (b) and (c), are sections on the coronal plane.

2.3.1 Data acquisition

In these experiments, a total of 367 sections from 20 TRAP mice [57] are collected according to the institutional animal care and use committee (IACUC) approved protocol. The data collection procedure is illustrated in Fig. 2.6. The geometric distortions mainly occur at the brain sectioning and tissue clearing steps. The vibratome used for tissue sectioning is the Leica vibratome VT1200, and the vertical deflection of the VT1200 is less than 1 μm [58]. When slicing tissues with the thickness of 200 μm, the vertical deflection is less than 1%. After brain sectioning, lipids are extracted from the tissue to increase the depth of light penetration by tissue clearing [45]. During tissue clearing, slices are incubated in 1% acrylamide, 0.25% VA044 solution with nitrogen under vacuum for 20 minutes, and then transferred to incubator with 37°C. This procedure builds the hydrogel-matrix needed for further clearing step. Slices are then cleared by incubation in 8% SDS buffer at 37°C. The cleared tissue is placed in refractive

index matching solution (RIMS) for imaging. Two channels of fluorescent signal are collected. The red fluorescent protein, tdTomato, is expressed in the mouse brain during seizure spreading. After brain sectioning and clearing, tissues are stained for neuronal marker NeuN (green). Images are obtained using a Zeiss 780 confocal microscope with C-Apochromat objective under 10X magnification. Excitations for the green and the near-red are provided by Argon 488 and DPSS 561 laser lines, and the emission windows are 500-562 nm for the green and 571-624 nm for the near-red. All images are acquired with an optical section separation (z-interval) of 10 μm. Fig. 2.7 shows one section sample. Since the tdTomato channel has higher signal intensity and contrast, the following reconstructions are based on the information from the tdTomato channel.

2.3.2 Evaluation criteria

Evaluating the quality of reconstructed brains is challenging because the lack of ground truth [1, 23, 38]. Visual inspection, manually selected landmarks, and structure smoothness are three commonly used evaluation approaches. Visual inspection of 3D volumes requires extensive human intervention [23, 36]. Similarly, evaluation with manually selected landmarks is not scalable [39]. The third evaluation metric based on the structure smoothness can be further classified into texture-based [59–61] and feature-based [37, 62, 63]. Texture-based metrics are not suitable for reconstruction evaluation with multilayer sections. The quality of reconstructed brains with multilayer sections is reflected at the interfaces between two physical sections, while texture-based metrics equally weight patches and tend to be influenced by the false texture at section joints. Since our work aims at registering multilayer sections to each other, the edge-based tensor voting evaluation [12] is most suitable to reflect the structure consistence among multilayer sections. In addition, this metric differs from the cost functions adopted by our registration methods, and thus is more objective.

Fig. 2.8 illustrates the three steps of the evaluation scheme [62]: token extraction, tensor voting and correlation. Tokens are the points that lie on edges [62]. Edges are extracted on imaging planes, and Fig. 2.8 (b) is the cross-section view of the 3D edge stack. In the

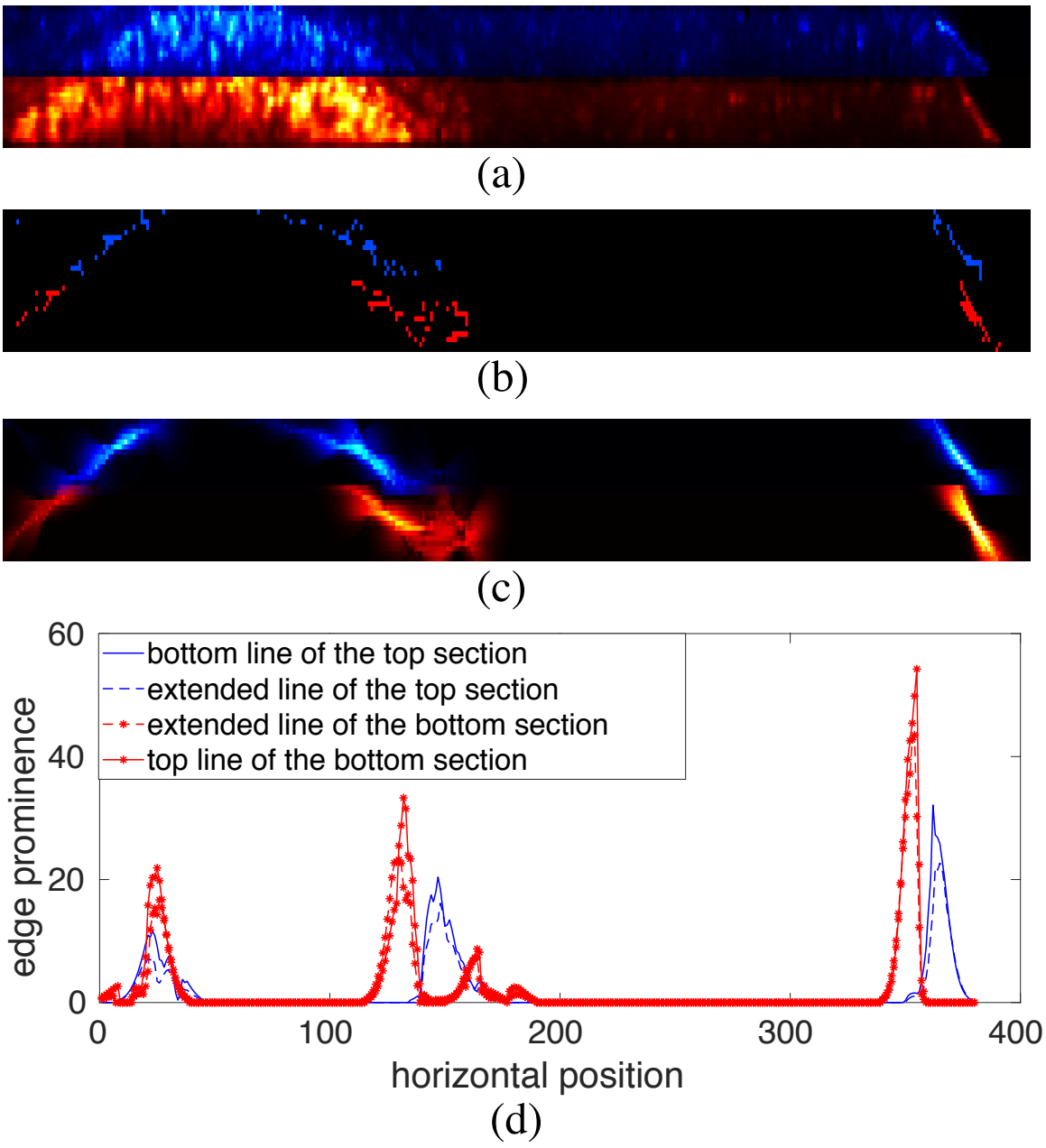


Figure 2.8: (a) is a cross-section view of two aligned sections. (b) is the cross-section view of extracted edges. (c) is the edge prominence map voted by edge points in (b). The structures in the top section extend one layer downward, and the structures in the bottom section extend one layer upward. Therefore, the edge prominence in the top section and the bottom section has two layers overlapped in (c). (d) plots the edge prominence in overlapped layers in (c).

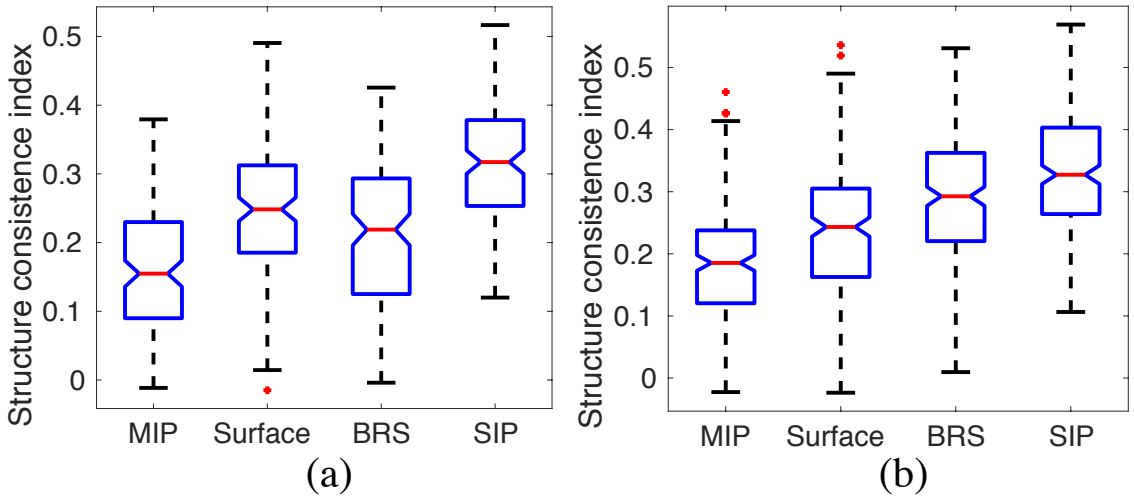


Figure 2.9: Structure consistence index of registered section pairs. (a) Sections imaged on the horizontal plane. (b) Sections imaged on the coronal plane.

Table 2.2: Structure consistence index of section-to-section alignment

		MIP	Surface	BRS	SIP
Horizontal	Median	0.1536	0.2669	0.2167	**0.3172**
	Mean	0.1622	0.2681	0.2059	**0.3185**
Coronal	Median	0.1837	0.2562	0.2922	**0.3298**
	Mean	0.1829	0.2482	0.2870	**0.3355**

second step of tensor voting, tokens communicate with each other and agree on a significant structure [62]. The output of token communication is referred as edge prominence as shown in Fig. 2.8 (c). At last, the zero-normalized cross-correlation of the edge prominence at the overlapped regions is used as the structure consistence index for one cross-section view. Since the quality of the reconstructed brain is reflected by the cross-section views [30, 36, 37], the overall reconstruction quality is the average of the structure consistence indexes of all the cross-section views. Two key parameters for this evaluation metric are the tensor voting scale and the overlap range [62]. The larger the tensor voting scale, the blurrier the edge prominence map (Fig. 2.8 (c)) will be. The tensor voting scale is set as 10, and the overlap range is set as 2 for the example in Fig. 2.8 and for all the following experiments.

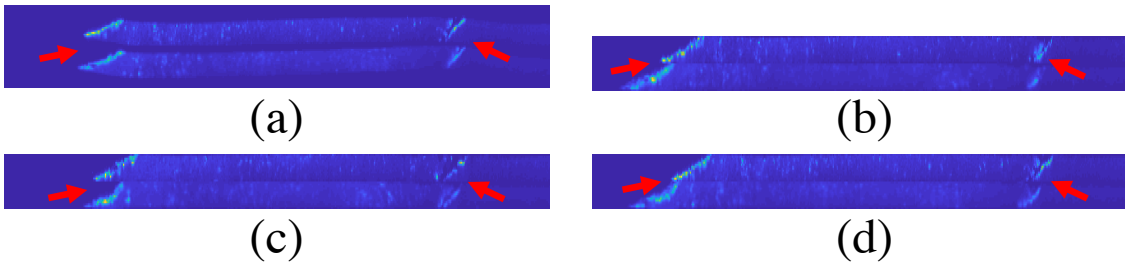


Figure 2.10: Examples of section-to-section registration with four different representatives. (a) MIP. (b) Surface. (c) BRS [1]. (d) Proposed SIP.

2.3.3 Section-to-section alignment

In this experiment, all the sections are flipped according to the manually labeled ground truth before registration. The adjacent section pairs are aligned to each other with the four different representatives of the multilayer sections: MIP, Surface, BRS and SIP. The sections are classified by their imaging planes: horizontal plane and coronal plane. There are 134 pairs of adjacent sections in the group of horizontal plane, and 213 pairs in the group of coronal plane.

Fig. 2.9 shows the quality of registered section pairs, and Table 2.2 summarizes the quantitative results. Single factor ANOVA is conducted to compare the performance of the four different multilayer representatives. The mean structure consistence indexes of these four methods are significantly different with $p < 10^{-36}$ for horizontal sections, and $p < 10^{-46}$ for coronal sections. Since sections are flipped according to the ground truth, this experiment mainly compares the accuracy of structure position reflected by different representatives. Maximum intensity projection of a multilayer section blends structures from different layers, and thus provides the lowest structure consistence index. After tissue flattening, structures in surface layers often show minor offsets compared with the structural trend exhibited by intra-section regions. Propagated surfaces correct such structure offsets in the surface layers and achieve the best alignment result. One interesting result is that the performances of Surface and BRS [1] are different on coronal and horizontal sections. This is because surface layers reflect the shape changing better than selected middle layers, and the shape growing trend is more significant among horizontal sections. However, the drawback of Surface is that the signal intensity and

contrast are weak. The following auto-flip experiment demonstrates how this drawback affects the reconstruction.

Fig. 2.10 shows one example to illustrate how the proposed tissue flattening and SIP improve the performance of adjacent section alignment. Registered with MIP, two sections in Fig. 2.10 (a) pose a noticeable offset. Minor offsets still exist in the aligned sections with surface layers, but is corrected by the structure-based intensity propagation. The best reference layers selected by BRS [1] are the 10th and the 8th layers for the first and the second multilayer sections, and therefore the registration does not capture the growing trend.

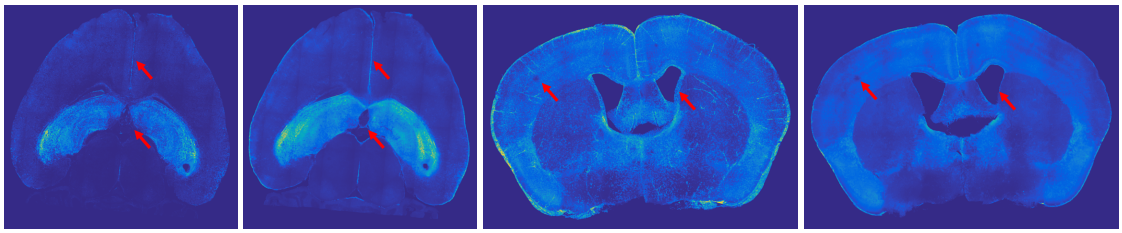
2.3.4 Flip detection

Although experimenters carefully track and label orientations of sections, inconsistencies can happen during large data collection, especially when tissue clearing is involved. Among the 367 raw sections in our experiment, 169 sections (46%) are flipped according to the ground truth. Incorrectly flipped sections are a major source of artifacts in brain reconstruction, especially for whole brain reconstruction with a series of sections: one incorrectly flipped section may influence flip decisions of all the following sections.

In this part, automatic flip detections with different representatives of multilayer sections are compared. Adjacent sections are registered, but only the first section is flipped according to the ground truth. The flip status of the second section is automatically decided according to affine registration results. In this experiment, flip detection is evaluated for each section pair, and one incorrect flip decision does not influence the following flip decisions. Table 2.3 reports the flip detection results of the four versions in terms of Type I error (falsely flipped sections) and Type II error (falsely unflipped sections). Flip detection based on the propagated surfaces achieves highest accuracy. An interesting phenomenon is that Surface achieves much worse flip detection accuracy than others. This is because although surface layers reflect the shape changing more accurately, they do not preserve the intensity information well. For flip detection, the asymmetric intensity information is crucial.

Table 2.3: Incorrect flip detection number

		MIP	Surface	BRS	SIP
Horizontal (134)	Falsely flipped	2	26	2	**1**
	Falsely unflipped	**1**	16	3	**1**
Coronal (213)	Falsely flipped	8	42	4	**2**
	Falsely unflipped	7	25	7	**2**



(a) (b) (c) (d)

Figure 2.11: The surface layer after tissue flattening in (a) shows less signal intensity and structure contrast compared with the propagated surface in (b). Compared with the maximum intensity projection in (c), the propagated surface in (d) removes the structures that only stay in certain layers and preserves the major structures. (a) and (b) are sections on the horizontal plane. (c) and (d) are sections on the coronal plane.

Fig. 2.11 shows different multilayer section representatives for flip detection. Compared with the propagated surface in Fig. 2.11 (b), the surface layer in Fig. 2.11 (a) lessens structure contrast. On the contrary, the maximum intensity projection in Fig. 2.11 (c) contains many structures that could mislead flip detection, such as the neurons. Only propagated surface layers preserve the consistent information among sections that is helpful to flip detection, and also get rid of the distracting information that only shows in certain layers.

2.3.5 Whole brain reconstruction

In this part, 20 mouse brains are reconstructed with and without the ground truth of flip status. Table 2.4 summarizes the reconstruction quality by different approaches. The same structure consistence index introduced in Section 2.3.2 is adopted to measure the whole brain reconstruction quality. As stated before, for the whole brain reconstruction, one wrong flip can influence the following flips. Therefore, a wrong flip at the beginning tends to cause more incorrect flip detections in the following sections, such as the sixth brain reconstructed

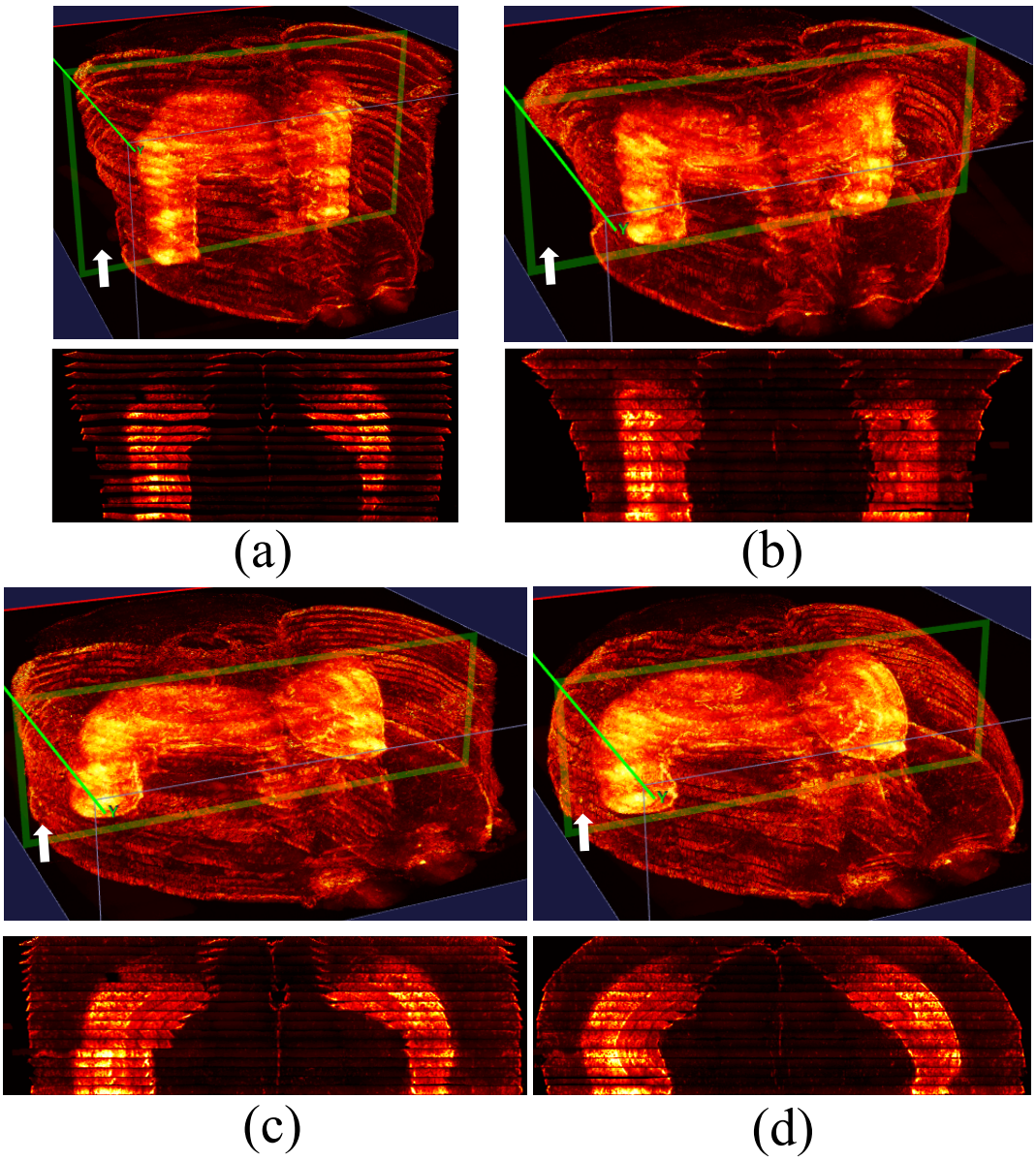


Figure 2.12: Reconstructed Brain 1, sectioned on the horizontal plane. The arrows indicate the ventral-dorsal direction. The frames indicate the positions of the cross-section views. The cross-section views are on the coronal plane. (a) MIP: 0.2746. (b) Surface: 0.2820. (c) BRS: 0.2560. (d) SIP: 0.4419.

with the propagated surface. Single factor ANOVA confirms that the difference between the four reconstruction approaches is significant with $p < 10^{-7}$ for the 20 brains. Fig. 2.12 and Fig. 2.13 show two examples of the reconstructed brains with automatic flip detection. Brain

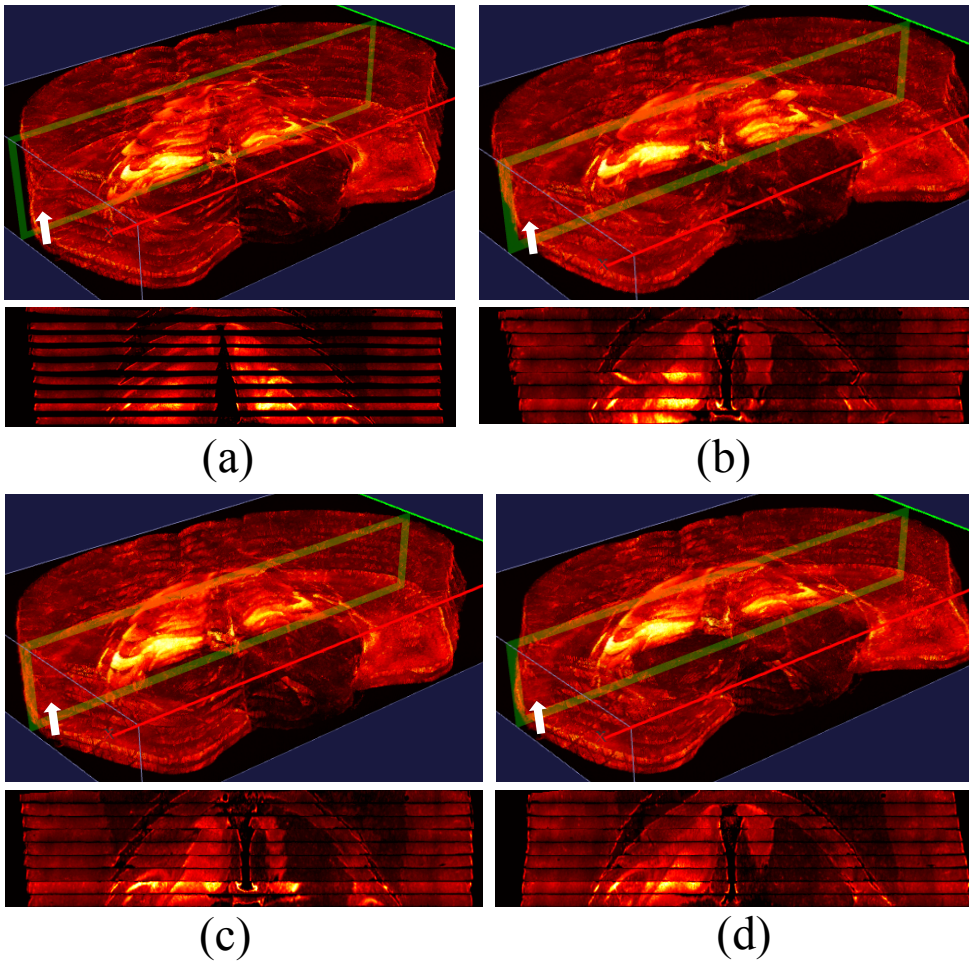


Figure 2.13: Reconstructed Brain 11, sectioned on the coronal plane. The arrows indicate the posterior-anterior direction. The frames indicate the positions of the cross-section views. The cross-section views are on the horizontal plane. (a) MIP: 0.3991. (b) Surface: 0.3419. (c) BRS: 0.4079. (d) SIP: 0.5240.

reconstruction with propagated surfaces improves the reconstruction quality in two aspects: more natural outer contour shape from the 3D overview and more consistent structure changing from the cross-section views.

2.4 Discussion

In this work, two structure correction methods are proposed for brain reconstruction with multilayer tissue sections: tissue flattening and structure-based intensity propagation. Tissue

Table 2.4: Structure consistence index of whole brain reconstruction

# Brain	Quality with flip groundtruth				Quality with auto-flip				# incorrect flip with auto-flip			
	MIP	Surface	BRS	SIP	MIP	Surface	BRS	SIP	MIP	Surface	BRS	SIP
1	0.2660	0.3184	0.2678	**0.4464**	0.2746	0.2820	0.2560	**0.4419**	2	12	6	**1**
2	0.2246	0.2922	0.2608	**0.3325**	0.2246	0.2760	0.2113	**0.3325**	**0**	9	13	**0**
3	0.1663	0.3145	0.2650	**0.3760**	0.1888	0.2574	0.2401	**0.3760**	9	10	10	**0**
4	0.2199	0.2308	0.2185	**0.3042**	0.2199	0.2213	0.1796	**0.3042**	**0**	4	12	**0**
5	0.1449	0.2419	0.1648	**0.2762**	0.1422	0.2250	0.2009	**0.2762**	4	6	7	**0**
6	0.2265	0.3064	0.2892	**0.3426**	0.2179	**0.3291**	0.2752	0.3064	18	**4**	10	16
7	0.2458	0.2812	0.2841	**0.3374**	0.1941	0.2675	0.2641	**0.3374**	18	10	5	**0**
8	0.1968	0.2468	0.2617	**0.2798**	0.1696	0.2262	0.2467	**0.2798**	10	11	6	**0**
9	0.2053	0.2729	0.2262	**0.2935**	0.1748	0.2596	0.1749	**0.2935**	13	8	7	**0**
10	0.3016	0.3016	0.3667	**0.4215**	0.2348	0.2763	0.3478	**0.4215**	5	12	9	**0**
11	0.3991	0.3723	0.4079	**0.5240**	0.3991	0.3419	0.4079	**0.5240**	**0**	2	**0**	**0**
12	0.2647	0.3250	0.3441	**0.3996**	0.2952	0.3179	0.3426	**0.3996**	25	9	1	**0**
13	0.1953	0.2413	**0.3312**	0.2746	0.1762	0.2678	0.3073	**0.3097**	19	**9**	15	10
14	0.2257	0.2775	0.3371	**0.3636**	0.1987	0.2762	0.3043	**0.3636**	20	11	21	**0**
15	0.2021	0.3236	0.3215	**0.3668**	0.2049	0.3332	0.3322	**0.3571**	18	10	9	**3**
16	0.3000	0.2607	0.2480	**0.3406**	0.3124	0.2979	0.2619	**0.3406**	2	8	5	**0**
17	0.2841	0.2770	0.3031	**0.3205**	0.2841	0.2538	0.2720	**0.3205**	**0**	13	21	**0**
18	0.1868	0.3177	0.3367	**0.3564**	0.1690	0.3321	0.3022	**0.3520**	12	11	18	**3**
19	0.2341	0.3130	0.3410	**0.3541**	0.2510	0.2750	0.3098	**0.3541**	21	12	12	**0**
20	0.2291	0.2262	**0.3207**	0.2640	0.2298	0.2409	**0.2649**	0.2640	19	6	12	**0**
mean	0.2359	0.2871	0.2948	**0.3487**	0.2281	0.2779	0.2751	**0.3477**	10.75	8.85	9.95	**1.65**

flattening projects the warped multilayer sections onto the bottom surfaces. Structure-based intensity propagation extends the intensity information within 3D section to surface layers. After tissue flattening and structure-based intensity propagation, the propagated surfaces serve as robust representatives of multilayer tissue sections, and facilitate the following registration and flip detection. The proposed methods can be incorporated into existing registration-based reconstruction frameworks as a preprocessing step. Experiments on 367 brain sections from 20 mouse brains verify the effectiveness of the proposed methods. Section-to-section experiments evaluate the performance improvement on registration and flip detection by the proposed methods. The whole brain reconstruction is evaluated in the last experiment. With reconstructed 3D mouse brains, observation and analysis of each single section can be done in an unified 3D coordinates for the whole mouse brain. For example, activated cell centers during SE from different sections can be aligned to form a graph with an unique 3D topology with the reconstructed brain.

Chapter 3

Auxiliary Imaging Modality[1]

Besides high resolution microscopy, other imaging modalities provide valuable data for brain research as well. With MRI, brains are imaged without slicing. The intact brain data are helpful for cross brain registration and the evaluation of brain reconstruction. However, the resolution of MRI is less than microscopy data, and acquiring high quality MRI data is important. In this work, we mainly focus on improving the quality of MRI images with parameter selection. The proposed methods can be used for the quality assessment of other auxiliary imaging modality as well. For example, image enhancement on brain images acquired by the low resolution microscopy [39], such as denoising and deblurring, could benefit from the proposed parameter selection pipeline. Measuring the perceptual image quality by subjective experimentation is time-consuming and expensive, so designing an image quality assessment (IQA) algorithm that agrees with the human visual system (HVS) [64–66] is a foundational image processing objective.

IQA algorithms are classified based on the amount of information from the reference image (the distortion-free image) that is required: full-reference (FR), reduced-reference (RR) and no-reference (NR). FR-IQA [3, 14, 67–69] is a relatively well-studied area. Traditional methods like mean squared error (MSE) and signal-to-noise ratio (SNR) are used as the standard signal

[1] ©2016 IEEE. Reprinted with permission from Haoyi Liang and Daniel Weller, "Comparison-based image quality assessment for selecting image restoration parameters", IEEE Transactions on Image Processing, November, 2016

fidelity indexes [70]. A more sophisticated FR-IQA algorithm, Structural Similarity Index Method (SSIM) [14], considers the structure information in images and performs well in different applications [70–73]. RR-IQA algorithms [5, 74–76] require some statistical features of the reference image, such as the power spectrum, and measure the similarity of these features from the reference image and the distorted image. NR-IQA algorithms usually adopt two different approaches. The first kind of NR-IQA [77–81] algorithms have an approach similar to that of RR-IQA. The difference is that rather than extracting the features from the reference image, this kind of NR-IQA algorithm extracts statistical features from a training set. The second kind of NR-IQA algorithm [73, 82, 83] adopts a local approach to quantifying structure as a surrogate for quality. A common implementation of the second approach is calculating local scores by analyzing the coherence of image gradients. The overall score is synthesized by taking the average of the local scores.

Among these three kinds of IQA algorithms, speed and accuracy generally decrease from FR-IQA, RR-IQA to NR-IQA progressively. However, reference images do not exist in many cases. In applications like parameter selection for image restoration algorithms, a set of distorted images with the same image content are available. A novel comparison-based IQA (C-IQA) framework is proposed in this thesis to make full use of these differently distorted images. A parallel two-step framework is adopted in C-IQA. First, a residual image is calculated by taking the difference between two input images, and the quality of the residual image is evaluated. Next, the contribution from two input images to the residual image is calculated. Finally, a simple procedure combines the first two parts: the input image that mainly contributes to high quality residual patches receives positive scores, while the input image that is more responsible for the distorted residual patches receives negative scores. Depending on the type of the distortion, different quality indexes, such as the blockiness index [84], can be used in the first part and a multi-metric fusion scheme [85–87] can further improve the versatility of the proposed framework.

It is worth differentiating RR-IQA and C-IQA since they both use extra data beyond a single distorted image. The most significant distinction between RR-IQA and C-IQA is that

the extra data required by most RR-IQA algorithms is distortion-free, while the extra data that C-IQA has is usually distorted. Therefore, RR-IQA algorithms specifically treat one image as the distorted image to be evaluated and the other information as truth or a reference. However, C-IQA treats the two input images equally without any prior knowledge of the quality of input images. In [5], Wu et al. take RR-IQA as a measurement of the fidelity of the distorted image to the original image. The RR-IQA method proposed in [74] depends on a distortion-free ancillary channel to transmit the features extracted from the original image. In [75], Soundararajan et al. point out that the output of their RR-IQA method is a single positive value that does not indicate if the input is better or worse than the reference image. On the contrary, the final output of our proposed C-IQA is a single real value that indicates which one of the input image is better and by how much. In brief, RR-IQA methods rely on the integrity of the extra data, while C-IQA is able to decide the relative quality only with two input images of the same scene.

The rest of the chapter is organized as follows. Section 3.1 introduces and compares different NR-IQA methods. Section 3.2 elaborates on the details of C-IQA. The algorithm used for image reconstruction and the framework of parameter trimming are introduced in Section 3.3. In Section 3.4 experiments are conducted on two widely used IQA databases, LIVE [2] and CSIQ [3], to verify the performance of C-IQA on parameter selection. Section 3.5 reviews the novelty and experimental results of the proposed C-IQA, and discusses further work on comparison-based IQA framework.

3.1 Existing NR-IQA methods

In [73], Shnayderman et al. classify NR-IQA algorithms into two types: global approaches and local approaches. The underlying difference between these two methods are the features used by different NR-IQA algorithms. Statistical features, such as the distribution of wavelet coefficients [77, 78], are extracted for global approaches. Local approaches usually rely on the structure information, such as the edge prominence [73, 82]. Usually, a regression model is adopted to synthesize the statistical features into an overall image quality, while the structure

features are able to reflect the image quality directly. The assumption of the statistical feature distribution is changed when considering the difference of two images, while the structure indexes that reflect the coherence of local gradients still work for the difference of two images. Therefore, the proposed C-IQA methods make use of a structure index, MetricQ. In the following part, we briefly review different NR-IQA methods and introduce one particular local structure index MetricQ [82].

3.1.1 Approaches with statistical features

The rationale of statistical feature-based NR-IQA methods [77–81] is that the distributions of natural scene statistics (NSS) share certain common characteristics among distortion-free images, and distortions will change these characteristics. For example, it is widely accepted that the wavelet coefficients of a natural image can be modeled by a generalized Gaussian distribution (GGD) [88, 89].

The main advantage of statistical features is that most of them are not dedicated to a specific distortion since the NSS features are a high-dimensional vector designed to be sensitive to various distortions. However, because of the high dimensionality of the statistical feature space, it is difficult to individually interpret and analyze these features quantitatively, and thus feature selection is largely an empirical work.

3.1.2 Approaches with structure indexes

Because human eyes are highly sensitive to the gradient in images, and the information in images can be well represented by their gradient [14, 82, 90], structure indexes usually reflect the spatial gradient information. Unlike the statistical indexes, most structure indexes represent the local quality directly without involving the learning process. However, the amount of the gradient, or total variation, itself is not a stable indicator of the quality [73]. Previous works [73, 82, 91] have shown that assessing the concentration of the gradient direction in an image is a promising way to evaluate the image quality. Among these works, MetricQ [82] shows encouraging results choosing denoising parameters. The underlying rationale of MetricQ

is that the more concentrated the gradient direction is, the better the quality of the patch is. It is a reasonable assumption since both of the two most common distortions, noise and blurring, disperse the distributions of the gradient direction.

3.1.3 MetricQ

The local quality index used by MetricQ is based on singular values of the local gradient matrix, which have been widely used as low level features in different image processing problems, such as tracking feature selection [92], recognition [93] and image quality assessment [73]. For each $n \times n$ local patch (w), the gradient matrix is

$$G = \begin{bmatrix} \vdots & \vdots \\ p_x(k) & p_y(k) \\ \vdots & \vdots \end{bmatrix}, \tag{3.1}$$

in which $p_x(k)$ and $p_y(k)$ are the gradients of the k^{th} pixel in the patch w on x and y directions. The singular value decomposition (SVD) of the gradient matrix, G, is defined as

$$G = USV^T = U \begin{bmatrix} s_1 & 0 \\ 0 & s_2 \end{bmatrix} \begin{bmatrix} V_1 & V_2 \end{bmatrix}^T, \tag{3.2}$$

where U and V are both orthonormal matrices. Vector V_1 is of size 2×1 and corresponds to the dominant direction of the local gradient; V_2 is orthogonal to V_1 and thus represents the edge direction. Singular values, s_1 and s_2, represent the luminance variances on V_1 and V_2 respectively. Intuitively, a large s_1 and a small s_2 indicate a prominent edge in the local patch. The gradient matrix, G, used in Eq. 3.1 and Eq. 3.2 is the 2D version of the gradient matrix used in Eq. 2.1

In MetricQ [82], two indices reflect the quality of a local patch: Image Content Index and Coherence Index. Image Content Index is defined as

$$Q = s_1 \frac{s_1 - s_2}{s_1 + s_2}, \tag{3.3}$$

and Coherence Index is defined as

$$R = \frac{s_1 - s_2}{s_1 + s_2}. \tag{3.4}$$

Q reflects the structure prominence in a local patch and R is used to determine whether a local patch is dominated by noise. The overall score of an image is calculated by

$$AQ = \frac{1}{MN} \sum_{i,j:R(i,j)>\tau} Q(i,j), \tag{3.5}$$

where $M \times N$ is the size of the image and τ is the threshold to decide whether a local patch is dominated by noise. $Q(i,j)$ and $R(i,j)$ are the Image Content Index and Coherence Index of the local patch centered at (i,j) in the image. A simplified interpretation of (3.5) is that AQ is the average structure index of local patches that have meaningful image content.

3.2 Comparison-based image quality assessment

Previous works on IQA [14, 65, 66, 79, 94] show that IQA performance can be significantly improved by taking advantage of the characteristics of HVS. For example, the structural information that human eyes are highly sensitive to is used by SSIM [14]. Traditional NR-IQA algorithms also try to exploit HVS features and make reasonable assumptions about natural scene images. However, one important aspect of HVS is ignored: comparison. In subjective IQA experiments [3], volunteers are required to evaluate the quality of an image by comparing it with a reference image, rather than giving an absolute score for the image. Although in most image processing applications, the reference image does not exist, a set of differently degraded images are available. In these cases, extending existing state-of-the-art FR-IQA and RR-IQA algorithms to comparison-based IQA algorithms is a natural thought. However, different from FR-IQA and RR-IQA algorithms, neither of the two input image qualities is known in the comparison-based IQA framework. As a result, in a comparison-based IQA algorithm, we not only measure the difference between two input images, but also assess the quality of the difference.

3.2.1 Framework of C-IQA

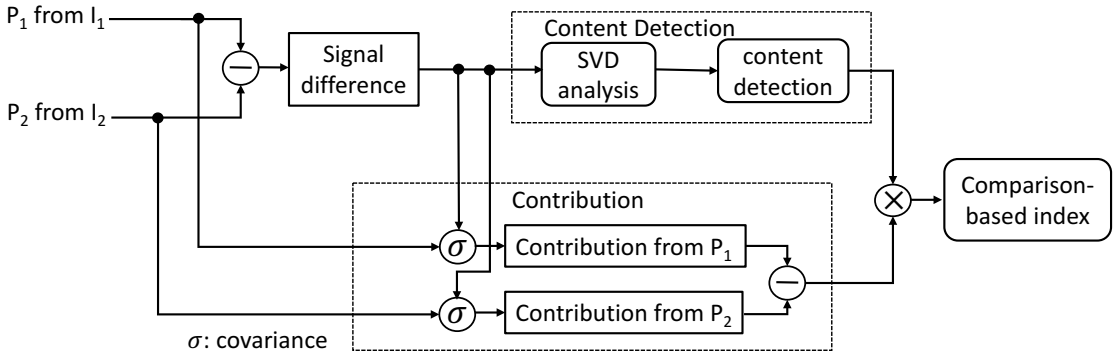


Figure 3.1: Flow Chart of the Comparison-based IQA: P_1 and P_2 are local patches from input images, I_1 and I_2, at the same location respectively. The Content Detection module determines whether there is a meaningful structure in the difference patch; the Contribution module calculates which patch mainly contributes to the difference patch; the Distortion Sensitivity Weighting module compensates the distortion sensitivity difference of patches with various texture complexities. The output, comparison-based index, indicates the relative quality of P_1 based on P_2.

As shown in Fig. 3.1, C-IQA has two input images, I_1 and I_2, and the output indicates the relative quality of I_1 based on I_2. No prior knowledge about the quality of two input images is known to C-IQA, and the relative quality can be either positive or negative depending on whether I_1 is better than I_2. We refer to the second image in C-IQA as the base image to distinguish it from the reference image in FR-IQA and RR-IQA. The implemented C-IQA method consists of two basic modules: Content Detection and Contribution. The third module, Distortion Sensitivity Weighting, is optional and its description is deferred to Section 3.2.3. In the rest of the chapter, we refer to the comparison-based IQA variation composed by the two basic modules as CQ and the variation with three modules as CDQ. Content Detection determines whether the difference between two input images contains any meaningful structure, and Contribution decides which image mainly contributes to the difference. CQ composes these two modules by the criterion that the input image that contributes to a structured difference is better and the input image that contributes to a random difference is worse. The Distortion Sensitivity Weighting module added in CDQ adjusts the distortion sensitivity difference among

patches with different texture complexity [14, 95].

Content Detection

The Content Detection module is based on the Coherence Index put forward in MetricQ [82]. Different from MetricQ, this index is calculated with the difference image between two input images in C-IQA. In MetricQ, limited by the information provided by a single input image, the algorithm does not know the texture complexity in the original image. Therefore, it is hard for an algorithm to estimate how concentrated the gradient should be. However, by mimicking the comparative way HVS works, C-IQA removes the main image content in the images by taking the difference, and thus it is easy for the Content Detection module to differentiate the patches with noisy and structured content.

Algorithm 3 Content Detection

$D_p = P_1 - P_2$
$G = [d_x(D_p)\ d_y(D_p)]$
$USV^T = SVD(G)$
$C_{ind} = \frac{s_1 - s_2}{s_1 + s_2}$ ▷ $s_1 > s_2$
if $C_{ind} > C_{thresh}$ **then**
  $is\_stru = 1$ ▷ structure
else
  $is\_stru = -1$ ▷ noise
end if

In Alg. 3, P_1 and P_2 are two patches of size $n \times n$ from I_1 and I_2 respectively, G is the same 2-column gradient matrix defined in (3.1), $SVD(G)$ represents taking the SVD operation on G, and s_1 and s_2 are the singular values of G. C_{thresh} is a constant threshold to binarize C_{ind}. The binary output is_stru indicates whether there is a meaningful structure in the difference of local patches.

Contribution from patches

Once the difference is classified into noise or structure, the Contribution module is designed to find out which of the two input images mainly contributes to the difference image. In

our implementation, the luminance-normalized covariance between the input image and the difference image is used to measure the contribution.

Algorithm 4 Contribution

$D_p = P_1 - P_2$
$M_p = max(\frac{mean(P_1)+mean(P_2)}{2}, \frac{1}{n \times n})$
$ctri1 = cov(P_1, D_p)$
$ctri2 = cov(P_2, -D_p)$
$ctri = \frac{ctri1-ctri2}{M_p}$

In Alg. 4, $mean(P_i)$ calculates the average of the local patch, and $cov(x_1, x_2)$ calculates the covariance between two input patches,

$$cov(x_1, x_2) = \frac{(x_1 - mean(x_1))^T (x_2 - mean(x_2))}{n^2 - 1},$$

x_1 and x_2 are vectorized patches of size $n^2 \times 1$. The output $ctri$ represents that P_1 contribute to D_p more than P_2 does by how much. A negative $ctri$ means that P_2 mainly contribute to D_p.

The comparative quality index for each local patch is calculated by

$$C_Q = is\_stru \cdot ctri.$$

The overall comparative quality of I_1 based on I_2 is

$$CQ(I_1, I_2) = \frac{1}{M \times N} \sum_{i,j=(n/2):(M-n/2)} C_Q(i, j),$$

where $C_Q(i, j)$ is the local comparative quality index centered at (i, j) in the image, $n \times n$ is the size of the local patch and $M \times N$ is the size of the image. Patches that are outside the boundary of the image are not included in the calculation. A positive $CQ(I_1, I_2)$ means I_1 is better than I_2, and the absolute value quantifies the quality difference. Due to the anti-symmetric design

of the algorithm, $CQ(I_1, I_2) = -CQ(I_2, I_1)$.

3.2.2 Justification of CQ

Inspired by Li's work [96] which claims that an IQA model should be based on three quantities: edge sharpness, random noise level and structure noise, we classify the distortions by residual images, the difference between a distorted image and the original image. In our classification, distortions can be categorized into two types: introducing a random residual image, or introducing a structured residual image. In most cases, random residual images correspond to noise-like distortions and structured residual images correspond to blurring-like distortions. In this part, we prove how C-IQA works under these two distortions.

Assume I_{true} is the original image, and I_1, I_2 are two distorted images. The residual images are calculated by,

$$e_i = I_i - I_{true},\ i = 1, 2.$$

Similarly, for each patch we have

$$e_{Pi} = P_i - P_{true},\ i = 1, 2.$$

Random residual image Residual images behave like noise in this case. If we assume I_1 is more severely distorted than I_2, then we have $E[\|e_{P1}\|_2^2] > E[\|e_{P2}\|_2^2]$. The expectation of the

local comparative quality index is

$$\begin{aligned}
E[C_Q] &= E[ctri \cdot is_stru] \\
&= E[(ctri1 - ctri2) \cdot is_stru] \\
&= E[cov(P_1, P_1 - P_2) - cov(P_2, P_2 - P_1)] \\
&\quad \cdot E[is_stru] \\
&= -E[cov(P_{true} + e_{P1}, e_{P1} - e_{P2}) \\
&\quad - cov(P_{true} + e_{P2}, e_{P2} - e_{P1})] \\
&= -E[2 \cdot cov(P_{true}, e_{P1} - e_{P2}) \\
&\quad + cov(e_{P1}, e_{P1}) - cov(e_{P2}, e_{P2})] \\
&= -E[cov(e_{P1}, e_{P1})] + E[cov(e_{P2}, e_{P2})] \\
&\quad < 0.
\end{aligned}$$

The three most important properties in the derivation are the irrelevance between P_{true} and e_{Pi}, the randomness of e_{Pi}, and independence of is_stru and $ctri1$, $ctri2$. The result $E[C_Q] < 0$ agrees with our assumption that I_1 is more severely distorted than I_2. When I_2 is more severely distorted, the same proof shows $E[C_Q] > 0$.

Structured residual image If the residual images show structured information, the most probable reason is that the image is distorted by a blurring-like distortion. Because the blurring filter acts as a low-pass filter, the residual images show a structure that is inversely related to the original image [97] to smoothen the high contrast on the edges.

Without loss of generality, we assume more blurring happens in I_1 than I_2, which means $E[|e_{P1}|] > E[|e_{P2}|]$. The expectation of the local comparative quality index is

$$\begin{aligned}
E[C_Q] &= E[ctri \cdot is_stru] \\
&= E[(ctri1 - ctri2) \cdot is_stru] \\
&= E[cov(P_1, P_1 - P_2) - cov(P_2, P_2 - P_1)] \\
&= E[cov(P_{true} + e_{P1}, e_{P1} - e_{P2}) \\
&\quad - cov(P_{true} + e_{P2}, e_{P2} - e_{P1})] \\
&= E[cov(2 \cdot P_{true}, e_{P1} - e_{P2}) \\
&\quad + cov(e_{P1} + e_{P2}, e_{P1} - e_{P2})] \\
&= E[cov(2 \cdot P_{true} + e_{P1} + e_{P2}, e_{P1} - e_{P2})] \\
&\quad < 0.
\end{aligned}$$

The most important step in this derivation is the last step. Since $E[|e_{P1}|] > E[|e_{P2}|]$, $e_{P1} - e_{P2}$ also demonstrates a structure that is inversely related to the original image as e_{Pi}. As long as the distortion is not severe enough to remove the structure in the original image, $2 \cdot P_{true} + e_{P1} + e_{P2} = P_1 + P_2$ is positively related to the original image. As a result, $E[cov(2 \cdot P_{true} + e_{P1} + e_{P2}, e_{P1} - e_{P2})] < 0$, which agrees with our assumption that I_1 is more severely distorted than I_2. Following the same steps, we can show $E[C_Q] > 0$ if I_2 is more severe distorted than I_1.

3.2.3 Distortion sensitivity weighting

We have proven that only with Content Detection and Contribution, the CQ can give correct results if both of the two input images are distorted by one distortion, either noise-like distortion or blurring-like distortion. However, another important property of HVS is missed in CQ: the response of HVS to the same distortion is texture-dependent. One example of this HVS property is that after being distorted by the same amount of Gaussian noise, the distortion

in the image with simpler texture is more obvious. In this part, we first investigate such texture-based response of CQ and then design a weighting module to adjust the distortion sensitivity of CQ to different textures. We refer to the improved C-IQA method with Distortion Sensitivity Weighting module as CDQ.

In CQ, Content Detection is a qualitative module that detects the meaningful structure and the Contribution module quantifies the relative quality. Therefore, the Contribution module may implicitly include distortion sensitivity weighting. We design an experiment to explore the relation between the texture complexity and the output of Contribution. In this experiment, 140 patches of size 101×101 with homogeneous texture are selected from LIVE [2] and CSIQ [3], and eight samples of these patches are shown in Fig. 3.2. As the representatives of blurring-like and noise-like distortions, a bilateral filter and Gaussian noise with the same parameters are applied to each patch. According to the Weber-Fechner law [98], we use luminance-normalized total variation as the perceived texture complexity, $T\_ind = \frac{TV(P)}{mean(P)}$, where $TV(P)$ is the total variation in the original patch and $mean(P)$ is the average of the original patch. The relation between texture complexity, T_ind, and the output of Contribution module, $ctri$, are plotted in Fig. 3.3. Each circle in Fig. 3.3 represents a patch sample. It is clear that $ctri$ is almost linear related to texture complexity, T_ind, when blurring happens. On the contrary, T_ind shows no relation with $ctri$ when the distortion is noise. The reason for this is that blurring is a highly image-dependent distortion, and the residual image is more prominent at areas where total variation is high. After figuring out the blurring sensitivity compensation mechanism in CQ, we need to design an algorithm to compensate the sensitivity difference to noise.

Because noise-like distortion tends to increase the total variation while blurring-like distortion tends to decrease the total variation, Alg. 5 uses the output of Content Detection to synthesize $T1\_ind$ and $T2\_ind$ into T_ind. After texture complexity estimation, we transfer T_ind to the smoothness index, S_{ind}, and compensate the sensitivity to noise.

In CDQ, the comparative quality index for each local patch is

$$CD_Q = is\_stru \cdot ctri \cdot weight.$$

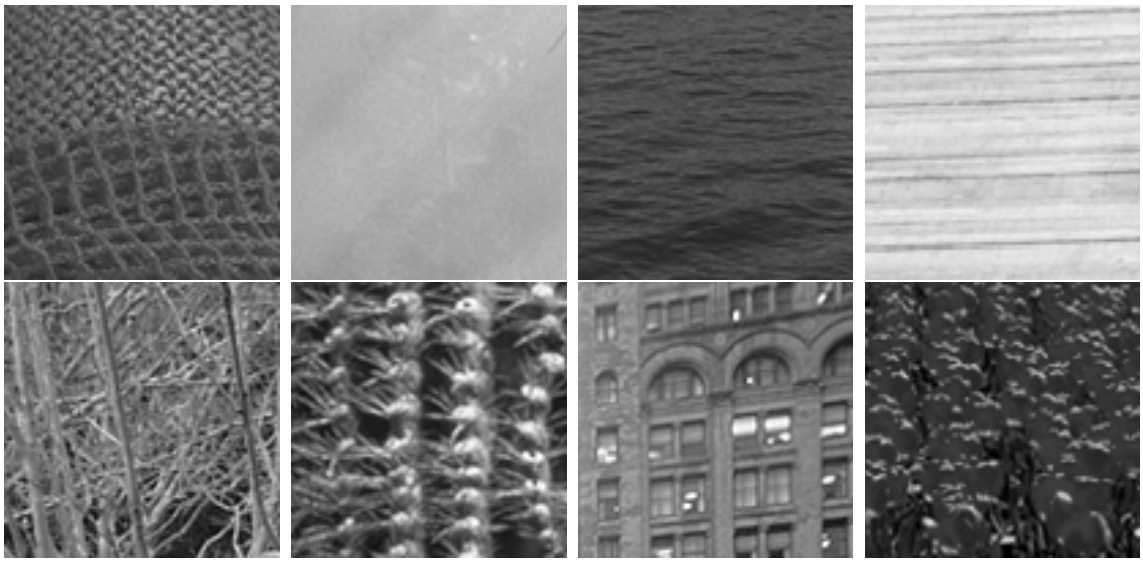
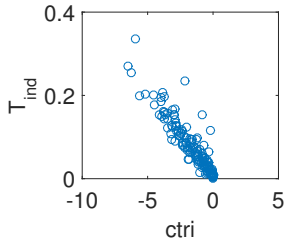


Figure 3.2: Patch samples with various texture complexity are selected from LIVE [2] and CSIQ [3] to verify the difference of distortion sensitivity in CQ. The assumption is that patches with complex texture are more robust to noise, while patches with flat texture are more robust to blurring.

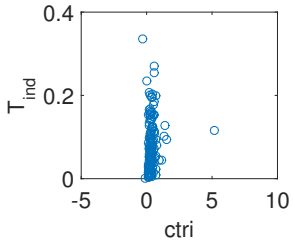
The overall comparative quality of I_1 based on I_2 is calculated by taking the average of local comparative quality index as CQ does.

3.2.4 Comparison between CDQ and SSIM

SSIM consists of three components: structure (loss of correlation), luminance (mean distortion) and contrast (variance distortion). In CDQ, the outputs of Content Detection and Distortion Sensitivity Weighting provide the quality of the difference image. The luminance and the contrast of an input image together determine the contribution of the input image to the difference image. Therefore, Content Detection and Distortion Sensitivity Weighting of CDQ together play the role of the structure part in SSIM. The difference is that without knowing which image has the better quality, CDQ has to analyze the quality of the structure in the difference image, rather than only measuring the structure distance as SSIM does. The Contribution module in CDQ is similar to the functions of luminance and contrast parts together in SSIM.



(a) Blurring ($\rho = -0.92$)



(b) Noise ($\rho = 0.10$)

Figure 3.3: Relations between the output of Contribution module, *ctri*, and texture complexity, *T_ind*. Each circle in the figure represents a sample patch. All the sample patches are degraded by the same amount of distortion for blurring and noise.

Algorithm 5 Distortion Sensitivity Weighting

$T1\_ind = \frac{TV(P_1)}{mean(P_1)}$
$T2\_ind = \frac{TV(P_2)}{mean(P_2)}$
if is_stru $= 1$ **then**
  $T\_ind = max\{T1\_ind, T2\_ind\}$;
else
  $T\_ind = min\{T1\_ind, T2\_ind\}$;
end if
$S\_ind = log(1 + \frac{1}{C_1 \times T\_ind})$
if is_stru $= 1$ **then**
  $weight = 1$
else
  $weight = -S\_ind$
end if

3.3 Parameter selection

As the motivation of C-IQA mentioned in the introduction, most image processing algorithms contain user-defined parameters (these image processing algorithms are referred as "target algorithms" in the following to differ from IQA algorithms). Parameter selection [82, 99–106] is of importance to these target algorithms. By parameter selection, some of these target algorithms [103, 104] achieve a faster convergence rate; some [101, 102] obtain a better restored image.

In this section, we first introduce an image reconstruction algorithm and illustrate the importance of parameter selection with this reconstruction algorithm. Next, a boosted parameter

selection framework for iterative image processing algorithm, parameter trimming [106], is introduced. In the following experimental section, we show target algorithms with the parameter trimming framework benefit from the parameters selected by CDQ.

3.3.1 Image reconstruction

Total variation (TV) reconstruction [107] is aimed at minimizing the cost function,

$$E_\beta(x) = \beta\|Dx\|_1 + \frac{1}{2}\|Sx - y\|_2^2 \,, \tag{3.6}$$

where x is the reconstructed image, y is the observed incomplete data set, S is the system matrix, D represents the difference matrix, and the TV regularizer $\|Dx\|_1$ combines gradients on two directions isotropically. In our implementation, $S = R\mathcal{F}$, where R represents the subsampling matrix and $\mathcal{F}$ represents the Fourier transform matrix. The regularization parameter β controls the sharpness of the reconstructed result. Large β oversmooths the reconstructed image, while small β leaves residual noise. A proper β is crucial to the performance of TV reconstruction. Split Bregman iteration [108] is used to solve (3.6). By making the replacement $d \leftarrow Dx$ and introducing the dual variable b, the split formulation of (3.6) becomes:

$$\begin{gathered} \min_{x,d} \beta\|d\|_1 + \frac{1}{2}\|Sy - y\|_2^2 + \frac{\mu}{2}\|d - Dx - b\|_2^2 \,, \\ s.t.\ d = Dx. \end{gathered} \tag{3.7}$$

The Split Bregman iteration solution to (3.7) is Alg. 6. In Alg. 6 we use the notation $K = (R^T R - \mu\mathcal{F}D^T D\mathcal{F}^{-1})$, $L_k = (\mathcal{F}^T R^T y + \mu D^T(d^k - b^k))$ and $s^k = \sqrt{|Dx^k + b^k|^2}$. μ is set as 0.01β to ensure a fast convergence rate.

To illustrate the necessity of parameter selection of TV reconstruction, the Brain image [4] is reconstructed with 30 values of β. These candidate values of β are uniformly sampled from 1.22×10^{-6} to 10 in logarithmic scale and three of the reconstructed results are shown in Fig. 3.4. The image quality indexes during the convergence process are plotted in Fig. 3.5(a) where

Algorithm 6 Split Bregman

Initialize: $x^0 = 0, d^0 = b^0 = 0$

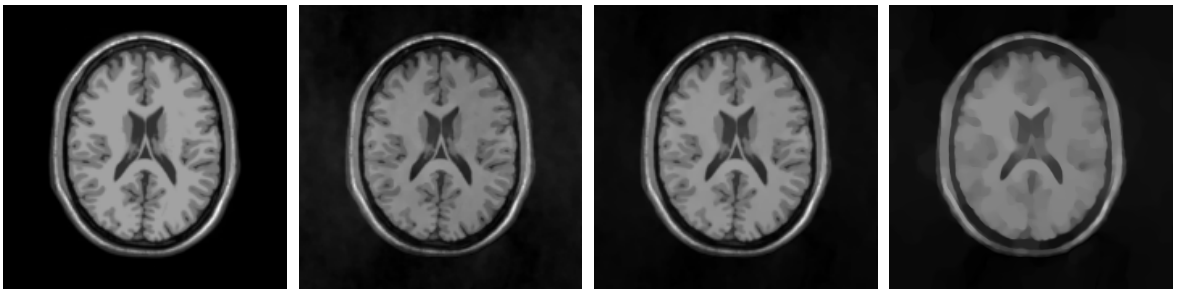
while stop criterion is not satisfied **do**

$$x^{k+1} = \mathcal{F}^{-1} K^{-1} L_k$$

$$d_{k+1} = \max(s^k - \frac{1}{\mu}, 0) \frac{Dx^k + b^k}{s^k}$$

$$b^{k+1} = b^k + (Dx^k - d^{k+1})$$

end while



(a) (b) (c) (d)

Figure 3.4: (a): original Brain image [4]; (b): reconstructed result with $\beta = 1.22 \times 10^{-6}$; (c): reconstruction result with $\beta = 4.46 \times 10^{-1}$; (d): reconstructed result with $\beta = 10$.

each line corresponds to one parameter candidate. The final reconstructed image qualities are plotted in Fig. 3.5(b). From Fig. 3.5, it is clear that if parameters that do not have the potential to achieve good results are terminated before convergence, considerable computation will be saved. This is the intuition of the parameter trimming in the next section.

3.3.2 Parameter trimming

A traditional approach to parameter selection [99–102] is selecting the parameters after the convergence of all the target algorithm instances. However, since either the target algorithms converge quickly [82, 102] or the NR-IQA algorithm is time-consuming [101], computational efficiency is not considered in previous works. In situations where target algorithms converge slowly or the set of parameter candidates is large, assessing image qualities and selecting the best parameter after all the algorithm instances converge would be too time-consuming to

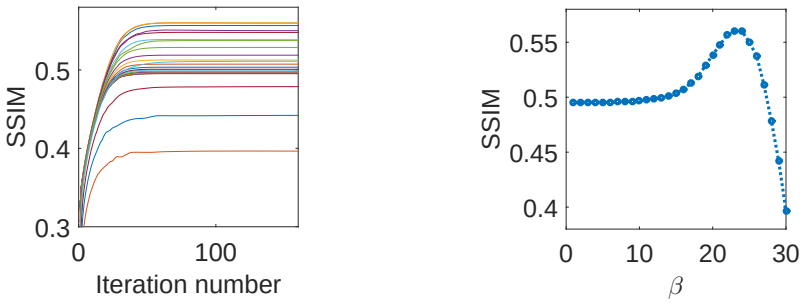


Figure 3.5: (a): Each line corresponds to an algorithm instance with a different regularization parameter. (b): Qualities of reconstructed results with different regularization parameters after 160 iterations.

be practical. Instead of placing the quality monitor at the output end, we first proposed a parameter trimming framework [106] that integrates the quality monitor into the target algorithms. In this section, we use image reconstruction as the application to illustrate the parameter trimming framework.

Assume I_m^i is the reconstructed result of the m^{th} parameter candidate at the i^{th} iteration. The trimming decision is made based on three indexes, q_m^i, g_m^i and p_m^i, which are the reconstructed quality, the quality increasing gradient and the prediction of the quality of I_m^i respectively. Because the image quality index we use here is a comparison-based index, the definitions of the these three indexes are modified to fit CDQ into the parameter trimming framework in [106]. Denoting the best reconstructed result at the i^{th} iteration is $best_i$, it satisfies $CDQ(I^i_{best_i}, I^i_{best_i-1}) \geq 0$ and $CDQ(I^i_{best_i}, I^i_{best_i+1}) \geq 0$. The three indexes used for parameter trimming, q_m^i, g_m^i and p_m^i, are defined as,

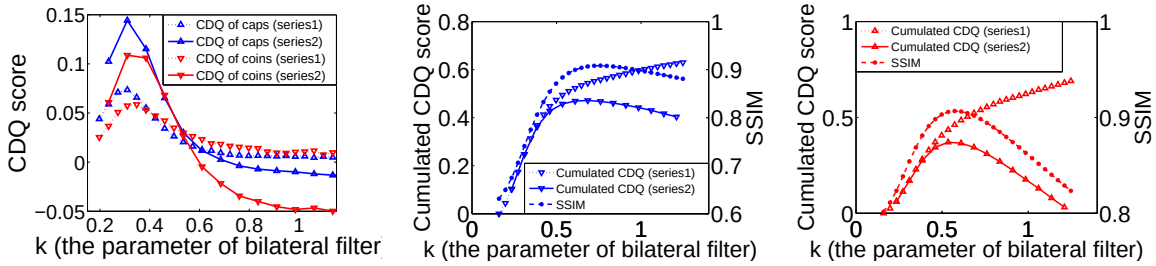
$$q_m^i = CDQ(I_m^i, I^i_{best_i}),$$

$$g_m^i = CDQ(I_m^i, I^{i-1}_{best_{i-1}}) - CDQ(I_m^{i-1}, I^{i-1}_{best_{i-1}}),$$

$$p_m^i = q_m^i + pre_{len} \cdot g_m^i.$$

(a) "caps" (b) "coinsinfountain"

Figure 3.6: Images in the gray scale [2] for the illustration of minimum resolution.



(a) (b) (b)

Figure 3.7: Minimum resolution of Comparison-based IQA algorithm: (a) CDQ scores of denoised images compared with their previous images (series1) and the one before previous images (series2); (b) SSIM scores and cumulated CDQ scores of "caps" in (a); (c) SSIM scores and cumulated CDQ scores of "coinsinfountain" in (a).

We set $pre_{len} = 4$ in all the experiments. More examples of the reconstruction and trimming process are shown in Section 3.4.3.

3.4 Experiments

We first introduce a key property, minimum resolution, that is unique to C-IQA in Section 3.4.1. In the next two parts, more comprehensive experiments are conducted to verify the effectiveness of C-IQA for parameter selection. The other NR-IQA algorithms that we use to compare CQ/CDQ with are DIIVINE (DII) [77], BRISQUE (BRI) [78], MetricQ (MQ) [82] and Anisotropy (Ani) [79]. Although RR-IQA methods are not suitable for parameter selection where the original image is not available, we include one RR-IQA [5] method to compare with C-IQA. The reason is that for some applications, such as delivering and decompressing visual

data sent to networked devices [78], both C-IQA and RR-IQA are practical. In [5], Wu. et al proposed a RR-IQA method that uses two numbers containing the information of the order and disorder parts of a reference image to help evaluate the quality of the distorted image. A widely accepted FR-IQA algorithm, SSIM [14], is used as the ground truth to evaluate the performance of other IQA algorithms. Two IQA databases used in the experiments are LIVE [2] and CISQ [3]. Parameters in CQ/CDQ are set as $C_{thresh} = 0.12$, $C_1 = 4.6$ and the size of a local patch is 9×9 pixels.

3.4.1 Minimum resolution of C-IQA

Since the comparison-based IQA is a brand-new approach, new properties arise. In this section, we illustrate the minimum resolution of C-IQA and corresponding solutions based on two images from LIVE [2] as shown in Fig. 3.6.

Similar to HVS, IQA algorithms are not able to make a convincing quality comparison between images whose difference is sufficiently small. In this part, we define the minimum mean squared difference (MSD) between two images required to make a convincing quality comparison as the minimum resolution. It is worth noticing that minimum resolutions vary over different distortions and different IQA algorithms.

For the traditional single-image-input NR-IQA algorithms, minimum resolutions can be regarded as the minimum MSD required to ensure consistency on a series of increasingly distorted images. However, under the comparison-based framework, a distorted image has different scores compared with different base images. We cannot refer to the consistency to define the minimum resolution for a comparison-based IQA algorithm. The minimum resolution for comparison-based IQA is defined as the minimum MSD required to preserve transitivity among a series of increasingly distorted images. We conduct an experiment on the images in Fig. 3.6 to demonstrate the transitivity.

Assume I_{org} is the original image, and I_1 is created by adding Gaussian noise to I_{org}. A series of gradually filtered images, $(I_1, I_2, \cdots, I_N)$, are denoised by bilateral filters [109], $BF_{(r,d)}$, where r and d are the variances of Gaussian range kernel for smoothing differences in intensities

and Gaussian spatial kernel for smoothing differences in coordinates. For simplicity, we reduce the parameters of bilateral filters to one by fixing the ratio between r and d, $BF_k = BF_{(0.1k,3k)}$. In Fig. 3.7(a), we show the CDQ scores of each image compared with its previous one in the denoised sequence (*series*1) and the CDQ scores compared with the one before its previous one (*series*2). We can see that CDQ scores in *series*1 are always positive, but pass 0 in *series*2. Therefore, the denoised image qualities monotonically increases in *series*1, but reach a peak in *series*2. In Fig. 3.7(b) and Fig. 3.7(c), we plot the cumulated CDQ scores in *series*1 and *series*2. It is clear that the cumulated CDQ scores fail to characterize the trend of image quality in *series*1, but successfully reflect the peak in *series*2. In this example, the MSD between adjacent images in *series*1 is below the minimum resolution of the bilateral filter, but the MSD between adjacent images in *series*2 is above the minimum resolution of the bilateral filter.

There are two ways to avoid the unwanted result of operating below minimum resolution. First, increase the MSD between adjacent images by increasing the parameter steps. Second, avoid comparing the adjacent images in a series of increasingly distorted images. The Key Image algorithm introduced in the Section 3.4.2 is an implementation of the second approach.

3.4.2 Experiment verification for parameter selection

Because the main motivation of Comparison-based IQA is parameter selection for image processing algorithms, two common image processing problems, image reconstruction and image denoising are used to demonstrate the parameter selection ability of the proposed C-IQA. The algorithm used for image reconstruction is the Split Bregman approach to total variation reconstruction [108]; the algorithm used for image denoising is the bilateral filter [109]. The optimal parameters of these two algorithms on different images are selected by SSIM, different NR-IQA algorithms, RR-IQA and C-IQA. The parameters selected by SSIM are compared with the ones selected by other IQA algorithms to evaluate the performance of other IQA algorithms.

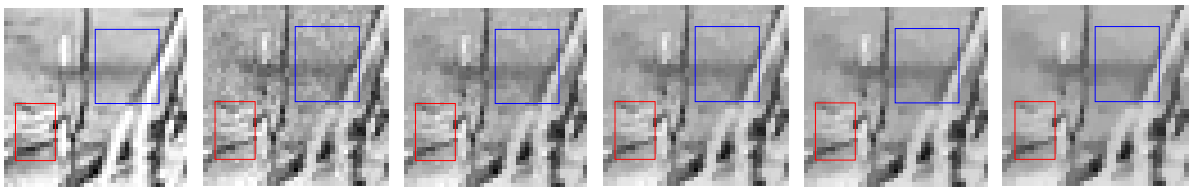
(a) Original "log_seaside" for CSIQ (b) Reconstructed "log_seaside" with β^*

Figure 3.8: One example of image reconstruction parameter selection. The best regularization parameter for this image, $\beta^* = 2.81 \times 10^{-2}$. The area highlighted by the red box is enlarged in Fig. 3.9.

Parameter selection for TV reconstruction

The algorithm used for image reconstruction is introduced in Section 3.3.1. In this experiment, 70% Fourier transform data are used to reconstruct the image and in order to be more realistic, Fourier transform data are distorted by Gaussian noise. The SNR is kept at 20 dB in all reconstruction experiments. All 30 regularization parameter candidates are uniformly selected between $[10^{-5}, 10^{-1}]$ in logarithmic scale.

One reconstruction example, "log_seaside", from CSIQ is shown in Fig. 3.8. The highlighted area in Fig. 3.8 is shown in details in Fig. 3.9 for the original image and reconstructed results with different regularization parameters, β. The RR-IQA method [5] selects $\beta_1 = 8.53 \times 10^{-4}$ as the optimal parameters; DIIVINE and BRISQUE select $\beta_2 = 1.49 \times 10^{-2}$; Anisotropy selects $\beta_3 = 2.04 \times 10^{-2}$; SSIM, CQ and CDQ select $\beta^* = 2.81 \times 10^{-2}$; MetricQ selects $\beta_4 = 3.86 \times 10^{-2}$. It is clear from Fig. 3.9 that as β increases, noisy component disappears and blurring occurs. Fig. 3.10 and Fig. 3.11 show how CQ works by comparing different reconstructed results. In Fig. 3.10, the reconstructed result of β_1, Fig. 3.9 (b), is compared with the result with optimal

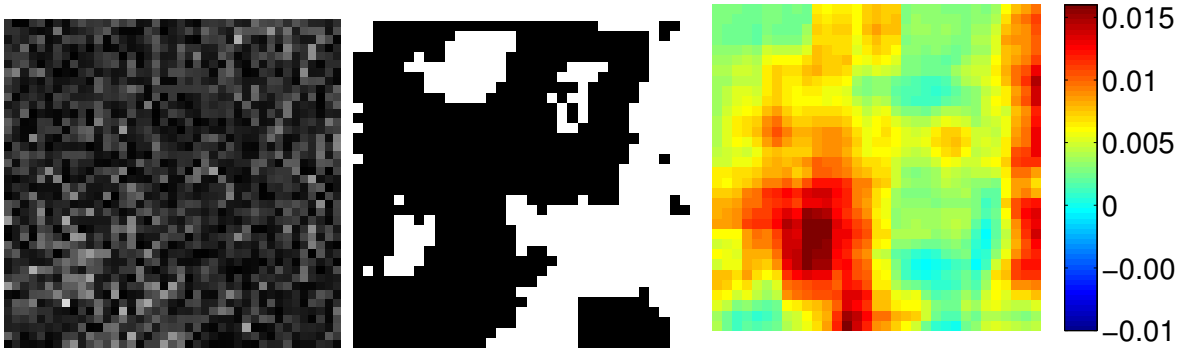


(a) Patches from the original image | (b) Reconstructed result with $\beta_1 = 8.53 \times 10^{-4}$ | (c) Reconstructed result with $\beta_2 = 1.49 \times 10^{-2}$ | (d) Reconstructed result with $\beta_3 = 2.04 \times 10^{-2}$ | (e) Reconstructed result with $\beta^* = 2.81 \times 10^{-2}$ | (f) Reconstructed result with $\beta_4 = 3.86 \times 10^{-2}$

Figure 3.9: Patches from the highlighted areas in Fig.3.8. The regularization parameter of the total variation term is β. RR-IQA [5] selects β_1, DIIVINE and BRISQUE select β_2, Anisotropy selects β_3, CQ and CDQ select β^*, and MetricQ select β_4. As β increases, noise is suppressed as shown in blue rectangles, while subtle structures is blurred as shown in red rectangles.

parameter, β^*, Fig. 3.9 (e). From Fig. 3.10 (a), it is clear that the difference patch shows a noise pattern. The black areas in Fig. 3.10 (b) indicate the areas that are likely to be taken as noise. Fig. 3.10 (c) shows the contribution difference from reconstructed results with β_1 and β^* to Fig. 3.10 (a). It should be noticed that the contribution difference in white areas in Fig. 3.10 (b) tends to be assigned a much smaller absolute value. Therefore, the CQ index of Fig. 3.9 (b) based on Fig. 3.9 (e) is a negative number that indicates Fig. 3.9 (b) is worse. Similarly, the comparison between reconstructed results with β_4 and β^* are shown in Fig. 3.11. A clear structured difference in Fig. 3.11 (a) is supported by the majority white area in Fig. 3.11 (b). The negative value in Fig. 3.11 (c) means that the contribution to the structured difference comes from β^*. It is worth to notice that although both the comparisons between results of β_1 and β^*, and β_4 and β^* lead to negative values that show β^* is better, the decision-making processes are different. When comparing β_1 and β^*, the difference is noisy and mainly comes from β_1; while when comparing β_4 and β^*, the difference is structured and mainly comes from β^*.

In the datasets of LIVE and CSIQ, there are 59 original images and each original image corresponds to 30 reconstructed results with different regularization parameters. The SSIM index difference between the best images chosen by SSIM and the one chosen by other IQA



(a) Patches difference (b) Content Detection (c) Contribution difference

Figure 3.10: CQ index of Fig. 3.9 (b) (β_1) based on Fig. 3.9 (e) (β^*). (a) is the difference between the image with β_1 and β^*. The white in (b) stands for structured areas in (a) and the black stands for noisy area. (c) shows the contribution difference between β_1 and β^*. On this local patch, $CQ(P_{\beta_1}, P_{\beta^*}) = -1.83 \times 10^{-3}$.

Table 3.1: Accuracy of parameter selection for image reconstruction

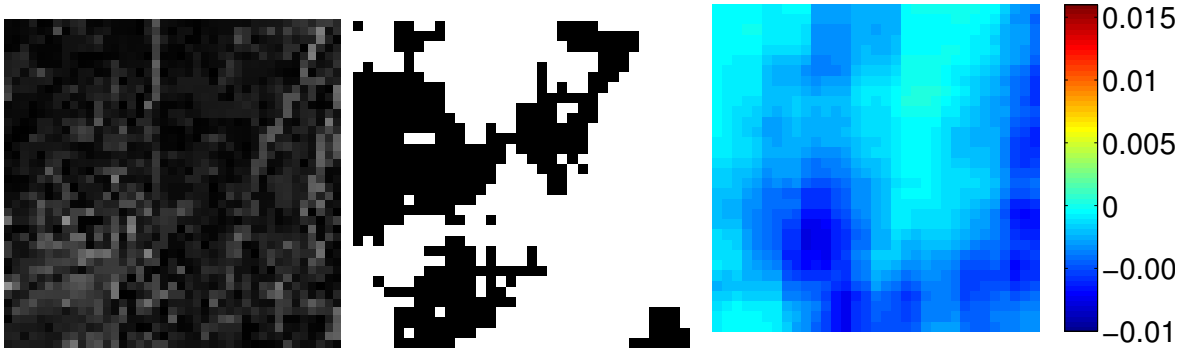
		DIIVINE	BRISQUE	MetricQ	Anisotropy	RR-IQA	**CQ**	**CDQ**
LIVE	median of all SSIM differences	1.59×10^{-2}	2.03×10^{-2}	2.42×10^{-2}	8.88×10^{-2}	4.28×10^{-2}	2.97×10^{-3}	**0**
	average of all SSIM difference	3.45×10^{-2}	3.57×10^{-2}	5.07×10^{-2}	1.09×10^{-1}	5.45×10^{-2}	9.91×10^{-3}	$\mathbf{7.76 \times 10^{-3}}$
	average of non-outliers	2.59×10^{-2}	2.42×10^{-2}	5.07×10^{-2}	1.09×10^{-1}	5.45×10^{-2}	7.02×10^{-3}	$\mathbf{2.07 \times 10^{-3}}$
CSIQ	median of all SSIM difference	3.63×10^{-2}	3.43×10^{-2}	2.44×10^{-2}	3.66×10^{-2}	4.30×10^{-2}	$\mathbf{1.73 \times 10^{-3}}$	$\mathbf{1.73 \times 10^{-3}}$
	average of all SSIM difference	5.19×10^{-2}	4.97×10^{-2}	4.25×10^{-2}	4.30×10^{-2}	4.77×10^{-2}	$\mathbf{1.11 \times 10^{-2}}$	1.12×10^{-2}
	average of non-outliers	4.72×10^{-2}	3.49×10^{-2}	2.19×10^{-2}	3.81×10^{-2}	4.77×10^{-2}	6.02×10^{-3}	$\mathbf{5.38 \times 10^{-3}}$

algorithms is used to evaluate other IQA methods in this experiment. The SSIM difference of each IQA algorithm is plotted in Fig. 3.12. More quantitative evaluation of different IQA algorithms are provided in Table 3.1. Both the results in Fig. 3.12 and Table 3.1 show that the comparison-based methods, CQ and CDQ, have the best accuracy of selecting reconstruction parameters.

Parameter selection for bilateral filter

A series of increasingly denoised images, $I_1, I_2, \cdots, I_{30}$, are created for each image the same as Section 3.4.1. The SSIM index of the most oversmoothed image I_{30} is between 0.85 ± 0.01.

Because the MSD between the adjacent images are below minimum resolution of bilateral filtering, Alg. 7 is adopted to select the best result. Key images are a set of images among which the MSD is greater than the minimum resolution. Alg. 7 first separates the 30 increasingly denoised images into a few segments by key images and selected the best key image with



(a) Patches difference (b) Content Detection (c) Contribution difference

Figure 3.11: CQ index of Fig. 3.9 (f) (β_4) based on Fig. 3.9 (e) (β^*). (a) is the patch difference between the image with β_4 and β^*. The white areas in (b) corresponds to structured area. In (c), the negative values means that the contribution comes from β^*. On this local patch, $CQ(P_{\beta_4}, P_{\beta^*}) = -2.50 \times 10^{-4}$.

Table 3.2: Accuracy of parameter selection for bilateral filter

		DIIVINE	BRISQUE	MetricQ	Anisotropy	RR-IQA	**CQ**	**CDQ**
LIVE	median of all SSIM differences	1.65×10^{-2}	2.89×10^{-2}	$\mathbf{2.36 \times 10^{-3}}$	6.95×10^{-2}	7.98×10^{-3}	7.67×10^{-3}	3.80×10^{-3}
	average of all SSIM differences	2.53×10^{-2}	3.31×10^{-2}	6.73×10^{-3}	8.23×10^{-2}	1.07×10^{-2}	1.43×10^{-2}	$\mathbf{6.05 \times 10^{-3}}$
	average of non-outliers	2.53×10^{-2}	3.31×10^{-2}	4.63×10^{-3}	8.23×10^{-2}	1.07×10^{-2}	1.12×10^{-2}	$\mathbf{3.93 \times 10^{-3}}$
CSIQ	median of all SSIM differences	1.59×10^{-2}	3.84×10^{-2}	$\mathbf{2.83 \times 10^{-3}}$	2.18×10^{-2}	7.22×10^{-3}	5.28×10^{-3}	4.05×10^{-3}
	average of all SSIM differences	3.40×10^{-2}	3.86×10^{-2}	7.65×10^{-3}	4.30×10^{-2}	2.41×10^{-2}	1.09×10^{-2}	$\mathbf{6.24 \times 10^{-3}}$
	average of non-outliers	3.40×10^{-2}	3.86×10^{-2}	6.34×10^{-3}	3.17×10^{-2}	9.75×10^{-3}	6.36×10^{-3}	$\mathbf{5.32 \times 10^{-3}}$

CQ/CDQ. The MSD difference between key images are lower bounded by K_{thresh}. Next, images in the two segments that are adjacent to the best key image are evaluated based on the two key images on the ends. By doing so, we avoid comparing the adjacent images directly. K_{thresh} is set as 3.0 in this experiment.

The SSIM difference between the best denoised image chosen by SSIM and the one chosen by other IQA methods is plotted in Fig. 3.13. Table 3.2 shows more quantitative results of different IQA algorithms. Both CDQ and MetricQ give satisfying results for bilateral denoising parameter selection.

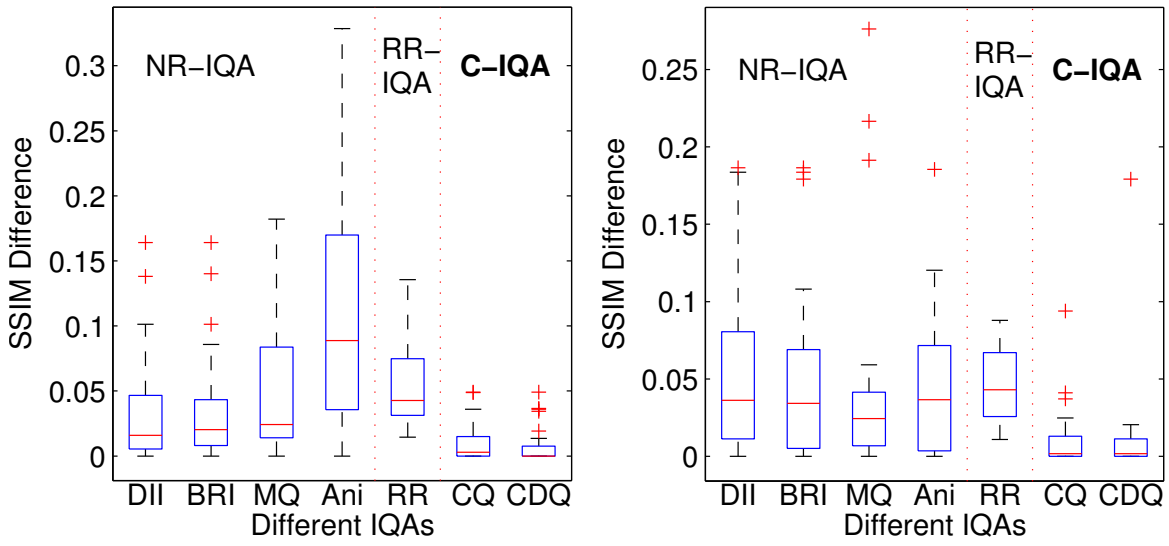
In order to better analyze the performance of comparison-based methods, two of the outliers of denoising parameter selection by C-IQA are shown in Fig. 3.14. Comparison-based methods tend to choose the parameters that lead to over-smoothed denoised results. This is a common challenge for all the NR-IQA algorithms in our experiments. For the lack of texture complexity information from the original image, NR-IQA algorithms are easy to confuse the fine texture

Algorithm 7 Key Image

Key Images Selection;
$key\_img = \{1\}$
$key_{num} = 1$
for $i = 1 : N$ **do**
  if $MSD(I_i, I_{pre_{key}(key_{num})}) > K_{thresh}$ **then**
    $key\_img = \{key\_img, i\}$
    $key_{num} = key_{num} + 1$
  end if
end for

Key Images Comparision;
for $i = 2 : (key_{num} - 1)$ **do**
  if $C - IQA(I_{key\_img(i)}, I_{key\_img(i-1)}) > 0$ and
    $C - IQA(I_{key\_img(i)}, I_{key\_img(i+1)}) > 0$ **then**
    $best_{key} = i$
    break;
  end if
end for

Best Image Selection;
$start_{num} = key\_img(best_{key} - 1)$
$end_{num} = key\_img(best_{key} + 1)$
for $i = start_{num} : end_{num}$ **do**
  $score_{start}(i) = C - IQA(I(i), I(start_{num}))$
  $score_{end}(i) = C - IQA(I(i), I(end_{num}))$
end for
$best_{img} = max(score_{start} + score_{end})$



(a) LIVE (b) CSIQ

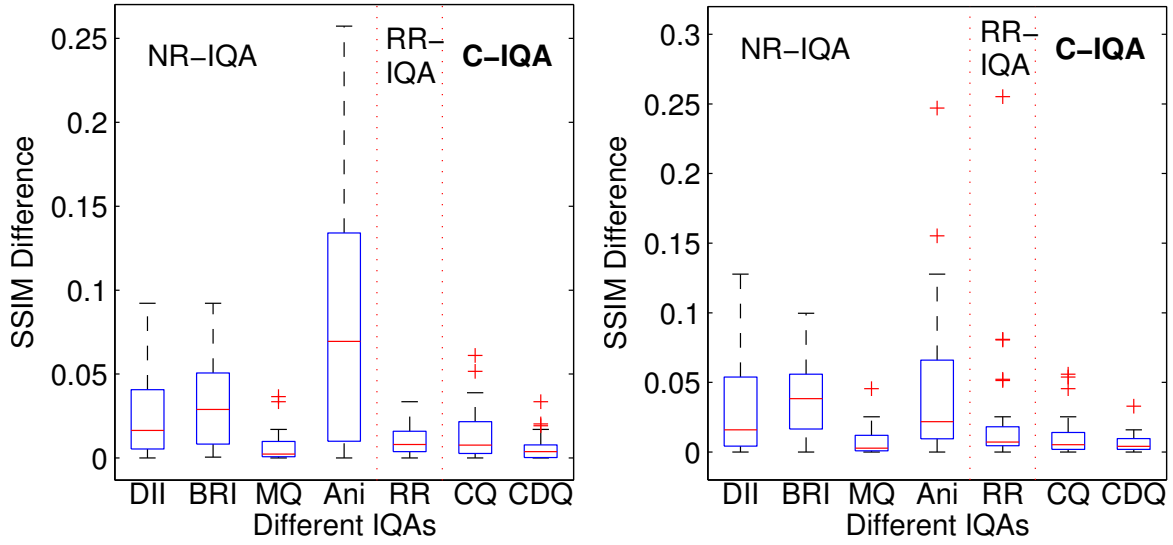
Figure 3.12: The SSIM differences of the best images chosen by different NR, RR, and Comparison-based IQA methods for image reconstruction parameter selection.

with noise component. On the contrary, the RR-IQA in [5] is good at handling images with different global texture complexity because an index that indicates the energy in the disorder part in the original image is available for the image quality assessment.

Parameter selection for BM3D

Similar to the settings of parameter selection for bilateral filter, a series of increasingly denoised images, $I_1, I_2, \cdots, I_{30}$, are created with BM3D [110]. The SSIM index of the most oversmoothed image I_{30} is between 0.85 ± 0.01. Alg. 7 is adopted to select the best result.

The SSIM difference between the best denoised image chosen by SSIM and the one chosen by other IQA methods is plotted in Fig. 3.15. Table 3.3 shows more quantitative results of different IQA algorithms. It should be noticed that the performance of MetricQ decreases significantly compared with the results of bilateral filters. The reason is that BM3D mainly removes the subtle details as the parameter increases, while MetricQ takes these details as noise. On the contrary, bilateral filters blur major structures as well when the parameter increases. The results of MetricQ and C-IQA on BM3D also reveal that although the implemented C-IQA



(a) LIVE (b) CSIQ

Figure 3.13: The SSIM differences of the best images chosen by different NR, RR, and Comparison-based IQA methods for the bilateral filter parameter selection.

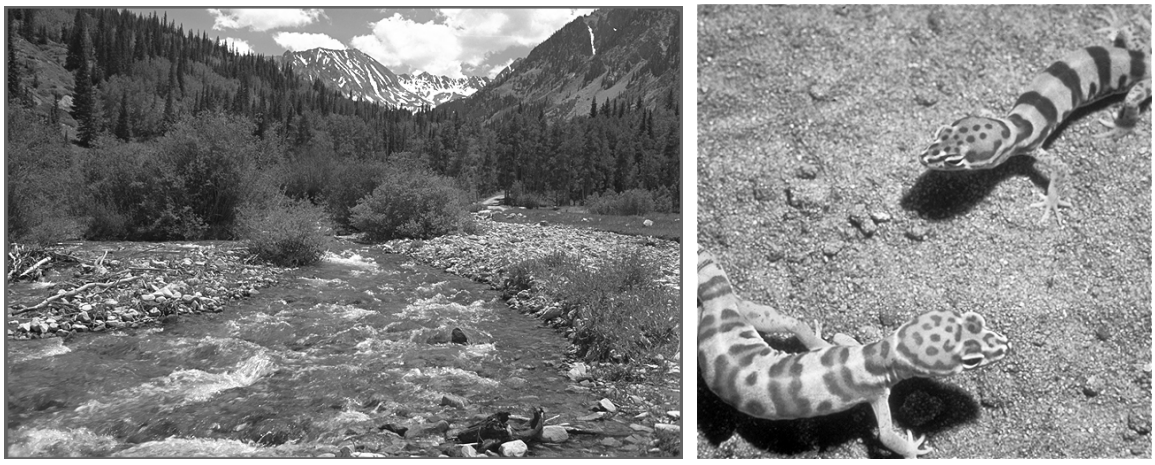
Table 3.3: Accuracy of parameter selection for BM3D

		DIIVINE	BRISQUE	MetricQ	Anisotropy	RR-IQA	**CQ**	**CDQ**
LIVE	median of all SSIM differences	1.25×10^{-2}	$\mathbf{5.45 \times 10^{-3}}$	7.15×10^{-2}	7.43×10^{-3}	7.43×10^{-3}	1.67×10^{-2}	6.23×10^{-3}
	average of all SSIM differences	1.94×10^{-2}	9.50×10^{-3}	6.43×10^{-2}	1.74×10^{-2}	1.02×10^{-2}	2.24×10^{-2}	$\mathbf{7.52 \times 10^{-3}}$
	average of non-outliers	1.11×10^{-2}	7.25×10^{-3}	6.08×10^{-2}	1.23×10^{-2}	9.45×10^{-3}	1.70×10^{-2}	$\mathbf{6.16 \times 10^{-3}}$
CSIQ	median of all SSIM differences	2.33×10^{-2}	7.25×10^{-3}	6.05×10^{-2}	4.34×10^{-3}	4.57×10^{-3}	8.00×10^{-3}	$\mathbf{8.41 \times 10^{-4}}$
	average of all SSIM differences	2.79×10^{-2}	1.72×10^{-2}	5.49×10^{-2}	9.16×10^{-3}	5.78×10^{-3}	2.05×10^{-2}	$\mathbf{3.05 \times 10^{-3}}$
	average of non-outliers	2.43×10^{-2}	9.48×10^{-3}	5.49×10^{-2}	6.67×10^{-3}	4.88×10^{-3}	1.62×10^{-2}	$\mathbf{1.94 \times 10^{-3}}$

makes use of MetricQ, the performance of C-IQA is significantly improved based on MetricQ due to the comparison framework.

3.4.3 Application in parameter trimming

In this section, we combine CDQ with the parameter trimming framework and show that considerable computation can be saved while preserving the accuracy of parameter selection. In this part, all the parameter settings for image reconstruction are the same as the settings in Section 3.4.2. Fig. 3.16(a) shows one example image in parameter trimming. The SSIM indexes in Fig. 3.16(b) and (c) are only used to demonstrated the convergence process. Fig. 3.16(b) shows the parameter selection after all the algorithm instances with different parameters



(a) "stream" (LIVE) (b) "geckos" (CSIQ)

Figure 3.14: Outliers of parameter selection

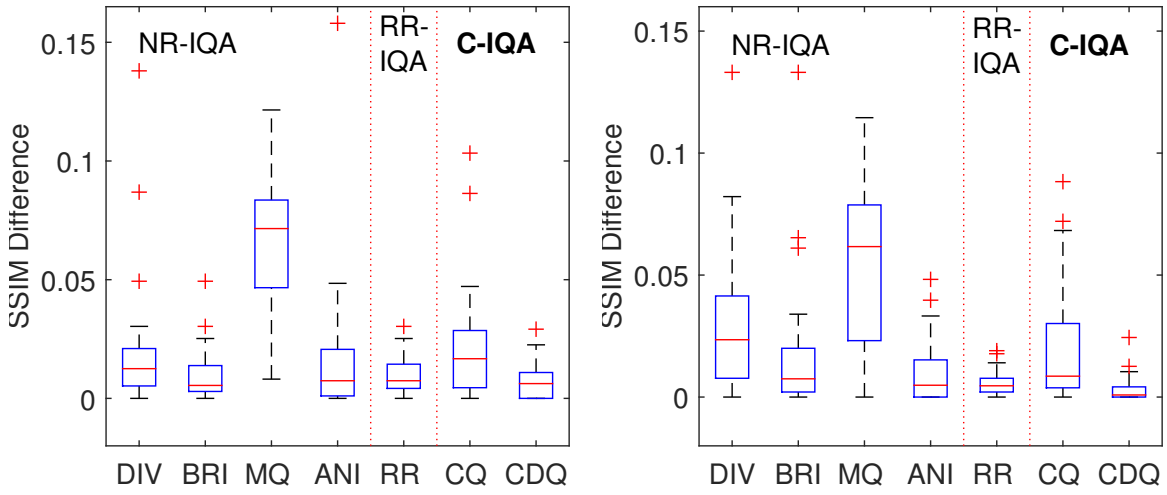
Table 3.4: Computation saved by parameter trimming

	Ave # of iteration without parameter trimming	Ave # of iteration with parameter trimming	saved computation (%)
LIVE	4651.9	847.7	81.78
CSIQ	4565.6	941.1	79.39

converge. Fig. 3.16(c) illustrates the parameter trimming process with CDQ. From Fig. 3.16, we can see that the trimming decision based on CDQ achieves the goal of terminating the iteration of parameters that are far from the best choice. On LIVE [2], all the parameters selected with parameter trimming are the same as the parameters selected after convergence; on CSIQ [3], only one of the best parameters selected by parameter trimming is different from the one selected after convergence. From Table 3.4, it is clear that considerable computation is saved with parameter trimming.

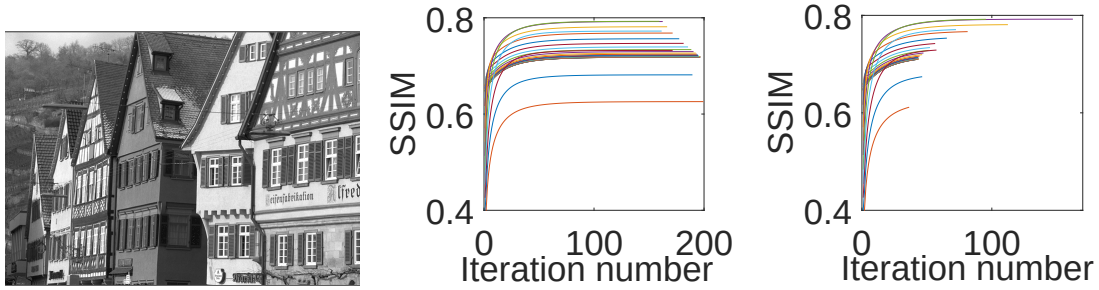
3.5 Discussion

Motivated by the parameter selection for image restoration algorithms, for the first time we proposed a comparison-based IQA framework. The comparison-based method implemented in this work includes three primary modules, Content Detection, Contribution and Distortion



(a) LIVE (b) CSIQ

Figure 3.15: The SSIM differences of the best images chosen by different NR, RR, and Comparison-based IQA methods for the BM3D parameter selection.



(a) "buildings" (from LIVE) (b) Convergence without parameter trimming (c) Convergence with parameter trimming

Figure 3.16: Comparison between convergence with and without parameter trimming on "buildings"

Sensitivity Compensation. One important property that is unique to comparison-based IQA, minimum resolution, is analyzed. At last, the comparison-based IQA compares favorably with other NR-IQA and RR-IQA algorithms on two widely used databases for image reconstruction and bilateral filter parameter selection.

With the proposed parameter selection method by comparison-based image quality assessment, both enhancement of the microscopy data and reconstruction of the MRI data result in high quality images, and further benefit visualization and analysis of these image data.

Chapter 4

Cell Detection for Reconstructed Brains

With high resolution microscopy data, cell analysis for brain tissues are discussed in this chapter. The cell number and distribution pattern reflect many underlying biomedical mechanisms. In SE, the activation topology of granule cells in the dentate gyrus is an important topic. Recently, data-driven methods receive more attention [8, 111, 112] for two reasons. First, the success of deep learning in the area of object classification and detection [113] provides practical experience on the architecture design and training of convolutional neural networks (CNN) that can be adapted to other domains. Packages such as Pytorch [114], and regularization techniques such as dropout [115] and residual convolution [116] make building customized CNNs easier. Secondly, huge amounts of imaging data are generated with advanced microscopy technology. Models trained with more imaging data are more robust and accurate.

Because microscopy data differ from natural scene images in many ways, transferring solutions for natural scene images to cell analysis should be performed carefully. For cell analysis, the objects of interest, cells, usually are of a smaller size compared with the objects in natural scene images. Very deep network architectures are not necessary for cell detection [112], since the typical size of a single cell is within 50×50 pixels in microscopy images. The benefit

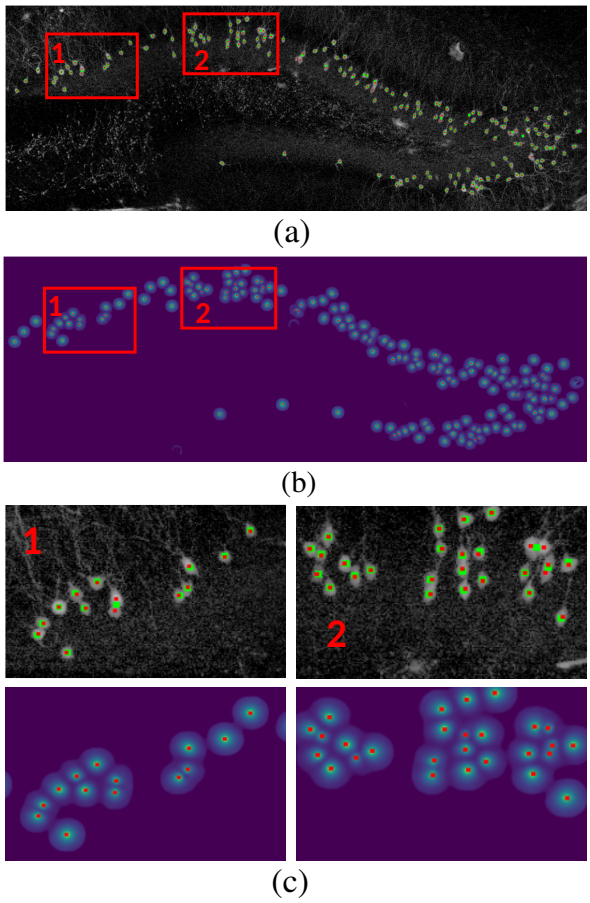


Figure 4.1: Cell detection with the proposed repel coding for the granule cells in mouse brain tissues. (a) Input image. (b) Output of the network. (c) Zoomed-in patches. Green dots are the labeled cell centers, and red dots are the detected cell centers.

of designing a deeper network is marginal if the receptive field is large enough to cover a single cell. Deeper networks also have more trainable parameters, and this could even hurt the performance by over-fitting. In biomedical research, cell analysis is a highly customized task. The morphology could change dramatically among different types of cells. Even for the same type of cell, the appearance changes with different tissue preparation protocols, imaging equipment and imaging protocols. Instead of training an omnipotent cell analysis framework, the ability to adapt the analysis model to a specific cell type and imaging modality is desirable in practice.

The success of deep learning on numerous computer vision tasks is widely perceived as

a result of improved learning architectures and big data. CNNs and many regularization techniques, such as max out [117], drop out [115] and residual convolution [116], are practical components for building neural networks. The other factor, data enhancement, is discussed less especially for the output data. In this paper, instead of focusing on the data enhancement at the input side of a learning network, the coding scheme at the output side is discussed. Compared with the input data, the output data has more freedom to be designed to suit our applications. For cell analysis, one type of the raw labels is a sparse 2D picture where non-zero pixels indicate centers of cells. The widely used cell density function for cell counting can be taken as an augmentation of the output data [6]. The cell density usually is a smoothed version of the raw dot labels. Similarly, cropping the whole images into patches can be taken as filtering the raw dot labels with a rectangle filter.

To better understand the effect of output data enhancement, two criteria of raw label coding are discussed in this paper: entropy and reversibility. The entropy measures how suitable the coding scheme is for training, and the reversibility measures how robust inverting the coding to raw labels is when predicting. Different raw label coding methods are trading-off between these two criteria. For applications with different variations of shape, size and crowdedness of the cells, the coding scheme should be designed accordingly. The coding scheme with both high entropy and high reversibility is always preferred. However, a trade-off between these two indexes has to be made in most cases depending on the training loss and the overall performance. For example, if the training loss with a coding scheme converges fast, but the accuracy is not ideal after inversing the outputs to the raw data, then a coding scheme with lower entropy but higher reversibility should be considered. We also propose a new coding scheme that balances these two criteria well in most cases. Fig. 4.1 shows one example of granule cell detection in the mouse brain with the proposed coding scheme. However, it should be emphasized that we are not proposing an optimal coding scheme for all types of cells. The coding scheme should be carefully designed based on each application.

The rest of this paper is organized as follows. Section 4.1 reviews recent cell counting and detection works with CNNs, and highlights one common step in these works: raw dot label

enhancement. Section 4.2 details the proposed coding evaluation criteria, and provides a cell center coding scheme based on the design criteria. Section 4.3.2 verifies the proposed coding method with four cell datasets and two network architectures. At last, Section 4.4 reviews the proposed method, and discusses future work in cell analysis.

4.1 Related works

4.1.1 Object Detection in Natural Scenes and Cell Analysis

Object detection for natural scene images usually outputs a group of bounding boxes to represent the location and the size of detected objects. Both R-CNN [118], YOLO [119] and their variants adopt this box representation. Some widely used detection benchmarks also use the box labeling, such OTB [120] and COCO [121]. By representing an object with a bounding box by four numbers, the number of output nodes in a detection network is significantly reduced. For cell analysis, the center dots and contours of cells are more common labeling formats rather the bounding box. Since labeling contours for training costs more time, dot labeling is used more often if the cell morphology is not of particular importance. In this paper, we focus on the dot labeling for cell counting and detection.

Among the learning-based cell detection works, the raw dot labels usually are pre-processed before being set as the output data for training. Cropping whole images into small patches [9,112], and transforming the dot labels into a density representation [6, 7] are two common approaches.

Designed for cell counting [112], the inputs of the CNN are patches of size 60×60 pixels, and the output is the total number of cells in this patch. The CNN treats number estimation as a regression problem rather than a classification one. The reason is that the appearance difference between patches with similar number of cells should be smaller than that between patches with great disparity in cell numbers [112]. By cropping a whole image into multiple patches, the quantity of training examples is greatly enhanced. With the implementation of fully convolutional networks [122, 123], another way to understand the pipeline of cropping and counting is that the raw dot labels are filtered with a moving average filter of size 60×60. In

Walach's work [7], raw dot labels convolved with a Gaussian kernel are used for cell detection, and raw dot labels convolved with a human-shape kernel are used for pedestrian detection.

Cell counting and detection are implemented by a CNN followed by a compressive sensing module [124]. The input to the CNN is a patch of size 200×200 containing multiple cells, and the output is a vector y that contains compressed center location information. The compressed location information is computed by a learned sensing matrix S with $y = Sx$, where x is a vector of raw dot labels. The length of y is much less than the length of x. As a result, decoding with a sparse prior is used to recover the exact position at prediction. The concept of coding the raw dot labels is emphasized [124], but the coding with compressive sensing is used less than simple spatial filter codings for two reasons. First, the sensing matrix, S, is an extra mapping relation learned by the neural network. Because of the size of S, a large amount of training data is required to avoid over-fitting and maintain good accuracy. Second, recovering the location information from the compressed vector can be time-consuming.

Unlike coding the raw dot labels by convolving with a filter [6, 112], proximity coding [9] is defined as,

$$C_{ij} = \begin{cases} \frac{1}{1+\alpha D_{ij}}, & \text{if } D_{ij} < r, \\ 0, & \text{otherwise}, \end{cases} \tag{4.1}$$

where D_{ij} is the Euclidean distance from the pixel (i, j) to the closest cell center. Because proximity coding preserves local maxima at cell centers, such coding can be used for cell detection as well [8].

From recent works on cell analysis, there are two observations. First, transformation of the raw dot labels is necessary before training the network. As a result, a corresponding inverse transformation is required at the prediction phase. For cell counting, integration over the outputs serves this purpose, and for cell detection, local maximum detection is used. Second, cell counting and cell detection share many common components. The tasks of cell counting and detection share the same pipeline [8]. A more detailed comparison between cell counting and detection is provided in the next part.

4.1.2 Counting vs. Detection

Object counting usually is preferred over detection in applications where objects are crowded and single objects are not distinguishable. In this case, detection will not provide accurate location information, and texture information is a crucial clue to estimate the object density [125, 126]. However, if each single object can be identified, the detection approach has more advantages. First, detection provides the location information lost in counting. Second, more complexity is involved in counting than detection for training. For the applications of cell analysis, usually the size of cells is much smaller than the whole image. This indicates that a single cell can be recognized with a smaller receptive field than the whole image required by counting. Larger receptive fields of CNNs usually mean more trainable parameters and deeper networks. As mentioned before, cell analysis is a highly customized application, and retraining a model to fit a specific imaging modality and cell morphology is more effective than an overly general model. For this reason, cell detection enables a smaller network design that is easier to train with limited labeled data. In general, detection is a more appropriate approach if objects are not heavily occluded and each single object is recognizable. The major challenge of detection is when objects are densely packed or partially occluded. In the next section, two criteria are proposed for the raw dot label coding for cell detection, and a new raw dot coding method, repel coding, is proposed to better tackle this challenges.

4.2 Proposed coding scheme

Table 4.1: Entropy and Reversibility of Different Coding Schemes

	Dot labels	Gaussian kernel	Avg. kernel	Proximity coding	**Repel coding**
E	0.0000	6.761	2.052×10^{-1}	4.901	6.472
R	1.000	4.583×10^{-3}	4.434×10^{-3}	1.198×10^{-2}	8.199×10^{-3}
R^5	1.000	1.089×10^{-1}	1.111×10^{-1}	1.286×10^{-1}	1.306×10^{-1}

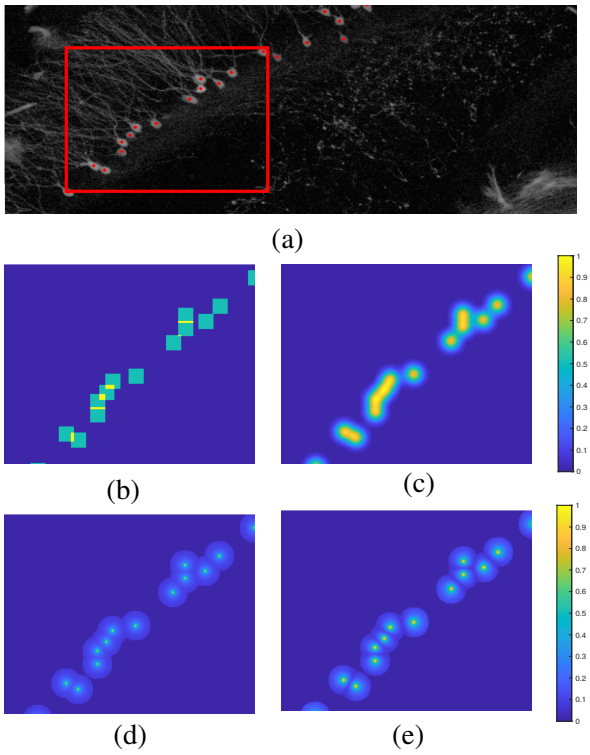


Figure 4.2: Illustrations of different coding schemes. (a) An example image with cell centers labeled with red dots. The rectangle in (a) indicates the zoomed-in patches in (b)-(e). (b) shows the coding with the rectangle kernel [6]. (c) shows the coding with the Gaussian kernel [6,7]. (d) shows the coding with proximity coding [8,9]. (e) is the proposed repel coding. The proposed center coding scheme is superior to proximity coding in two ways. First, the repel coding provides a slower response decay away from the cell center, which is helpful to the stability during training. Secondly, the repel coding suppresses the responses between two cell centers. This is important to recover the raw dot labels during prediction. Coding values in (b) to (e) are normalized to 0-1 for comparison.

4.2.1 Criteria

The two proposed criteria for the center coding are entropy and reversibility. At the training phase, entropy characterizes if the coding scheme is easy for the neural network to learn. At the prediction phase, reversibility measures if the coding scheme can recover the raw dot labels robustly.

Entropy The entropy of a coding scheme, C, is defined as

$$E_C = entropy(C_{ij}\ if\ C_{ij} \neq 0), \tag{4.2}$$

where C_{ij} represents the coded value at position (i, j). Zero values in the transformed coding are excluded here, since these regions usually are far from any cell. Entropy measures how evenly the non-zero values are distributed. An ideal coding scheme should distribute the coding values uniformly over a range. By doing so, the gradient backpropagation during the training phase is more robust. A similar concept is mentioned by modeling the counting problem as regression rather than classification [112]. The extreme case of low entropy coding is the raw dot labels, where the entropy is always zero.

Reversibility The reversibility of a coding C is defined as,

$$R_C = \frac{\sum_{i,j} M_{ij} \cdot C_{ij}}{\sum_{i,j} C_{ij}}, \tag{4.3}$$

where M is the mask defining the proximity region of cell centers. The binarized raw dot labels, or the dilated version of raw dot labels can be used for M. In the prediction phase, the output response is not identical to the ideal coding scheme. A robust coding scheme should be able to recover the original coding, raw dot labels, in challenging cases. Reversibility is a similarity measurement between the raw dot labels and the coded response. Because local maximum detection is used to recover the raw dot labels at the prediction phase, reversibility here is defined as the degree of energy concentration around the raw dot labels.

For cell detection, a coding scheme with large entropy and reversibility indexes is preferred. As an extreme case, the dot label itself has the maximum reversibility index. However, the raw dot label has the smallest entropy index. This means it is hard for the neural network to learn raw dot labels. On the other hand, coding by the Gaussian kernel has a larger entropy but a lower reversibility index. The result is that networks trained with coding by a Gaussian kernel converge fast and robustly in terms of loss value, but center recovery is obscured in the

prediction phase. More analysis on the entropy and reversibility trade-off is illustrated with experimental results in Section 4.3.2.

4.2.2 Repel Coding

The proposed coding scheme of raw dot labels is based on proximity coding defined in Eqn. 4.1. When proximity coding was first proposed [9], it was designed for cell counting. Because proximity coding produces local maxima at cell centers, it was also used for cell detection later [8]. However, one common challenge for cell detection is to distinguish two neighboring cells. For cell counting, only a global counting number is required. In other words, only the entropy is considered when coding raw dot labels for cell counting, but not the reversibility. In practice, we notice that proximity coding does not perform well for detection when cells are crowded. The response valley between two cells is not significant enough, and local maxima do not align with cell centers accurately during the prediction. Aiming at increasing the reversibility of proximity coding, the proposed repel coding is defined as,

$$
\begin{aligned}
D'_{ij} &= dist^1_{ij} \times \left(1 + dist^1_{ij}/dist^2_{ij}\right)^2, \\
C_{ij} &= \begin{cases} \frac{1}{1+\alpha D'_{ij}}, & \text{if } D'_{ij} < r, \\ 0, & \text{otherwise}, \end{cases}
\end{aligned}
\tag{4.4}
$$

where $dist^1_{ij}$ is the distance of the pixel (i, j) to its nearest cell center, and $dist^2_{ij}$ is the distance of the pixel (i, j) to its second nearest cell center. The intermediate variable D'_{ij} can be taken as $dist^1_{ij}$ suppressed by $dist^2_{ij}$.

In Fig. 4.2, examples of different coding schemes are illustrated. Comparing Fig. 4.2 (d) and Fig. 4.2 (e), it is obvious that the proposed repel coding forms a more significant valley between two neighboring cells than proximity coding. Table 4.1 provides the entropy and the reversibility of different coding schemes shown in Fig. 4.2. The entropy, E in Table 4.1, is calculated by separating the coded non-zero values into eight bins. The reversibility, R in Table 4.1, is calculated by using the raw dot labels as M in Eqn. 4.3. Because it is rare in

practice that the centers of two cells are 1 pixel away, the dilated reversibility, R^5 in Table 4.1, is calculated by dilating the raw dot labels with a disk of diameter of 5 pixels. The meaning of the entropy and the reversibility in Table 4.1 can be interpreted by comparing with the illustrations in Fig. 4.2. The coding scheme with the highest entropy is the Gaussian kernel, and the entropy of the proposed repel coding is slightly less than that of the Gaussian kernel. With visual inspection, the intensity variations of the Gaussian kernel and the repel coding are larger than the other codings. Measured by R^5, the proposed repel coding achieves the highest reversibility index except for raw dot labels. This also aligns with the visual inspection where cell centers with the repel coding are more prominent than those with proximity coding.

4.2.3 Relation to existing works

Besides different coding schemes for cell analysis, coding of the raw labeled data is also widely adopted for computer vision tasks with natural scenes. The anchor box introduced in YOLO v2 [127] resembles convolution kernels with different shapes [7]. The watershed transformation [128] for semantic segmentation is similar to the proposed repel coding. The difference is that in watershed transformation, boundary information is of interest, while for cell detection, center information is the final output. A two-step coding scheme [128] inspired by the watershed algorithm is effective in many semantic segmentation applications. The first step involves coding object boundaries. In this step, the coded response at each pixel is a two dimensional unit vector that points to the closest boundary pixel. The coding in the first step aims at maximizing the reversibility of the raw labels. In the second step, the coded response at each pixel is the distance from a pixel to its closest boundary pixel. The second step focuses more on entropy maximization. These two steps are cascaded in the watershed transformation pipeline.

In general, different coding schemes are different ways to transfer an end-to-end training framework to a stepwise implementation. Another way to understand coding the raw dot labels is taking the neural network as a signal processing system. As in the analog domain, designing filters with shape responses such as the unit impulse response is challenging. By coding the

raw dot labels, we impair the ideal response by smoothing it, but such smoothing is preferred sometimes because of its easier implementation.

4.3 Experiments

In this part, different coding schemes are tested for four types of cells and with two CNN architectures. Experimental results show that the proposed repel coding outperforms existing coding schemes both for cell counting and detection tasks. Discussion of the examples from the four types of cells provides some insights into different coding schemes.

4.3.1 Datasets

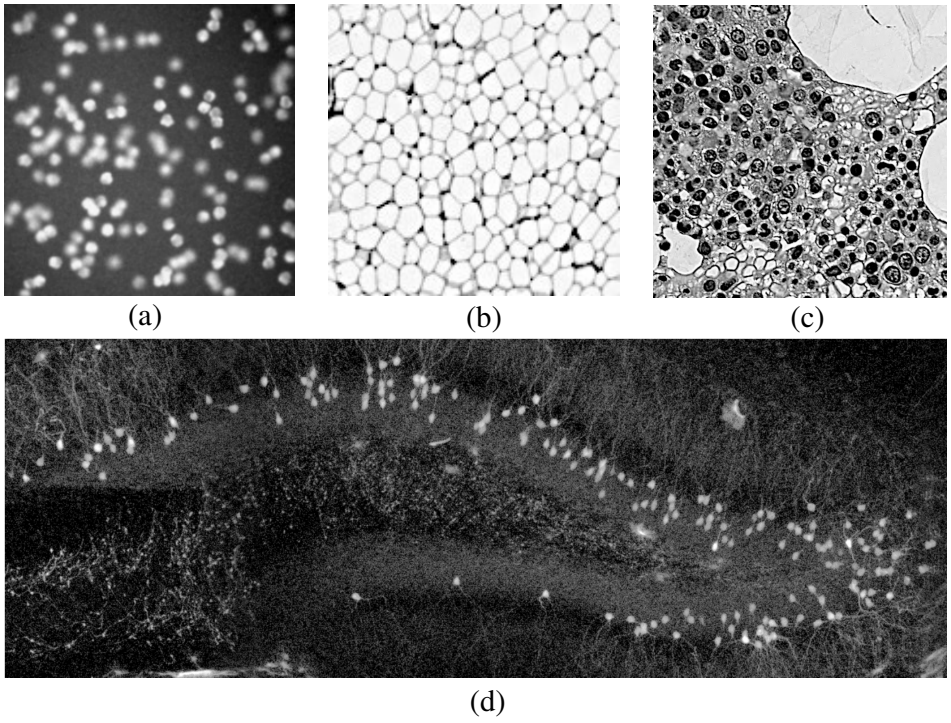


Figure 4.3: Sample images from the four datasets. (a) Synthetic Vgg cells. (b) Adipocyte cells. (c) Human bone marrow cells. (d) Granule cells in the mouse dentate gyrus.

Four datasets are evaluated in the experiments: granule cells in the mouse dentate gyrus (DG), human adipocyte cells (Adip), human bone marrow cells (HBM) and Vgg-generated synthetic cells (Vgg). Fig. 4.3 shows examples from these four datasets.

DG dataset

The DG dataset comprises mouse brain tissues stained by tdTomato after seizure. These brain tissues include the dentate gyrus in the V-shape, and the highlighted cells are granule cells. A Zeiss 780 confocal microscope with a C-Apochromat objective under 10X magnification is used to image the brain tissues. The DG dataset contains 26 high resolution dentate gyrute images, and the image size is from 452×942 to 732×1336. More details about the DG dataset can be found in our previous work [129].

Adip dataset

The Adip dataset [6] contains human subcutaneous adipose tissues obtained from the genotype tissue expression consortium. The available images of Adip dataset are 150×150, and the size of each single cell is within 32×32. These adipocyte cells are densely-packed as shown in Fig. 4.3 (b). The Adip dataset contains 200 images.

HBM dataset

The HBM dataset [130] includes images of healthy human bone marrow from eight different patients. The HBM dataset contains 44 images of size 600×600, and the objects of interest are the stained nuclei shown in Fig. 4.3 (c). Tissues in HBM dataset are stained with Hematoxylin and Eosin.

Vgg dataset

The Vgg dataset [131] is a group of synthetic cells. With a cell simulation method [132], cell detection algorithms can be easily evaluated with different imaging settings, such as cell overlap, out-of-focus blurring, and size variation. The Vgg dataset contains 200 images of size 256×256.

4.3.2 Experimental settings

Two network architectures are tested in these experiments. The training settings, including the cost function, the learning rate, and the optimization algorithm are kept the same through different experiments. The learning rate is set to 10^{-5}, the optimizer is Adam with default parameters [133], and the training batch size is set as eight for HBM, Adip and Vgg datasets, and two for the DG dataset, to fit in the 12GB memory of the NVIDA TITAN Xp used in the experiments.

Cost function

To be consistent with the previous works on cell counting and detection [7, 112, 124], the $L2$ norm is used as the cost function for training,

$$cost = \|y - y'\|_2^2,$$

where y is the ground truth output coding generated from the raw dot labels, and y' is the output of the CNN. Fig. 4.2 illustrates the four coding schemes.

Network architectures

Two CNNs based on the Unet [111] architecture are evaluated in the experiments. One CNN is the same as the FCNN-A [8], and the other one replaces the convolutional layers in the FCNN-A with residual convolution blocks. The overall architecture of the FCNN-A is shown in Fig. 4.4. The activation function used in all layers is rectified linear unit (ReLU). The receptive fields of these two CNNs are both 38×38.

Evaluation

Since coding schemes for both counting and detection are compared in the experiments, two measures are adopted here. For codings with the Gaussian kernel and the rectangle kernel, the integration over the outputs is taken as the total number of cells. For raw dot coding,

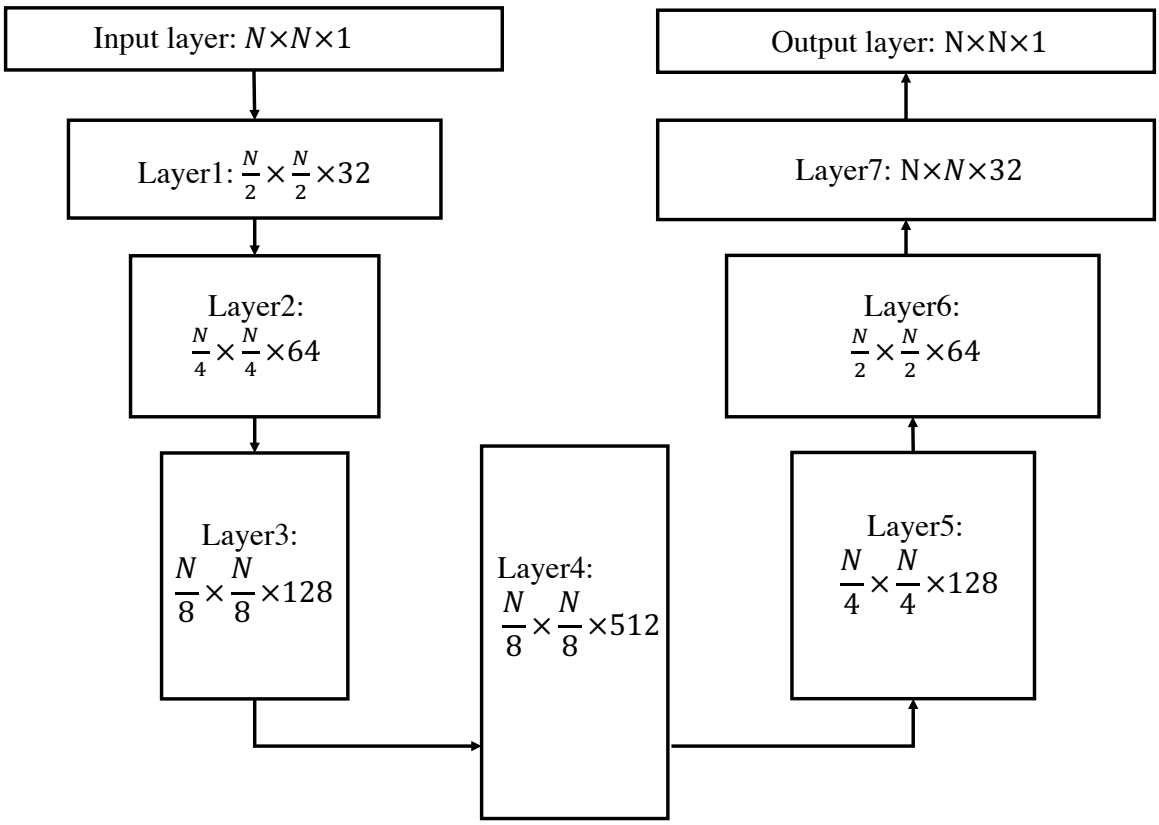


Figure 4.4: The architecture of the CNN in this experiment is based on a U-net with eight layers. If all layers are implemented with plane convolutional blocks, the network is the same as the FCNN-A [8].

proximity coding, and repel coding, cell centers are extracted by local maximum detection. The F1 score [15] is used as a comprehensive index to evaluate the detection accuracy. Alg. 8 summarizes the F1 score calculation. In Alg. 8, each non-paired detected cell center is matched to the closest non-paired ground truth cell center. Cell centers in the detected list, D_{list}, and the ground truth list, G_{list}, can be paired only once. A match with the distance less than the average radius of the cell is considered to be a successful match since the cell size variant within a dataset is not much. The average radius is 8 pixels for the DG dataset, 11 pixels for the Adip dataset, 15 pixels for the HBM dataset, and 11 pixels for the Vgg dataset.

Algorithm 8 F1 Score for Evaluation

Inputs:

D_{list}: List of detected centers

G_{list}: List of ground truth centers

Outputs:

$F1$: F1 score

Initialization:

1. $D_{matched} = [inf, ..., inf]$
2. $G_{matched} = [inf, ..., inf]$

Pair Match:

while D_{list} is not empty and G_{list} is not empty **do**

3. $D_{idx}, G_{idx}, dist = get\_closest\_pair(D_{list}, G_{list})$
4. $D_{matched}(D_{idx}) = dist$
5. $G_{matched}(G_{idx}) = dist$

end while

6. $Acc = \frac{\#(D_{matched} \leq thresh)}{\#D_{list}}$
7. $Rec = \frac{\#(G_{matched} \leq thresh)}{\#G_{list}}$
8. $F1 = \frac{1}{\frac{1}{Acc} + \frac{1}{Rec}}$

4.3.3 Results and analysis

With the four datasets, five coding methods, and two CNN implementations, 40 sub-experiments are evaluated. Each sub-experiment tests one coding method with one CNN implementation on one dataset. For each dataset, 80% percent of the data are used for training, and 20% are used for testing. Each sub-experiment is run five times with random training/test splitting, and the average performance is reported in Table 4.2 and Table 4.3. The proposed repel coding achieves the best performance in most sub-experiments. The only exception is the sub-experiment on the Vgg dataset with the FCNN-A implementation, where the performance of raw dot labels is slightly better than the proposed method. The reason may be that the cell variance in the Vgg dataset is less than in the other datasets, and thus is a less challenging dataset. Comparing the results in Table 4.2 and Table 4.3, another observation is that the performance of all the coding methods benefits from the residual convolution blocks. This result is expected since the increased effectiveness of the residual convolution block is demonstrated in previous cell

analysis works [7, 112].

To clarify why the proposed repel coding outperforms others, examples from four datasets are shown in Figs. 4.5-4.7. Fig. 4.5 shows an example from the DG dataset with the FCNN-A network. When two cells are close, proximity coding tends to merge the two centers. By comparing the prediction results with the illustrations in Fig. 4.2, we can find the reason. The proposed repel coding suppresses the responses of pixels that lie in the middle of two cell centers, and boosts the responses that are close to cell centers. The outputs from the Gaussian kernel coding and the rectangle kernel coding are as expected in Fig. 4.5, and do not have the ability to recover cell centers. Fig. 4.6 shows an example from the Vgg dataset trained by the CNN with residual convolutaion blocks. The advantage of the repel coding is obvious in the partially occluded regions. The outputs of the raw dot labels in Fig. 4.6 (a) are unstable, and tend to output duplicated cell centers. In addition, we find that training with raw dot labels can easily diverge if the training batch size is less than eight on Adip and Vgg datasets, and this does not happen with the other coding schemes. Image sizes of Adip and Vgg datasets are smaller than those of DG and HBM datasets. This may be due to the sparsity in the raw dot labels that leads to insufficient positive training examples. At last, Fig. 4.7 compares the performance of proximity coding and repel coding on Adip and HBM datasets. In these two datasets, occlusion is less common, but cell appearance varies more. Because the proposed repel coding has a larger reversibility index, the repel coding generally provides stronger responses around cell centers, resulting in more robust center detection.

Table 4.2: F1 score of the FCNN-A

	Dot label	Gaus. kernel	Rec. kernel	Proximity	**Repel**
DG	0.8526	0.8431	0.8431	0.9199	**0.92267**
Apip	0.8495	0.7918	0.7645	0.8437	**0.8784**
HBM	0.7752	0.8685	0.8110	0.7333	**0.8773**
Vgg	**0.9585**	0.9169	0.8728	0.9545	0.9583

Table 4.3: F1 score of the FCNN-A with res. blocks

	Dot label	Gaus. kernel	Rec. kernel	Proximity	**Repel**
DG	0.8887	0.8772	0.8940	0.9282	**0.9337**
Apip	0.8764	0.8177	0.7799	0.9028	**0.9038**
HBM	0.8370	0.8628	0.8121	0.8933	**0.9011**
Vgg	0.9520	0.9247	0.8992	0.9676	**0.9695**

Table 4.4: Detection results with FCNN with plane network

	Recall			Precision		
	Dot label	Proximity	Repel	Dot label	Proximity	Repel
DG	0.9167	0.9110	0.9347	0.8056	0.9373	0.9163
Apip	0.9127	0.7832	0.8926	0.8096	0.9338	0.8804
HBM	0.8345	0.6052	0.8507	0.7329	0.9479	0.9124
Vgg	0.9488	0.9184	0.9233	0.9692	0.9947	0.9970

4.4 Conclusion

In this chapter, after reviewing recent learning-based works on cell counting and detection, the common step of coding raw dot labels is extracted and discussed. Two center coding criteria are proposed: entropy and reversibility. These two criteria help predict the performance of a coding scheme at the training and prediction steps. A new coding scheme, repel coding, is proposed for a better balance with these two center coding criteria. Experimental results verify the effectiveness of repel coding for cell detection on four types of cells. In the future, we would like to explore more about the cell activation topology with the detected cell centers.

Table 4.5: Detection results with FCNN with res-block

	Recall			Precision		
	Dot label	Proximity	Repel	Dot label	Proximity	Repel
DG	0.9411	0.9249	0.9363	0.8483	0.9458	0.9356
Apip	0.9053	0.9165	0.9195	0.8650	0.9018	0.9005
HBM	0.8482	0.8793	0.9125	0.8377	0.9147	0.8963
Vgg	0.9398	0.9396	0.9438	0.9683	0.9979	0.9973

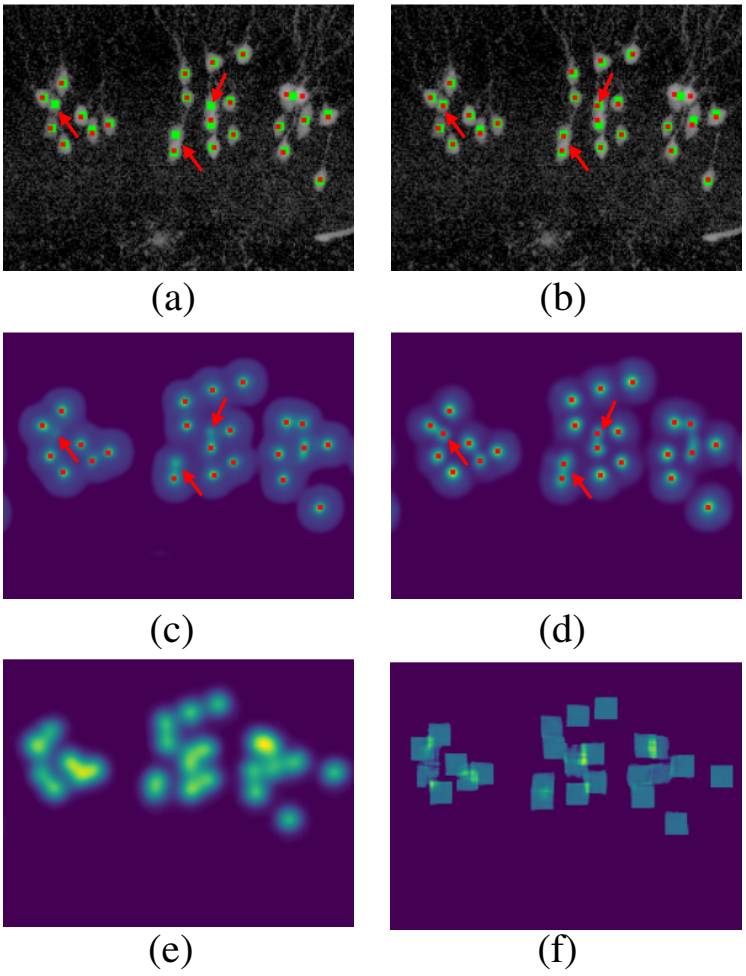


Figure 4.5: (a) Detection results with proximity coding on an example from the DG dataset. (b) Detection results with the proposed repel coding. (c) The proximity coding output. (d) The repel coding output. (e) The Gaussian kernel coding output. (f) The rectangle kernel coding output. The green dots indicate labeled cell centers, and the red dots represent detected cell centers.

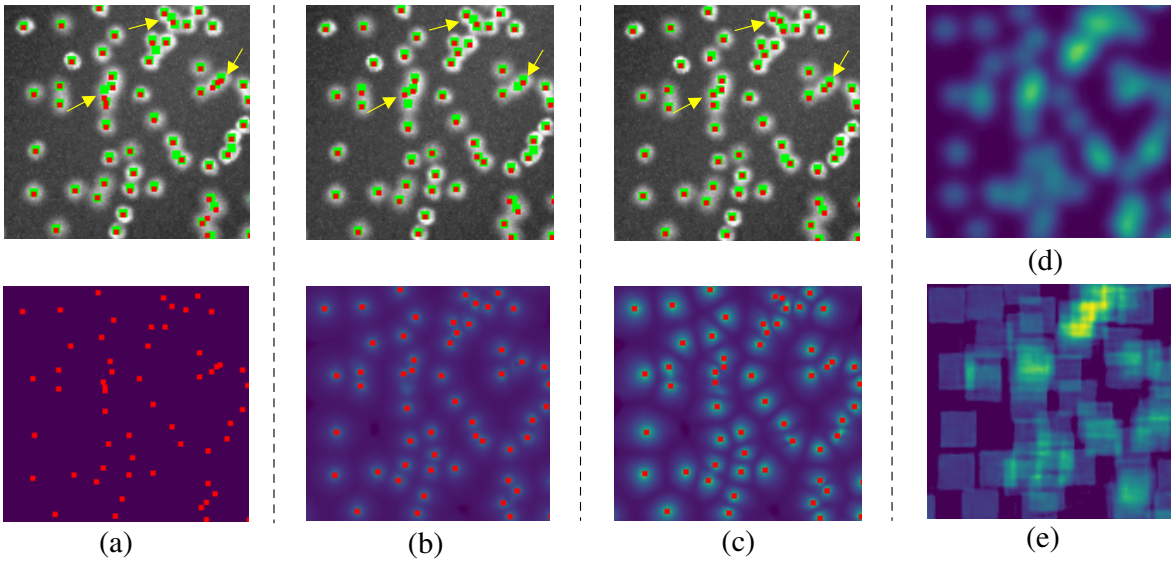


Figure 4.6: Different coding schemes on an example from the Vgg dataset. The green dots indicate labeled cell centers, and the red dots represent detected cell centers. The proposed method distinguishes adjacent cells in most cases. (a) Dot labels. (b) The proximity coding. (c) The repel coding. (d) The Gaussian kernel coding for counting. (e) The rectangle kernel for counting.

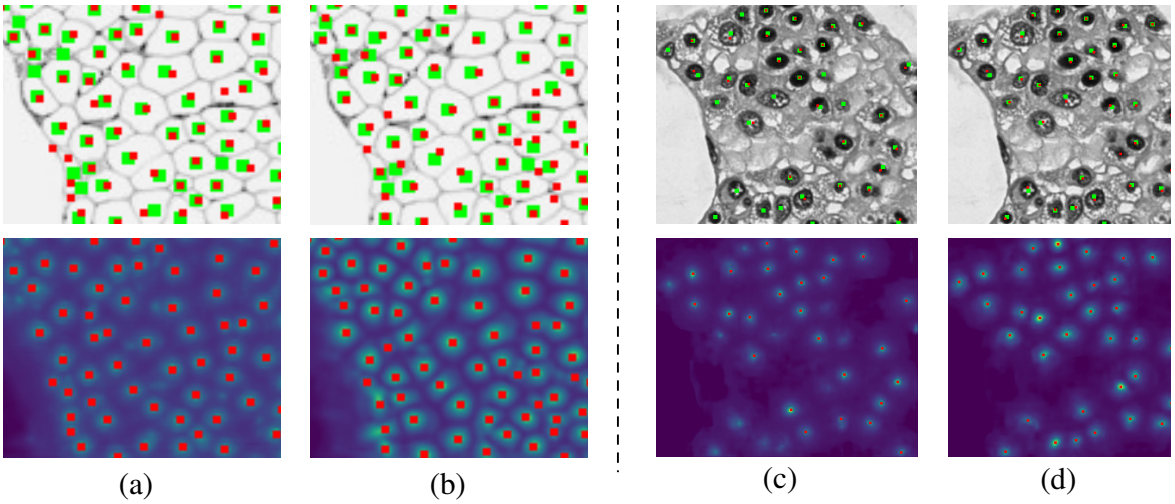


Figure 4.7: (a) and (b) are proximity and repel codings of an example from the Adip dataset. (c) and (d) are proximity and repel codings of an example from the HBM dataset. The green dots indicate labeled cell centers, and the red dots represent detected cell centers.

Chapter 5

Conclusions and Future work

5.1 Discussions and summary of the proposed works

3D mouse brain reconstruction with microscopy data In this work, two techniques are developed to reconstruct 3D mouse brains from series of brain tissues: tissue flattening and structure-based intensity propagation. During tissue preparation, brain tissues are warped by the tissue clearing process. Tissue flattening is designed to correct curved surfaces by projecting voxels onto flat surfaces. After solving the geometric distortion that is perpendicular to the cutting plane, 2D representatives are extracted from each single section for the following section registration. To obtain representatives that have strong signal-to-noise ratio and preserve prominent structure information for registration, we proposed the structure-based intensity propagation method. With this method, a median filter is applied to thick specimens along the directions of major structures in the brain tissues. Comprehensive experiments show that the proposed methods improve the flip detection accuracy and the structure consistence of brain reconstruction.

Recent developments in microscopy and tissue preparation technologies make the imaging of larger specimens with high resolution easier. For example, multiphoton microscopy [134] increases the penetration depth, and decreases the power of excitation laser beams. Therefore, imaging live specimens is much easier. With specimen expansion techniques [135], small

specimens are expanded to four times larger or even more. The whole specimens in these applications usually are imaged piece by piece, and how to align different pieces together is important. Our structure-based intensity propagation method is a potential technique for image stitching in these applications. With the proposed structure-based intensity propagation, the overlapping volumes among neighboring imaging areas increase, and thus image stitching will be more robust.

The 3D data is not unique to the microscopy. Other imaging modalities also provide more and more 3D data for biomedical research and applications. The light field cameras [136] are used for the spine reconstruction during surgeries [137]. In this work, MRI is done before surgery, and comprehensive surgery planes are made with the MRI data. During the surgery, 3D surfaces of the spine are acquired in real time with light field cameras. The proposed methods are also promising techniques for the 3D reconstruction with the light field camera.

Image quality assessment Most image restoration algorithms require one or more parameters to regulate the restoration process. For instance, the regularization parameter of image reconstruction [106] is selected by a no-reference image quality index [82]. However, most existing no-reference IQA algorithms output the estimated image quality based on a single distorted image, ignoring that different degraded images can provide more information together. The proposed comparison-based IQA method fills the gap between the increasing need of parameter selection for image processing algorithms and the lack of such an IQA algorithm that makes full use of the available information.

Besides the demonstrated applications of parameter selection for denoising and MRI reconstruction, the proposed comparison-based image quality assessment can be used in many other scenarios. As one of the important features of many cameras, auto-focus ensures the objects in pictures sharp. The proposed comparison-based image quality assessment method is able to evaluate images captured with different focal lengths and select the best imaging settings.

Cell center coding Deep convolutional neural networks (CNNs) demonstrate promising results on many object detection tasks. For the cell detection in this work, a CNN with the Unet [111] structure is used. However, as we test granule cell detection for the brain tissues in our project, existing pipelines do not provide accurate detection results when two cells are close to each other. To solve this, we focus on the ground truth output data for training. By suppressing the response values in the middle of two cell centers and deceasing the decay rate of response values away from the cell centers, the proposed repel coding improves the overall detection accuracy, especially in the densely packed regions. In biomedical research, cell analysis is a highly customized task. The morphology could change dramatically among different types of cells. Even for the same type of cell, the appearance changes with tissue preparation protocols, imaging equipment and imaging protocols. Instead of training an omnipotent cell analysis framework, the ability to adapt the analysis model to a specific cell type and imaging modality is desirable in practice. Because of the effective coding scheme, it is easy to retrain cell detection models for new types of cells. Besides the application on granule cell detection, experimental results show that the proposed repel coding scheme achieves promising results on three other cell datasets.

5.2 Future works

The future works will focus on incorporating more imaging techniques for the brain reconstruction and analysis. The current work mainly relies on the microscopy data. MRI and other non-invasive imaging techniques are promising data sources to guide registration and reconstruction of the microscopy data. With extracted cell centers, the topology formed activated cells is another topic we will investigate.

Auxiliary imaging modalities for brain registration Because of the high resolution and flexible choice of stains, microscopy data are the main imaging data in this work. However, the requirement of brain slicing before imaging introduces many challenges. More comprehensive

information can be acquired to facilitate the analysis if auxiliary imaging modalities are used in the future. For example, the brain reconstruction could be more robust if unwarped brain tissues are imaged before slicing [39]. A customized low resolution microscope is integrated with a vibrating microtome, and the 2D overviews are acquired before the slicing and the following tissue preparation steps [39]. With the low resolution microscopy data, brain reconstruction of the high resolution microscopy data has a reference image stack. Another common auxiliary data for brain analysis is MRI. The mouse brain atlas of the Allen Brain Institute [36] first reconstructs the 3D mouse brain with high resolution microscopy data, and then registers the cross-section views of the microscopy data to MRI brain templates. Since a general MRI template is used, the pipeline of mouse brain reconstruction [36] involves human visual inspection and many empirically set parameters. If all high resolution microscopy data have 3D MRI of the same brain as the reference, the reconstruction will be more accurate and robust in the future.

Topology Analysis With the reconstructed 3D mouse brains and the positions of activated cells, the next question to ask is whether the activation topology is related to memory formation in the mouse brain. To answer this, experiments on spatial working memory are conducted with the genetically engineered mice. It is known that the spatial recognition ability is largely related to the dentate gyrus in the brain. By studying the relations between the activation topology in the dentate gyrus and the environments where mice are raised, this question could be better understood in the future.

5.3 Publication list of this work

5.3.1 Journals

- **H. Liang** and D. S. Weller. "Comparison-based Image Quality Assessment for Selecting Image Restoration Parameters." IEEE Trans. on Image Processing, vol. 25, no. 11, pp. 5118-5130, Nov. 2016.

- **H. Liang**, Natalia Dabrowska, Jaideep Kapur, and Daniel S. Weller, "Structure-based Intensity Propagation for 3D Brain Reconstruction with Multilayer Section Microscopy." IEEE Trans. on Medical Imaging, October 2018, accepted.

- H. Jeelani, **H. Liang**, S. T. Acton and D. S. Weller. "Content-Aware Enhancement of Images with Filamentous Structures." IEEE Trans. on Image Processing, January 2019, accepted.

- **H. Liang**, Aijaz Naik, Cedric Williams, Jaideep Kapur, and Daniel S. Weller, "Enhanced Center Coding for Cell Detection with Convolutional Neural Networks." submitted.

5.3.2 Conferences

- **H. Liang** and D. S. Weller. "Regularization Parameter Trimming for Iterative Image Reconstruction." IEEE Asilomar Conf. on Signals, Systems, and Computers, 2015, pp. 755-759.

- **H. Liang** and D. S. Weller. "Edge-based Texture Granularity Detection." IEEE Int. Conf. on Image Process, 2016, pp. 3563-3567.

- **H. Liang** and D. S. Weller. "Denoising Method Selection by Comparison-based Image Quality Assessment." IEEE Int. Conf. on Image Process, 2016, pp. 3106-3110.

- **H. Liang**, S. T. Acton and D. S. Weller. "Content-Aware Neuron Image Enhancement." 2017 IEEE Int. Conf. on Image Processing, 2017, 3510-3514.

- **H. Liang**, N. Dabrowska, J. Kapur and D. S. Weller, "Whole Brain Reconstruction from Multilayered Sections of a Mouse Model of Status Epilepticus." 2017 IEEE Asilomar Conf. on Signals, Systems, and Computers. Pacific Grove, CA, USA, October 2017, 1260-1263.

5.3.3 Patents

- **H. Liang** and D. S. Weller. "System, Method and Computer Readable Medium for Quality Assessment for Parameter Selection." U.S. Application, Serial No. 62/254,335, filed Nov. 2015.

Bibliography

[1] U. Bagci and L. Bai. Automatic best reference slice selection for smooth volume reconstruction of a mouse brain from histological images. IEEE Trans. on Medical Imaging, 29(9):1688–1696, Sep. 2010.

[2] H. R. Sheikh, A. C. Bovik, L. Cormack, and Z. Wang. LIVE image quality assessment database release 2, 2005. http://live.ece.utexas.edu/research/quality.

[3] E. C. Larson and D. M. Chandler. Most apparent distortion: full-reference image quality assessment and the role of strategy. Journal of Electronic Imaging, 19(1):011006–1–011006–20, 2010.

[4] R. K. S. Kwan, A. C. Evans, and G. B. Pike. MRI simulation-based evaluation of image-processing and classification methods. IEEE Trans. Medical Imaging, 18(11):1085–1097, 1999.

[5] J. Wu, W. Lin, G. Shi, and A. Liu. Reduced-reference image quality assessment with visual information fidelity. IEEE Trans. Multimedia, 15(7):1700–1705, NOV. 2013.

[6] J. P. Cohen, G. Boucher, C. A. Glastonbury, H. Z. Lo, and Y. Bengio. Count-ception: Counting by fully convolutional redundant counting. In ICCV Workshop, pages 18–26, 2017.

[7] E. Walach and L. Wolf. Learning to count with cnn boosting. In ECCV, pages 660–676, 2016.

[8] W. Xie, J. A. Noble, and A. Zisserman. Microscopy cell counting and detection with fully convolutional regression networks. Computer methods in biomechanics and biomedical engineering: Imaging and Visualization, 6(3):283–292, 2018.

[9] Y. Xie, F. Xing, X. Kong, H. Su, and L. Yang. Beyond classification: structured regression for robust cell detection using convolutional neural network. In MICCAI, pages 358–365, 2015.

[10] J. L. Stringer and E. W. Lothman. Reverberatory seizure discharges in hippocampal-parahippocampal circuits. Exp Neurol, (116):198–203, 1992.

[11] E. W. Lothman et al. Functional anatomy of hippocampal seizures. Prog Neurobiol, (37):1–82, 1991.

[12] G. Medioni et al. Tensor voting: Theory and applications. In Proceedings of Reconnaissance des Formes et Intelligence Artificielle, pages 1–10, 2000.

[13] J. Dauguet et al. Three-dimensional reconstruction of stained histological slices and 3D non-linear registration with in-vivo MRI for whole baboon brain. Journal of Neuroscience Methods, 164(1):191–204, 2007.

[14] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. Image quality assessment: from error visibility to structural similarity. IEEE Trans. Image Processing, 13(4):600–612, Apr. 2004.

[15] Y. Sasaki. The truth of the F-measure. Teach Tutor mater, 1(5):1–5, 2007.

[16] S. W. Paddock (ed.) and K. W. Eliceiri. Laser scanning confocal microscopy: History, applications, and related optical sectioning techniques. Confocal Microscopy, Humana Press, New York, NY,, 1075:9–47, 2014.

[17] N. Pritchard et al. Corneal confocal microscopy predicts 4-year incident peripheral neuropathy in type 1 diabetes. Diabetes care, 38(4):671–675, 2015.

[18] T. Ju. Building a 3D atlas of the mouse brain. Rice University, Houston, TX, 2005.

[19] R. Tomer et al. Advanced CLARITY for rapid and high-resolution imaging of intact tissues. Nature Protocols, 9(7):1682–1697, 2014.

[20] E. S. Lein et al. Genome-wide atlas of gene expression in the adult mouse brain. Nature, 445(7124):168–176, 2007.

[21] S. W. Oh et al. A mesoscale connectome of the mouse brain. Nature, 508(7495):207–214, 2014.

[22] M. J. Hawrylycz et al. An anatomically comprehensive atlas of the adult human transcriptome. Nature, 489(7416):391–399, 2012.

[23] M. L. Berlanga et al. Three-dimensional reconstruction of serial mouse brain sections: solution for flattening high-resolution large-scale mosaics. Frontiers in neuroanatomy, 5(17):1–8, Mar. 2011.

[24] M. N. Economo et al. A platform for brain-wide imaging and reconstruction of individual neurons. Elife, 5(10566), 2016.

[25] J. W. Bohland et al. A proposal for a coordinated effort for the determination of brainwide neuroanatomical connectivity in model organisms at a mesoscopic scale. PLoS computational biology, 5.3(1000334):1–9, 2009.

[26] W. M. Abdelmoula et al. Automatic registration of mass spectrometry imaging data sets to the allen brain atlas. Analytical Chemistry, 86(8):3947–3954, 2014.

[27] A. Osokin et al. 3-D Mouse Brain Model Reconstruction from a Sequence of 2-D Slices in Application to Allen Brain Atlas. In International Meeting on Computational Intelligence Methods for Bioinformatics and Biostatistics. Springer, 2009.

[28] P. R. Hof. Comparative cytoarchitectonic atlas of the C57BL/6 and 129/Sv mouse brains. Elsevier, 2000.

[29] G. Paxinos and K. B. J. Franklin. The Mouse Brain in Stereotaxic Coordinates: Compact Second Edition. Gulf professional publishing, 2004.

[30] S. Ourselin et al. Reconstructing a 3D structure from serial histological sections. Image and Vision Computing, 19:25–31, 2000.

[31] N. Renier et al. iDISCO: A Simple, Rapid Method to Immunolabel Large Tissue Samples for Volume Imaging. Cell, 159(4):896–910, 2014.

[32] D. GoodSmith et al. Spatial representations of granule cells and mossy cells of the dentate gyrus. Neuron, 93(3):677–690, 2017.

[33] J. P. Neunuebel and J. J. Knierim. CA3 retrieves coherent representations from degraded input: Direct evidence for CA3 pattern completion and dentate gyrus pattern separation. Neuron, 81(2):416–427, Jan. 2014.

[34] E. T. Rolls. The mechanisms for pattern completion and pattern separation in the hippocampus. Frontiers in systems neuroscience, 7(74):1–21, Oct. 2013.

[35] M. Stille et al. 3D reconstruction of 2D fluorescence histology images and registration with in vivo MR images: Application in a rodent stroke model. Journal of Neuroscience Methods, 219(1):27–40, 2013.

[36] P. A. Yushkevich et al. 3D mouse brain reconstruction from histology using a coarse-to-fine approach. In International Workshop on Biomedical Image Registration. Springer, Berlin, Heidelberg, pages 230–237, 2006.

[37] T. Ju et al. 3D volume reconstruction of a mouse brain from histological sections using warp filtering. Journal of Neuroscience Methods, 1(156):84–100, Feb. 2006.

[38] A. Cifor, L. Bai, and A. Pitiot. Smoothness-guided 3-D reconstruction of 2-D histological images. NeuroImage, 56(1):197–211, 2011.

[39] M. Capek et. al. Volume reconstruction of large tissue specimens from serial physical sections using confocal microscopy and correction of cutting deformations by elastic registration. Microscopy research and technique, 72(2):110–119, 2009.

[40] P. Karen et al. Three-dimensional computer reconstruction of large tissue volumes based on composing series of high-resolution confocal images by gluemrc and linkmrc software. Microscopy research and technique, 65(5):415–422, 2003.

[41] S. Ghosh et al. Sensory maps in the olfactory cortex defined by long-range viral tracing of single neurons. Nature, 472(7342):217–222, 2011.

[42] K. J. Hayworth et al. Ultrastructurally smooth thick partitioning and volume stitching for large-scale connectomics. Nature Methods, 12(4):319–322, 2015.

[43] H. Liang, N. Dabrowska, J. Kapur, and D. S. Weller. Whole brain reconstruction from multilayered sections of a mouse model of status epilepticus. In 2017 IEEE Asilomar Conf. on Signals, Systems, and Computers, pages 1260–1263, 2017.

[44] C. Kenwright et al. 2D to 3D refinement of post mortem optical and mri co-registration. In MICCAI, pages 935–944, 2003.

[45] B. Yang, J. B. Treweek, R. P. Kulkarni, B. E. Deverman, C. K. Chen, E. Lubeck, S. Shah, L. Cai, and V. Gradinaru. Single-Cell Phenotyping within Transparent Intact Tissue through Whole-Body Clearing. Cell, 158(4):945–958, 2014.

[46] Y. Isogai, D. Richardson, C. Dulac, and J. Bergan. Optimized protocol for imaging cleared neural tissues using light microscopy. In Synapse Development, pages 137–153, 2017.

[47] M. R. Cronan et al. CLARITY and PACT-based imaging of adult zebrafish and mouse for whole-animal analysis of infections. Disease models and mechanisms, 8(12):1643–1650, 2015.

[48] D. S. Richardson et al. Clarifying tissue clearing. Cell, 162(2):246–257, 2015.

[49] S. Ohno et al. A morphological analysis of thalamocortical axon fibers of rat posterior thalamic nuclei: A single neuron tracing study with viral vectors. Cerebral cortex, 22(12):2840–2857, 2011.

[50] A. Roche et al. The correlation ratio as a new similarity measure for multimodal image registration. In MICCAI, pages 1115–1124, 1998.

[51] F. Maes, A. Collignon et al. Multimodality image registration by maximization of mutual information. IEEE Trans. on Medical Imaging, 16(2):187–198, 1997.

[52] A. Myronenko and X. Song. Intensity-based Image Registration by Minimizing Residual Complexity. IEEE Trans. on Medical Imaging, 29(11):1882–1891, 2010.

[53] R. G. Keys. Cubic convolution interpolation for digital image processing. IEEE Trans. on Acoustics, Speech, and Signal Processing, 29(6):1153–1160, 1981.

[54] A. F. Frangi et al. Multiscale vessel enhancement filtering. In Medical Image Computing and Computer-Assisted Intervention, pages 130–137, 1998.

[55] X. Feng and P. Milanfar. Multiscale principal components analysis for image local orientation estimation. In IEEE Asilomar Conf. on Signals, Systems, and Computers, pages 478–482, 2002.

[56] J. H. Vilkinson and C. Reinsch. Handbook for Automatic computation Vol. II: Linear Algebra. Springer, 1971.

[57] C. J. Guenthner et al. Permanent genetic access to transiently active neurons via trap: Targeted recombination in active populations. Neuron, 78(5):773–784, 2013.

[58] D. M. Mathis, J. L. Furman, and C. M. Norris. Preparation of acute hippocampal slices from rats and transgenic mice for the study of synaptic alterations during aging and amyloid pathology. Journal of visualized experiments, 49(2330):1–9, 2011.

[59] M. Capek, J. Janacek, and L. Kubinova. Methods for compensation of the light attenuation with depth of images captured by a confocal microscope. Microscopy research and technique, 69(8):624–635, 2006.

[60] G. N. Laszlo and J. Udupa. On standardizing the mr image intensity scale. Magn. Reson. Med., 42(6):1072–1081, 1999.

[61] A. Madabhushi and J. Udupa. New methods of mr image intensity standardization via generalized scale. Med. Phys., 33(9):3426–3434, 2006.

[62] M. Yigitsoy and N. Nassir. Structure propagation for image registration. IEEE Trans. on Medical Imaging, 32(9):1657–1670, Sep. 2013.

[63] E. Guest and R. Baldock. Automatic reconstruction of serialsections using the finite element method. Bioimaging, 3(4):154–167, 1995.

[64] E. P. Simoncelli and B. A. Olshausen. Natural image statistics and neural representation. Annual Review of Neuroscience, pages 1193–1216, 2001.

[65] J. Lubin. A human vision system model for objective picture quality measurements. In Broadcasting Convention, pages 498–503, 1997.

[66] S.J.P. Westen, R.L. Lagendijk, and J. Biemond. Perceptual image quality based on a multiple channel HVS model. In Proc. Int. Conf. Acoustics, Speech, Signal Processing, volume 4, pages 2351–2354, 1995.

[67] H. R. Sheikh, M. F. Sabir, and A. C. Bovik. A statistical evaluation of recent full reference image quality assessment algorithms. IEEE Trans. Image Processing, 15(11):3441–3451, Nov. 2006.

[68] Z. Wang, E. P. Simoncelli, and A. C. Bovik. Multiscale structural similarity for image quality assessment. In Proc. IEEE Asilomar Conf. on Signals, Systems, and Computers, volume 2, 2003.

[69] S. J. Daly. Visible differences predictor: an algorithm for the assessment of image fidelity. In Symposium on Electronic Imaging: Science and Technology, pages 179–206, 1992.

[70] Z. Wang and A. C. Bovik. Modern image quality assessment. 2006.

[71] H. R. Sheikh and A. C. Bovik. Image information and visual quality. IEEE Trans. Image Processing, 15(2):430–444, Feb. 2006.

[72] Z. Wang and Q. Li. Information content weighting for perceptual image quality assessment. IEEE Trans. Image Processing, 20(5):1185–1198, May. 2011.

[73] A. Shnayderman, A. Gusev, and A. M. Eskicioglu. An SVD-based grayscale image quality measure for local and global assessment. IEEE Trans. Image Processing, 15(2):422–429, Feb. 2006.

[74] Q. Li and Z. Wang. Reduced-reference image quality assessment using divisive normalization-based image representation. IEEE J. Sel. Topics Signal Processing, 3:202–211, Apr. 2009.

[75] R. Soundararajan and A. C. Bovik. RRED indices: Reduced reference entropic differencing for image quality assessment. IEEE Trans. Image Processing, 21, Feb. 2011.

[76] Z. Wang, G. Wu, H. R. Sheikh, E. P. Simoncelli, E. Yang, and A. C. Bovik. Quality-aware images. IEEE Trans. Image Process, 15(6):1680–1689, Jun. 2006.

[77] A. M. Krishna and A. C. Bovik. Blind image quality assessment: From natural scene statistics to perceptual quality. IEEE Trans. Image Processing, 20(12):3350–3364, Dec. 2011.

[78] A. Mittal, A. K. Moorthy, and A. C. Bovik. No-reference image quality assessment in the spatial domain. IEEE Trans. Image Processing, 21(12):4695–4708, Dec. 2012.

[79] S. Gabarda and G. Cristbal. Blind image quality assessment through anisotropy. J. Opt. Soc. Am. A, 24(12):B42–B51, 2007.

[80] M. A. Saad, A. C. Bovik, and C. Charrier. A DCT statistics-based blind image quality index. In IEEE Signal Process. Lett., volume 17, pages 583–586, Jun. 2010.

[81] Z. Wang, H. R. Sheikh, and A. C. Bovik. No-reference perceptual quality assessment of JPEG compressed images. In Proc. Int. Conf. Image Processing, volume 1, pages 477–480, 2002.

[82] X. Zhu and P. Milanfar. Automatic parameter selection for denoising algorithms using a no-reference measure of image content. IEEE Trans. Image Processing, 19(12):3116–3132, Dec. 2010.

[83] Q. Wu, Z. Wang, and H. Li. A highly efficient method for blind image quality assessment. In Proc. Int. Conf. Image Processing, pages 339–343, 2015.

[84] L. Li, Y. Zhou, W. Lin, J. Wu, X.Zhang, and B. Chen. No-reference quality assessment of deblocked images. Neurocomputing, 177:572–584, 2016.

[85] T. Liu, W. Lin, and C.-C. J. Kuo. A multi-metric fusion approach to visual quality assessment. In Third Int. Workshop on Quality of Multimedia Experience, pages 593–600, 2011.

[86] T. Liu, K. Liu, J. Lin, W. Lin, and C.-C. J. Kuo. A ParaBoost method to image quality assessment. IEEE Trans. Neural Networks and Learning Systems, pages 1–15, Dec. 2015.

[87] Q. Wu and et al. Blind image quality assessment based on multi-channel features fusion and label transfer. Pattern recognition, 26(3):425–440, 2015.

[88] R. L. Joshi, V.J. Crump, and T. R. Fischer. Image subband coding using arithmetic coded trellis coded quantization. IEEE Trans. Circuits Syst. Video Technol, 5:515–523, Dec. 1995.

[89] S. M. LoPresto and K. Ramchandran. Image coding based on mixture modeling of wavelet coefficients and a fast estimation-quantization framework. Data Comparession Conference, Snowbird, UT, pages 221–230, Mar. 1997.

[90] A. Liu, W. Lin, and M. Narwaria. Image quality assessment based on gradient similarity. IEEE Trans. Image Processing, 21(4):1500–1512, Apr. 2012.

[91] Y. Wang, Q. Zhang, and B. Li. Structure-preserving image quality assessment. Int. Conf. Multimedia and Expo, pages 1–6, 2015.

[92] J. Shi and C. Tomasi. Good features to track. In Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., pages 72–77, 1994.

[93] Z. Hong. Algebraic feature extraction of image for recognition. Pattern recognition, 24(3):211–219, 1991.

[94] H. R. Sheikh, A. C. Bovik, and G. Veciana. An information fidelity criterion for image quality assessment using natural scene statistics. IEEE Trans. Image Processing, 14(12):2117–2128, Dec. 2005.

[95] S. A. Golestaneh and L. J. Karam. Reduced-reference quality assessment based on the entropy of DNT coefficients of locally weighted gradients. In Proc. Int. Conf. Image Processing, 2015.

[96] X. Li. Blind image quality assessment. In Proc. Int. Conf. Image Processing, volume 1, pages 449–452, 2002.

[97] A. Buades, B. Coll, and J.M. Morel. A non local algorithm for image denoising. In Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., volume 2, pages 60–65, 2005.

[98] L. L. Thurstone. A law of comparative judgment. Psychological review, pages 273–286, 1927.

[99] B. Peng and O. Veksler. Parameter selection for graph cut based image segmentation. In Proc. Br. Mach. Vis. Conf., pages 16.1–16.10, 2008.

[100] Y. Yitzhaky and E. Peli. A method for objective edge detection evaluation and detector parameter selection. IEEE Trans. Pattern Anal. Machine Intell., 25:1027–1033, Aug. 2003.

[101] M. S. C. Almeida and M. T. Figueiredo. Parameter estimation for blind and non-blind deblurring using residual whiteness measures. IEEE Trans. Image Processing, 22(7):2751–2763, Jul. 2013.

[102] J. L. Herraiz, M. Ieee, S. Gabarda, G. Cristobal, and S. Member. Automatic parameter selection in PET image reconstruction based on no-reference image quality assessment. In Nuclear Science Symposiwn and Medical Imaging Conference Record, pages 3371–3374, 2012.

[103] D. Weller, S. Ramani, and J. A. Fessler. Augmented Lagrangian with variable splitting for faster nonCartesian L1-SPIRiT MR image reconstruction. IEEE Trans. Med. Imag., 33(2):351–361, Feb. 2014.

[104] G. Demoment. Image reconstruction and restoration: overview of common estimation structures and problems. IEEE Trans. Acoust. Sp. Sig. Proc., 37(12):2024–2036, Dec. 1989.

[105] T. Wiegand and B. Girod. Lagrange multiplier selection in hybrid video coder control. In Proc. Int. Conf. Image Processing, volume 3, pages 542–545, 2001.

[106] H. Liang and D. S. Weller. Regularization parameter trimming for iterative image reconstruction. In Proc. IEEE Asilomar Conf. on Signals, Systems, and Computers. IEEE, 755-759.

[107] L. Rudin and S. Osher. Total variation based image restoration with free local constraints. In Proc. Int. Conf. Image Processing, volume 1, pages 31–35, 1994.

[108] T. Goldstein and S. Osher. The split Bregman method for L1-Regularized problems. SIAM Imaging Science, 2:323–343, Jan. 2009.

[109] C. Tomasi and R. Manduchi. Bilateral filtering for gray and color images. In Proc. Int. Conf. Image Processing, pages 839–846, 1998.

[110] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian. Image denoising by sparse 3-D Transform-Domain collaborative filtering. IEEE Trans. Image Process., 16(8):2080–2096, Aug. 2007.

[111] O. Ronneberger, P. Fischer, and T. Brox. U-net: Convolutional networks for biomedical image segmentation. In MICCAI, pages 234–241, 2015.

[112] Y. Xue, R. Nilanjan, J. Hugh, and G. Bigras. Cell counting by regression using convolutional neural network. In ECCV, pages 274–290, 2016.

[113] A. Krizhevsky, I. Sutskever, and G. E. Hinton. Imagenet classification with deep convolutional neural networks. In NIPS, pages 1097–1105, 2012.

[114] A. Paszke, S. Gross, S. Chintala, G. Chanan, E. Yang, Z. Devito, Z. Lin, A. Desmaison, L. Antiga, and A. Lerer. Automatic differentiation in pytorch. In NIPS Workshop, 2017.

[115] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov. Dropout: a simple way to prevent neural networks from overfitting. The Journal of Machine Learning Research, 15(1):1929–1958, 2014.

[116] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In CVPR, pages 770–778, 2016.

[117] I. J. Goodfellow, D. Warde-Farley, M. Mirza, A. Courville, and Y. Bengio. Maxout networks. 2013. arXiv 1302.4389.

[118] R. Girshick, J. Donahue, T. Darrell, and J. Malik. Rich feature hierarchies for accurate object detection and semantic segmentation. In CVPR, pages 580–587, 2014.

[119] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi. You only look once: Unified, real-time object detection. In CVPR, pages 779–788, 2016.

[120] Y. Wu, J. Lim, and M. H. Yang. Online object tracking: A benchmark. In CVPR, pages 2411–2418, 2013.

[121] T. Y. Lin, M. Maire, S. Belongie, L. Bourdev, R. Girshick, J. Hays, P. Perona, D. Ramanan, C. L. Zitnick, and P. Dollar. Microsoft COCO: Common objects in context. In ECCV, pages 740–755, 2014.

[122] J. Long, E. Shelhamer, and T. Darrell. Fully convolutional networks for semantic segmentation. In CVPR, pages 3431–3440, 2015.

[123] K. He, G. Gkioxari, P. Dollar, and R. Girshick. Mask R-CNN. In ICCV, pages 2980–2988, 2017.

[124] Y. Xue and N. Ray. Cell detection with deep convolutional neural network and compressed sensing. 2017. arXiv 1708.03307.

[125] C. Shang, H. Ai, and B. Bai. End-to-end crowd counting via joint learning local and global count. In ICIP, pages 1215–1219, 2016.

[126] H. Idrees, I. Saleemi, C. Seibert, and M. Shah. Multi-source multi-scale counting in extremely dense crowd images. In CVPR, pages 2547–2554, 2013.

[127] J. Redmon and A. Farhadi. YOLO9000: Better, Faster, Stronger. In CVPR, pages 7263–7271, 2017.

[128] M. Bai and R. Urtasun. Deep watershed transform for instance segmentation. In CVPR, pages 2858–2866, 2017.

[129] H. Liang, N. Dabrowska, J. Kapur, and D. S. Weller. Structure-based Intensity Propagation for 3D Brain Reconstruction with Multilayer Section Microscopy. IEEE Transactions on Medical Imaging, 2018, in press.

[130] P. Kainz. You Should Use Regression to Detect Cells. In MICCAI, pages 276–283, 2015.

[131] V. Lempitsky and A. Zisserman. Vgg cell dataset from learning to count objects in images. 2010.

[132] A. Lehmussola, P. Ruusuvuori, J. Selinummi, H. Huttunen, and O. Yli-Harja. Computational framework for simulating fluorescence microscope images with cell populations. IEEE Transactions on Medical Imaging, 26(7):1010–1016, 2007.

[133] D. P. Kingma and J. L. Ba. Adam: A method for stochastic optimization. 2014. arXiv 1412.6980.

[134] W. R. Zipfel et. al. Nonlinear magic: multiphoton microscopy in the biosciences. Nature biotechnology, 21(11):1369–1377, 2003.

[135] R. Gao et. al. Cortical column and whole-brain imaging with molecular contrast and nanoscale resolution. Science, 363(245), 2019.

[136] T. Wang et. al. Light field video capture using a learning-based hybrid imaging system. ACM Transactions on Graphics, 36(4), 2017.

[137] Proprio vision. https://www.propriovision.com/technology.

[138] A. K. Moorthy and A. C. Bovik. A two-step framework for constructing blind image quality indices. IEEE Signal Process. Lett., 17(5):513–516, May. 2010.

[139] Z. Wang, A. C. Bovik, and B. L. Evan. Blind measurement of blocking artifacts in images. In Proc. Int. Conf. Image Processing, volume 3, pages 981–984, 2000.

[140] J. Caviedes and F. Oberti. A new sharpness metric based on local kurtosis, edge and energy information. Signal Processing: Image Communication, 19:147–161, 2004.

[141] D. L. Ruderman. The statistics of natural images. In Network: Computation in Neural Systems, volume 5, pages 517–548, 1994.

[142] P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. IEEE Trans. Pattern Anal. Machine Intell., 12(7):629–639, Jul. 1990.

[143] M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester. Image inpainting. In Proc. Conf. Computer Graphics and Interactive Techniques, pages 417–424, 2000.

[144] P. C. Teo and D. J. Heeger. Perceptual image distortion. In Proc. SPIE, pages 127–141, 1994.

[145] F. XJ. Lukas and Z. L. Budrikis. Picture quality prediction based on a visual model. IEEE Trans. Communications, 30(7):1679–1692, Jul. 1982.

[146] A. P. Bradley. A wavelet visible difference predictor. IEEE Trans. Image Processing, 8(5):717–730, May 1999.

[147] A. Buades, B. Coll, and J. M. Morel. A review of image denoising algorithms with a new one. Multiscale Modeling and Simulation, 4:490–530, 2005.

[148] T. H. Cormen, C. E. Leiserson R. L. Rivest, and C. Stein. Introduction to Algorithm. The MIT Press, 2001.

[149] J. Liang, L. Zhang, and J. Yang. Mixed noise removal by weighted encoding with sparse nonlocal regularization. IEEE Trans. Image Process., 23(6):2651–2662, Apr. 2014.

[150] H. R. Sheikh, A. C. Bovik, L. Cormack, and Z. Wang. LIVE image quality assessment database release 2, 2005. http://live.ece.utexas.edu/research/quality.

[151] M. Elad and M. Aharon. Image denoising via sparse and redundant representation over learned dictionaries. IEEE Trans. Image Process., 15(12):3736–3745, 2006.

[152] J. Mairal, F. Bach, J. Ponce, G. Sapiro, and A. Zisserman. Non-local sparse models for image restoration. In Proc. ICCV, pages 2272–2279, 2009.

[153] B. Goossens, H. Luong, A. Pizurica, and W. Philips. An improved non-local denoising algorithm. In Proc. Int. Workshop on Local and Non-Local Approximation in Image Processing (LNLA), pages 143–157, 2008.

[154] W. Zuo, L. Zhang, C. Song, and D. Zhang. Texture enhanced image denoising via gradient histogram preservation. In Proc. CVPR, pages 1203–1210, 2013.

[155] H. Liang and D. S. Weller. Comparison-based image quality assessment for selecting image restoration parameters. IEEE Trans. on Image Processing, 25(11):5118–5130.

[156] F. Chen, L. Zhang, and H. Yu. External patch prior guided internal clustering for image denoising. In Proc. ICCV, pages 603–611, 2015.

[157] T. Dai, C. Song, J. Zhang, and S. Xia. PMPA: A patch-based multiscale products algorithm for image denoising. In Proc. ICIP, pages 4406–4410, 2015.

[158] Q. Wu, Z. Wang, and H. Li. A highly efficient method for blind image quality assessment. In Proc. ICIP, pages 339–343, 2015.

[159] J. Xu, L. Zhang, W. Zuo, D. Zhang, and X. Feng. Patch group based nonlocal self-similarity prior learning for image denoising. In Proc. ICCV, pages 244 – 252, 2015.

[160] D. Krishnan and R. Fergus. Fast image deconvolution using hyper-laplacian priors. In Proc. NIPS, pages 1033–1041, 2009.

[161] L. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithm. Physical D: Nonlinear Phenomena, pages 259–268, 1992.

[162] E. C. Larson and D. M. Chandler. Most apparent distortion: full-reference image quality assessment and the role of strategy. Journal of Electronic Imaging, 19(1):011006–1–011006–20, Mar. 2010.

[163] Q. Li and Z. Wang. Reduced-reference image quality assessment using divisive normalization-based image representation. IEEE J. Sel. Topics Signal Processing, 3:202–211, Apr. 2009.

[164] A. M. Krishna and A. C. Bovik. Blind image quality assessment: From natural scene statistics to perceptual quality. IEEE Trans. Image Process., 20(12):3350–3364, Dec. 2011.

[165] R. Soundararajan and A. C. Bovik. RRED indices: Reduced reference entropic differencing for image quality assessment. IEEE Trans. Image Process., 21(2):517–526, Feb. 2011.

[166] D. G. Lowe. Object recognition from local scale-invariant features. In ICCV, 1999.

[167] Y. Xie, F. Xing, X. Kong, H. Su, and L. Yang. Beyond Classification: Structured Regression for Robust Cell Detection Using Convolutional Neural Network. In MICCAI, pages 358–365, 2015.

[168] J. Matasa, O. Chuma, M. Urbana, and T. Pajdla. Robust wide-baseline stereo from maximally stable extremal regions. Image Vision Comput., 22(10):761–767, 2004.

[169] J. Lonsdale et. al. The genotype-tissue expression (GTEx) project. Nature genetics, 45(6):580–585, 2013.

[170] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In CVPR, pages 886–893, 2005.

[171] Y. Ke and R. Sukthankar. ORB: An efficient alternative to SIFT or SURF. In ICCV, pages 2564–2571, 2011.

[172] Y. Ke and R. Sukthankar. PCA-SIFT: A more distinctive representation for local image descriptors. In CVPR, page 2, 2004.

[173] T. F. Chan and L. A. Vese. Active contours without edges. IEEE Transactions on image processing, 10(2):266–277, 2001.

[174] C. Fu, S. Lee, D. J. Ho, S. Han, P. Salama, K. W. Dunn, and E. J. Delp. Fluorescence Microscopy Image Segmentation Using Convolutional Neural Network With Generative Adversarial Networks. 2018. arXiv 1801.07198.

[175] M. Sussman, P. Smereka, and S. Osher. A level set approach for computing solutions to incompressible two-phase flow. Journal of Computational physics, 114(1):146–159, 1994.

[176] M. Ignatiadis, M. Buyse, and C. Sotiriou. St Gallen international expert consensus on the primary therapy of early breast cancer: an invaluable tool for physicians and scientists. Annals of Oncology, 26(8):1519–1520, 2015.

[177] R. L. Camp and G. G. Chung D. L. Rimm. Automated subcellular localization and quantification of protein expression in tissue microarrays. Nature medicine, 8(11):1323–1328, 2002.

[178] E. Bernardis and S. X. Yu. Pop out many small structures from a very large microscopic image. Medical image analysis, 15(5):690–707, 2011.

[179] C. Arteta, V. Lempitsky, J. A. Noble, and A. Zisserman. Learning to detect cells using non-overlapping extremal regions. In MICCAI, pages 348–356, 2012.

[180] E. Meijering. Cell segmentation: 50 years down the road. IEEE Signal Processing Magazine, 29(5):140–145, 2012.

[181] T. Tikkanen, P. Ruusuvuori, L. Latonen, and H. Huttunen. Training based cell detection from bright-field microscope images. In Image and Signal Processing and Analysis, pages 160–164, 2015.

[182] S. Lee and P. Bajcsy. Trajectory fusion for three-dimensional volume reconstruction. Computer Vision and Image Understanding, (110):19–31, 2007.

[183] J. R. McIntosh J. R. Kremer, D. N. Mastronarde. Computer Visualization of Three-Dimensional Image Data Using IMOD. Journal of structural biology, 1(116):71–76, 1996.

[184] http://ncmir.ucsd.edu/downloads/mogrifier/mogrifier.shtm.

[185] J. R. Anderson et al. A computational framework for ultrastructural mapping of neural circuitry. In PLoS biology, volume 7(3), 2009.

[186] J. Canny. A computational approach to edge detection. IEEE Trans. Pattern Analysis and Machine Intelligence, 8(6):679–698, Nov. 1986.

[187] S. Mukherjee, B. Condron, and S. T. Acton. Tubularity flow field– a technique for automatic neuron segmentation. IEEE Trans. on Image Processing, 24(1):374–390, Jan. 2015.

[188] M. Jirkovska et. al. Three-dimensional reconstructions from non-deparaffinized tissue sections. Anatomy and embryology, 210(3):163–173, 2005.

[189] A. Diaspro. Confocal and two-photon microscopy foundation, applications, and advances. New York: Wiley., 2002.

[190] P. Dollar and C. L. Zitnick. Structure forests for fast edge detection. In ICCV, pages 1841–1848, 2013.

[191] D. Rueckert et al. Non-rigid registration using free-form deformations: application to breast MR images. IEEE Trans. on Medical Imaging, 18(8):712–721, 1999.

[192] A. Diaspro. Confocal and two-photon microscopy, foundation, applications, and advances. New York: Wiley, 2002.

[193] C. Sorzano and P. Thevenaz and M. Unser. Elastic Registration of Biological Images Using Vector-Spline Regularization. IEEE Trans. on Biomedical Engineering, 52(4):652–663, 2005.

[194] T. Rohlfing et al. Evaluation of atlas selection strategies for atlas-based image segmentation with application to confocal microscopy images of bee brains. NeuroImage, 21(4):1428–1442, 2004.

[195] R. B. Timothy et al. From marrow to brain: expression of neuronal phenotypes in adult mice. Science, 290(5497):1775–1779, 2000.

[196] D. Rueckert et al. Nonrigid registration using free-form deformations: Application to breast MR images. IEEE Trans. on Image Processing, 18(8), Aug. 1999.

[197] F. Maes et al. Multimodality image registration by maximization of mutual information. IEEE Trans. on Medical Imaging, 16(2):187–198, Apr. 1997.

[198] A. Myronenko and X. Song. Intensity-based image registration by minimizing residual complexity. IEEE Trans. on Medical Imaging, 29(11):1882–1891, Nov. 2010.

[199] Gabor Szekely M. Styner, C. Brechbuhler and Guido Gerig. Parametric estimate of intensity inhomogeneities applied to mri. IEEE Trans. on Medical Imaging, 19(3):153–165, Mar. 2000.

[200] S. Mukherjee and S. T. Acton. Oriented filters for vessel contrast enhancement with local directional evidence. In International Symposium on Biomedical Imaging, pages 503–506, 2015.

[201] R. Manniesing et al. Vessel enhancing diffusion: A scale space representation of vessel structures. Medical Image Analysis, 10(6):815–825, 2006.

[202] K. Krissian et al. Model-based detection of tubular structures in 3D images. Computer Vision and Image Understanding, 80(2):130–171, 2000.

[203] L. B. Cohen, B. M. Salzberg, and A. Grinvald. Optical methods for monitoring neuron activity. Annual review of neuroscience, 1(1):171–182, 1978.

[204] M. Sofka and C. V. Stewart. Retinal vessel centerline extraction using multiscale matched filters, confidence and edge measures. IEEE Trans. on Med. Imag., 25(12):1531–1546, 2006.

[205] T. Kohler et. al. Automatic No-reference quality assessment for retinal fundus images using vessel segmentation. In International Symposium on Computer-Based Medical Systems, pages 95–100. IEEE, 2013.

[206] H. Peng et al. BigNeuron: large-scale 3D neuron reconstruction from optical microscopy images. Neuron, 87(2):252–256, Jul. 2015.

[207] K. Rajkovic et al. Mathematical modeling of the neuron morphology using two dimensional images. Journal of Theoretical Biology, 390:80–85, 2016.

[208] H. Peng, F. Long, and G. Myers. Automatic 3D neuron tracing using all-path pruning. Bioinformatics, 27:i239–i247, 2011.

[209] H. Xiao and H. Peng. APP2: automatic tracing of 3D neuron morphology based on hierarchical pruning of a gray-weighted image distance-tree. Bioimage informatics, 29(11):1448–1454, 2013.

[210] T. Lee, R. L. Kashyap, and C. Chu. Building skeleton models via 3-D medial surface/axis thinning algorithms. Computer Vision, Graphics, and Image Processing, 56(6):462–478, 1994.

[211] N. Otsu. A threshold selection method from gray-level histograms. IEEE Trans. on Systems, Man, and Cybernetics, 9(1):62–66, 1979.

[212] H. Peng, Z. Ruan, F. Long, J. H. Simpson, and E. W. Myers. V3D enables real-time 3D visualization and quantitative analysis of large-scale biological image data sets. Nature Biotechnol, 28(4):348–353, 2010.

[213] P. Chatterjee and P. Milanfar. Is denoising dead? IEEE Trans. on Image Processing, 19(4):895–911, 2010.

[214] D. Krishnan and R. Fergus. Fast image deconvolution using hyper-laplacian priors. In NIPS, pages 1033–1041, 2009.

[215] S. T. Acton H. Liang and D. S. Weller. Content-aware neuron image enhancement. In ICIP, pages 3510–3514, 2017.

[216] L. Condat. A direct algorithm for 1-D total variation denoising. IEEE Signal Processing Letter, 20(11):1054–1057, 2013.

[217] L. Xu, C. Lu, Y. X, and J. Jia. Image smoothing via L_0 gradient minimization. ACM Transactions on Graphics, 30(6):174–186, Dec. 2011.

[218] M. Elad M. Aharon and A. Bruckstein. K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation. Trans. on Signal Process., 54(11):4311–4323, Nov. 2006.

[219] T. Lee, R. L. Kashyap, and C. Chu. Building skeleton models via 3-D medial surface/axis thinning algorithms. Computer Vision, Graphics, and Image Processing, 56(6):462–478, 1994.

[220] R. Timofte R. Rothe and L. V. Gool. Efficient regression priors for reducting image compression artifacts. In ICIP, pages 1543–1547, 2015.

[221] L. Xu J. Xie and E. Chen. Image denoising and inpainting with deep neural networks. In NIPS, pages 341–349, 2012.

[222] D. Pathak et al. Context encoders: Feature learning by inpainting. In CVPR, pages 2536–2544, 2016.

[223] C. Bao, J. Cai, and H. Ji. Fast sparsity-based orthogonal dictionary learning for image restoration. In ICCV, pages 3384–3391, 2013.

[224] Y. LeCun, Y. Bengio, and Geoffrey Hinton. Deep learning. In Nature, pages 436–444, 2015.

[225] M. Abadi et al. Tensorflow: Large-scale machine learning on heterogeneous distributed systems. In arXiv, 2016.

[226] Y. Jia et al. Caffe: Convolutional architecture for fast feature embedding. In ACMMM, pages 675–678, 2014.

[227] L. Vincent and P. Soille. Watershed in Digital Spaces: An Efficient Algorithm Based on Immersion Simulations. IEEE Trans. on Pattern Analysis and Machine Intelligence, 13(6):583–598, Jun. 1991.

[228] H. T. Nguyen, M. Worring, and R. Boomgaard. Watersnakes: Energy-Driven Watershed Segmentation. IEEE Trans. on Pattern Analysis and Machine Intelligence, 25(3):330–343, MAR. 2003.

[229] G. Hamarneh and X. Li. Watershed segmentation using prior shape and appearance knowledge. Image and Vision Computing, 27(1):59–68, 2009.

[230] S. Basu, A. Aksel, B. Condron, and S. T. Acton. TREE2TREE: Neuron Segmentation for Generation of Neuronal Morphology. ISBI, year =.

[231] C. A. Schneider, W. S. Rasband, and K. W. Eliceiri. NIH Image to ImageJ: 25 years of image analysis. Nature Methods, 9(7):671–675, 2012.

[232] W. M. Abdelmoula et al. Automatic Registration of Mass Spectrometry Imaging Data Sets to the Allen Brain Atlas. Analytical chemistry, 86(8):3947–3954, 2014.

[233] J. Kapur, N. Dabrowska, S. Joshi, and J. Williamson. Mapping seizure spread during status epilepticus at a single neuron resolution.

[234] S. W. Oh et al. A mesoscale connectome of the mouse brain. Nature, 508(7495):207, 2014.

[235] T. Chan and W. Zhu. Level set based shape prior segmentation. In CVPR, pages 1164–1170, 2005.

[236] M. Gastaud, M. Barlaud, and G. Aubert. Combining Shape Prior and Statistical Features for Active Contour Segmentation. IEEE Trans. CSVT, 14(5):726–734, 2004.

[237] E. Trinka et al. A definition and classification of status epilepticus–Report of the ILAE Task Force on Classification of Status Epilepticus. Epilepsia, (56):1515–1523, 2015.