平成30年度　　博士論文

Super-Resolution and Reconstruction of Medical Images

医用画像の超解像と再構成

名古屋工業大学　大学院　工学研究科
情報工学専攻

河村 直輝

Contents

Chapter 1

Introduction

1.1 Medical Imaging

Medical images are used to describe the anatomical structures, bodily functions, and lesions. Medical images are necessary for doctors to plan medical treatments in hospitals, and are also referred to in wide clinical research areas in order to observe and unravel unknown bodily functions and lesions. Nowadays, there is a variety of medical image modalities, and the respective imaging processes, matters, and objects are different by the modality.

For example, the X-ray photography, which is the projection imaging of the three dimensional (3D) bodily structures onto a 2D image, is widely used for medical-diagnosis. Since X-ray was detected in 1895 to today, it has been the most fundamental medical image modality to describe inside the living bodies. Histopathological, or pathological images are commonly used for the study of a pathology or a diagnosis. Given pathological images, one can observe the 2D cross-sections of the 3D anatomical structures with high spatial resolutions. They are often used for obtaining the gold-standard of the description of the very micro anatomical structures. They appeared in 17th century, and have been leveraged to understand the pathological bodily functions and lesions. In 1970s, the Computational Tomography (CT) imaging appeared, which is a 3D volume imaging using X-ray. The X-ray CT imaging is the fundamental modality for observing 3D anatomical structures. Positron Emission Tomography (PET) and Magnetic Resonance (MR) imaging appeared in 1980s. PET visualizes the physical function for each body part with a visualized 3D tomographic image. In addition, they can visualize the temporal change of medical reaction, where a PET image is represented by a four-dimensional image (3D-space + 1D-temporal change). A temporal series of PET images is widely used to analyze the dynamics of ligands in the human brain [1]. MR imaging is also one of important methods for 3D volume imaging with high contrast. It measures the density and reactions of hydrogen nuclei in a body when the body is exposed to several magnetic fields.

In spite of the broad usability of medical images, however, the improvements of image qualities are highly required, and are still under the developments for all the medical image modalities. For doctors to plan more and better medical treatments, or for researchers and computers to recognize or understand the images deeply, high quality imaging or measurement is important [2, 3, 4].

There are two principal kinds of limitations accompanied by the measurement of medical images. First, there are physical limitations of imaging equipment. For instance, lens apertures of optical cameras restrict the spatial resolutions of the observed images. The strength of the magnetic fields also restricts the spatial resolutions and signal-to-noise ratio (SNR) of MR images. Second, there are issues of health risks to our bodies. For example, some contrast agent is often injected into a patient's body in CT and MR imaging. In PET imaging, a patient's blood is sometimes sampled to analyze the medical reaction with PET images. Those approaches are widely employed to improve the image qualities instead of health risks. It is still always desired to obtain high quality images without those health risks, and there are many studies to avoid health risks during the measurements of medical images [2, 3, 4]. Apart from those physical and health limitations, there are also several trade-offs accompanied by the high quality imaging, such as imaging cost, and time spent for imaging.

In order to achieve the high quality imaging over those limitations, image restoration tasks such as denoising, tomographic reconstruction, and super-resolution are often desired and employed. Denoising is a typical task since many medical image modalities suffer from the terrible amount of noise [5, 6, 7]. Super-resolution is also pursued in several modalities [8, 9, 10] because the achievement of detail image descriptions over the measurements is one of the most important factors. In MR imaging, compressed sensing [11, 12, 13] is widely used and studied to save its imaging time/cost while keeping the image quality. In this paper, we focus on the several kinds of medical image reconstruction and super-resolution techniques. In general, those image retrieval tasks are described by solving inverse problems. We introduce a generalized representation of image inverse problems and how to solve them, in the next section.

1.2 Inverse Problems of Image Processing

There is a variety of image modalities by the purpose of the measurement such as digital color images, telescopic images, spectral images, and medical images described in Section 1.1. In many cases, however, a raw observed image is not enough to understand the details of the observed image or to achieve rich knowledge on it. Almost all modalities suffer from their degradation of the observed image such as noise, blurs, and artifacts caused during the imaging process. Given an observed image, the image retrieval techniques such

as denoising, deblurring, completion, and super-resolution try to obtain an ideal image which is free of any degradation. The ideal image, here, is defined depending on the imaging system and the modality while commonly satisfying the following conditions at least: (1) there is no structural location/shape gaps between the image and the substance to be imaged, and (2) the SNR diverges to ∞. If the ideal image is obtained, it is made good progress not only for humans to recognize and understand the image but also for computers to do.

Most of image restoration, image super-resolution, and image reconstruction tasks can be represented as solving inverse problems. In order to accomplish those tasks, first, an observation model is defined and formulated based on the respective observation process. Equation (1.1) shows the general form of an observation model [14]:

$$\boldsymbol{y} = f(\boldsymbol{x}) + \boldsymbol{\epsilon}, \tag{1.1}$$

where $\boldsymbol{y}$ is an observed image and $\boldsymbol{x}$ is the ideal image without any degradation. f is an observation function that degrades the ideal image. $\boldsymbol{\epsilon}$ is the noise or artifact model accompanied by the observation. For example, in case of super-resolution, f consists of blurring and downsampling operators which can be expressed as the matrices multiplication. In case of the tomographic reconstruction, f is the Radon transform and ϵ represents the noise accompanied with the Poisson distributed observation. Given the observed image $\boldsymbol{y}$, the goal is to solve Equation (1.1) with respect to $\boldsymbol{x}$. In many cases, however, solving Equation (1.1) is an ill-posed problem and the ideal, unique solution $\boldsymbol{x}$ cannot be obtained from just solving Equation (1.1). The ill-posedness of an image-processing task is mainly derived form the lack of constraint equations to constrain the solution space. The ill-posed nature varies depending on the observation model. For example, in the image completion or inpainting tasks [15, 16], the lack of image pixels makes them ill-posed. In the image super-resolution [8] and in the compressed sensing of MR images [11], the ill-posedness is caused by the lack of high-frequency components in the frequency domain rather than the downsample of image pixels. In the tomographic reconstruction tasks [17, 18], the ill-posedness is because of the terrible observation noise derived from the Poisson distribution.

In order to solve an ill-posed inverse problem, we have to constrain the solution space of $\boldsymbol{x}$ based on some additional priors of the image. These priors would vary depending on the image-processing tasks and on the image modalities, and we should employ appropriate priors based on them. In order to introduce the priors to constrain the solution space, the priors are modeled in the form of mathematical operations. There are typical three kinds of the prior models to solve the inverse problems. The typical examples of the prior models are introduced in the next paragraphs.

Projection onto Convex Sets (POCS)

The brief concept of the POCS is to constrain the solution space by the multiple convex sets where the ideal image reside. With the POCS, the input observed image is iteratively projected onto given convex subspaces until the convergence. The generalized POCS approaches can be formulated by the iterative model [19] as

$$\boldsymbol{x}^{k+1} = P_J \circ \cdots \circ P_{j+1} \circ P_j \circ \boldsymbol{x}^k, \tag{1.2}$$

where P_j is the j-th of J convex sets and $\circ$ is the operator to apply a projection. Starting from the input $\boldsymbol{x}^0$, POCS iteratively imposes the constraints of J convex sets to retrieve the ideal image. These convex subspaces should be defined based on the purpose of a task by users. POCS is originally established for super-resolution or phase retrieval methods [20, 21, 19, 22], and typical convex subspaces for the image retrieval are defined or established by [19, 22]. Some of them [19, 22, 23] have strong constraint performance because of their theoretical backgrounds. The details of the conventional POCS methods including one of a typical model are mentioned in Chapter 2.

Regularization

The regularization would be the most typical approach to solve an ill-posed problem. In general, the regularization techniques constrain the solution space by additional mathematical models based on the priors that the ideal image partially satisfy. Unlike the POCS models, the ideal image does not strictly satisfy the regularization model. The regularization supplements the lack of the observation constraint and prevents the image from over-fitting to the observation model at the same time.

Above all, the spatial smoothness is the widely used prior for the regularization. The smoothness of an image is usually evaluated by the L_q-norm of the image gradient-features. One of the typical regularization model is the Tikhonov regularization [24, 14], which is formulated as

$$\underset{\boldsymbol{x}}{\text{minimize}} \quad \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{y}\|_2^2 + \lambda \|\boldsymbol{L}\boldsymbol{x}\|_q^2, \tag{1.3}$$

where $\boldsymbol{y}$, $\boldsymbol{A}$, and $\boldsymbol{x}$ are respectively the observed image, the linear observation matrix, and the image to be restored. $\boldsymbol{L}$ is the feature mapping matrix, and q is the regularization metric. The first term evaluates the fidelity derived from the observance model, and the second term is the regularization in the feature space $\boldsymbol{L}\boldsymbol{x}$. By arguing Problem (1.3), the solution space is restricted between the data fidelity term and the regularization term. Two terms are balanced by the parameter λ.

The most fundamental regularizer is L_2 norm. Although it is easy to solve using the partial derivative, the L_2-based regularization isotropically degrades the feature weights: The higher the feature weights are, the more rapidly they are reduced. When the L_2-norm

is applied to the image gradients, all of the edges are simultaneously smoothed. Thus the L_2-norm is not suitable for several retrieval tasks such as super-resolution where it is desired to preserve edges and detail patterns.

In order to preserve edges, the L_0 regularization must be argued. However, it is in fact quite difficult to directly solve the L_0-norm because it is a combination optimization problem which is nonconvex, discontinuous in the all domain, and NP-hard. The L_0-norm is often approximated with the L_1-norm where the function is convex and continuous but have a small number of undifferentiable points. In recent years, it came to be able to efficiently be solved by using the Majorization Minimization (MM) algorithm [25] or the proximal operator [26]. When the L_1-norm is applied for the edge features, it is called the Total Variation (TV) norm. The details of the TV-norm are also presented in Chapter 2. Moreover in recent years, there are several methods proposed to directly solve the L_0-norm by somehow constraining the solution space [27, 28, 29]. It is also reported that L_0-norm regularization outperforms the other smoothness regularizations in several tasks while the nonconvexity is still remained [27, 28, 29]. Some of the applications of L_0-norm are presented in Chapter 4.

For a regularization model to represent the smoothness prior, the rank can also be leveraged. The low-rankness evaluates a global connectivity of the image space and can effectively reduce the observation noise. The rank function can be well approximated a convex function with the nuclear norm while it is originally nonconvex and NP-hard [30, 31].

Learning based models

Learning based approaches solve inverse problems by exploiting internal or external database [32, 33, 10]. In recent years, deep learning has come to be able to be implemented owing to both of hard and software improvements such as online bigdata, big-scale parallel processing units including GPU. Some complicated problems also became able to be solved with theoretical improvements such as batch-normalization [34] and ADAM [35]. Nowadays more and more deep learning based methods are applied to a variety of inverse problems [36, 37, 38, 39]. Priors derived from deep learning can be expressed as the learning the projection from input observed images to output ideal images. In contrast that the other prior modeling methods described in the former paragraphs need iterative processing, typical deep learning based methods can be processed as just a single projection once training has finished. Deep neural networks also contain prior models in their network architectures themselves. For example, Convolutional Neural Networks (CNN) would have architectural priors such as the local and uniform convolutions, and the shift-invariance of features.

Actually, several methods introduce effective priors not only by learning but also by the architectures and loss functions. For example of super-resolution, SRGAN [40] trains and

generates photorealistic images. The architecture is modeled with several contrivances to avoid the blurred result. LAPGAN [41] and LapSRN [42] exploits the Laplacian pyramid of images, where the high-resolutional images can be well represented as straightforward hierarchy summations of generated high-frequency patterns and a low-resolutional image. These methods with additional architectural priors could outperform simple CNN based methods.

The deep neural networks are widely leveraged also in medical applications, and their outperformences are reported [43, 36, 44, 45]. Learning-based methods, however, can achieve the ideal solution only when sufficient number and variation of training data is available and the target image can be regarded drawn from the probability distribution the training data represent. As for the super-resolution, for example, given a set of sufficiently large number of training CT images of healthy subjects, one can improve the spatial resolution of a CT image of a new healthy subject well but would be difficult to improve the resolution if a CT image of a subject with tumors. It should be noted that, in medical image processing, collecting sufficient number and variety of medical images for the training is challenging [43]. If the training dataset does not support the target image so well, the resultant solution is imposed a wrong bias derived from nonsupporting training data. It would be also difficult for humans to grasp the bias because of the black-box problem of deep learning.

1.3 Subjects in this paper

In this paper, we discuss how to introduce the priors as the mathematical models based on the modality natures in order to solve inverse problems related to some medical images. We deal with the following inverse problems associated with three principal medical image modalities and their tasks. For each tasks, the appropriate priors based on their respective natures are designed and introduced in the form of the optimization problems to be solved.

1. Super-resolution of MR images
2. PET image reconstruction
3. Blind-deconvolution of pathological images

For each task, the framework for solving the optimization problem is also mentioned or proposed. The remainder of this paper is organized as the followings.

In Chapter 2, super-resolution of MR images is argued. The proposed method leverage the prior that the backgrounds of MR images are essentially zero. The existing super-resolution method called the Gerchberg method can be effectively applied when the observation model and the fact that the object-backgrounds are zero is already known.

However the Gerchberg method does not assume the existence of the observation noise and the ringing artifacts. In the proposed method, the trouble is shotten by introducing the sparse regularizations. The main contribution includes how to introduce the sparse regularization to the algorithm.

In Chapter 3, the PET image reconstruction is argued. The PET image reconstruction is a task to reconstruct the PET image from a sinogram data. The sinogram is the raw projected data of a patient's body by the PET machine. In the proposed method, the PET observation model, the image spatial model, and the temporal model are introduced as the prior for the PET image reconstruction. In the PET observation model the data fidelity is evaluated with the Kullback-Leibler divergence based on the Poisson-distributed observation. In the image spatial model, the PET image is represented using the basis vectors, which leverage the priors of the nonnegativity and that the backgrounds are zero. As for the temporal model, the compartment model which is used in PET dynamics analysis area is introduced. The PET compartment model is a mathematical model that represents the dynamics of ligands injected in the patient's body while imaging. The compartment model is used in the proposed method to impose the temporal change of PET voxel values.

In Chapter 4, the blind deconvolution of the pathological images is argued. The objective is to remove the blur which we do not know profiles on. The problem to be solved is highly ill-posed as well as nonconvex because both the blur and the ideal image is unknown and lacking in constraints. In order to solve the problem, the high-resolutional and continuous sub-slice model is leveraged using the smooth deformation mapping. The priors based on edge-sparseness and the blur property are also introduced.

For each study above, the performance of the proposed method is evaluated using both synthetic and clinical data. The behaviors of the proposed models and their solvers are discussed comparing with those of the conventional methods.

Chapter 2

Super-Resolution of MR Images

MR imaging is one of the most important methods for observing 3D soft tissues with high contrast[2, 46, 47, e.g.]. However in order to assure sufficiently high SNR, MR images often have anisotropic spatial resolution: The spatial resolution along the through-slice direction is lower than the resolution along the in-plane direction as shown in Figure 2.1. The spatial resolution along the through-slice direction is mainly determined by the slice thickness and there is a trade-off between the slice thickness and the SNR of MR images. Increase of the slice thickness would degrade the spatial resolution along the through-slice direction, though the SNR of each slice image would be improved by the increase of the slice thickness because the quantity of hydrogen nuclei included in the measured slice increases and the magnitude of the signals emitted by the hydrogen nuclei also increases. This is a reason why slice thickness is often set as thick as several times the pixel size and the spatial resolution along the through-slice direction is lower than that of slice images. Therefore, the objective is to reconstruct isotropic MR images whose spatial resolution is equally high along any directions, given anisotropic MR images.

In this study, the new object super resolution technique is proposed to super-resolve MR images. The proposed method is a framework for incorporating the Gerchberg algorithm, which is one of the POCS methods, into a regularized optimization based method of super-resolution. In this framework, we can use the knowledge of the outer contour of a target and of the measured frequency range with the conventional regularizers simultaneously for computing higher spatial resolution images. Combining a TV regularization with the Gerchberg term, one can suppress ringing artifacts often generated by the Gerchberg method. Here, it should be noted that the incorporation of the Gerchberg method into regularized optimization based methods is not so straightforward because the Gerchberg method obtains high-resolution images not by explicitly minimizing some cost function but by iteratively projecting an image onto convex sets. The main contributions of the present study are as follows: (1) We reformulate the projections in the Gerchberg super-resolution algorithm using linear matrix equations, (2) we formulate a convex optimization

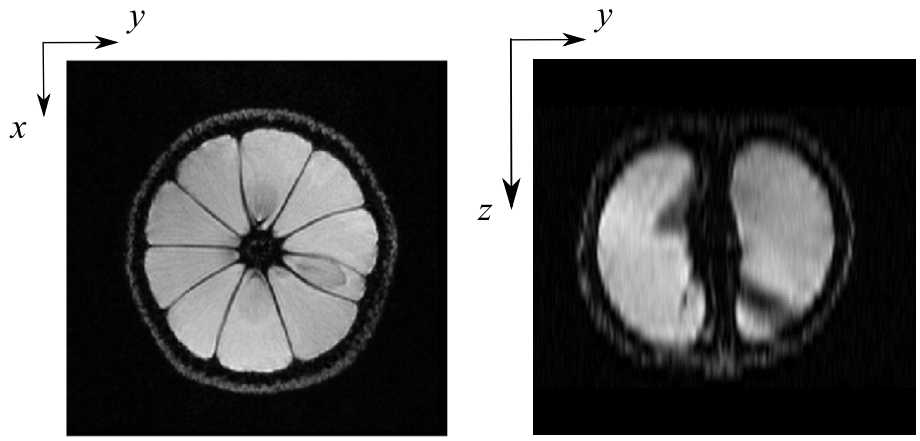


Figure 2.1 An example of observed MR image. The spatial resolution of the through-slice direction (right) is often lower than those of the in-slice directions (left). This is caused by the trade-off between the spatial resolution and SNR of MR images, and high SNR is often preceded.

problem, in which the reformulated projections and the low-TV/low-rank regularization are represented in a cost function and constraints, (3) we explicitly describe the algorithm for solving the convex optimization problem with the alternating direction method of multipliers (ADMM), and (4) we present extensive experimental evaluations conducted using the proposed method.

The remainder of this chapter is organized as follows. First in Section 2.1.1, we state the notations used in this study. We provide a problem statement of the MR-image super-resolution in Section 2.1.2. In Section 2.1.3, we review the Gerchberg algorithm and recent regularization-based approaches. The proposed method and the description of its explicit solvers are explained in Section 2.2. Variational experimental results are presented in Section 2.3 – 2.3.5. Finally, in Section 2.3.6, we discuss the behavior and various aspects of the proposed method.

2.1 Basic Materials

2.1.1 Notations

In this chapter, a vector is denoted by a bold small letter $\boldsymbol{a}$ and a matrix is denoted by a bold capital letter $\boldsymbol{A}$. A 3D tensor is denoted by a bold calligraphic letter $\boldsymbol{\mathcal{A}}$. The (s,t)-th entry of a matrix $\boldsymbol{A}$ is denoted by $\boldsymbol{A}_{st}$ and the (s,t,u)-th entry of a 3D tensor $\boldsymbol{\mathcal{A}}$ is denoted by $\boldsymbol{\mathcal{A}}_{stu}$

Given a vector $\boldsymbol{a}$, the tensor folding operator is denoted by $\mathrm{fold}(\boldsymbol{a}) : \boldsymbol{a} \in \mathbb{R}^{I_1 I_2 I_3 \times 1} \rightarrow \boldsymbol{\mathcal{A}} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$, and its adjoint operator is $\mathrm{vec}(\boldsymbol{\mathcal{A}}) : \boldsymbol{\mathcal{A}} \in \mathbb{R}^{I_1 \times I_2 \times I_3} \rightarrow \boldsymbol{a} \in \mathbb{R}^{I_1 I_2 I_3 \times 1}$. Given

a vector $\boldsymbol{a}$, its matricization is denoted by $\mathrm{mat}(\boldsymbol{a}) : \boldsymbol{a} \in \mathbb{R}^{IJ\times 1} \rightarrow \boldsymbol{A} \in \mathbb{R}^{I\times J}$. Given a tensor $\boldsymbol{\mathcal{A}}$, the i-th mode unfolding operator is denoted by $\mathrm{unfold}_i(\boldsymbol{\mathcal{A}}) : \boldsymbol{\mathcal{A}} \in \mathbb{R}^{I_1\times I_2\times I_3} \rightarrow \boldsymbol{A} \in \mathbb{R}^{I_i\times \prod_{l\neq i} I_l}$, and its adjoint operator is $\mathrm{fold}_i(\boldsymbol{A}) : \boldsymbol{A} \in \mathbb{R}^{I_i\times \prod_{l\neq i} I_l} \rightarrow \boldsymbol{\mathcal{A}} \in \mathbb{R}^{I_1\times I_2\times I_3}$.

Given that $\boldsymbol{A} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^{\mathrm{T}}$ is the singular value decomposition for a matrix $\boldsymbol{A}$, a singular value soft-thresholding operator [30, 31] is defined as

$$\mathrm{SVT}_\tau(\boldsymbol{A}) = \boldsymbol{U}\boldsymbol{\Sigma}_\tau\boldsymbol{V}^{\mathrm{T}}, \tag{2.1}$$

where $\boldsymbol{\Sigma}_\tau = \mathrm{diag}([\max(\sigma_1 - \tau, 0), \max(\sigma_2 - \tau, 0), \cdots, \max(\sigma_I - \tau, 0)]^{\mathrm{T}})$, and σ_ι is the ι-th singular value of $\boldsymbol{A}$. The operator $\circ$ is the Hadamard (element-wise) product.

2.1.2 Problem Statement

Without loss of generality, we can assume that a field of view (FOV) of an MR image is a cubic space. Let the side length of the cubic FOV be denoted by L and let the three mutually orthogonal directions corresponding to the sides of the cubic FOV be denoted by a X-axis, a Y-axis, and a Z-axis.

For simply describing the method, we assume that the slice thickness and the slice spacing are equal and that an MR image consists of n slice images, each of which has $N_c \times N_c$ voxels. It follows that the voxel size along the through-slice direction is given by $M = L/n$ and that the voxel size in each slice image is given by $m \times m$, where $m = L/N_c$. $M > m$ holds in many MR images in order to assure high SNR. [1] Let the scaling factor be denoted by β, where $\beta = M/m = N_c/n$. The spatial resolution along the through-slice direction is β times lower than the resolution along the in-plane directions in an MR image.

In the experiment here, we assume that multiple two MR images are given. When multiple MR images are given, it is assumed that the MR images are obtained with mutually orthogonal directions of slice-selective gradient. Let 3D tensors, $\boldsymbol{\mathcal{X}}_1 \in \mathbb{R}^{N_c\times N_c\times n}, \boldsymbol{\mathcal{X}}_2 \in \mathbb{R}^{N_c\times n\times N_c}$ denote MR images of a same FOV obtained with the slice-selective gradient parallel to the Z-axis, and the Y-axis, respectively. Let a tensor $\boldsymbol{\mathcal{I}} \in \mathbb{R}^{N_c\times N_c\times N_c}$ denote an $N_c \times N_c \times N_c$ isotropic noise-free MR image of the FOV obtained by an ideal MR scanner. It is assumed that any measured MR image of the FOV, $\boldsymbol{\mathcal{X}}_d$, can be generated from $\boldsymbol{\mathcal{I}}$ by appropriately eliminating higher frequency components in the corresponding direction of the slice-selective gradient followed by downsampling by $\beta = N_c/n$.

Let the Fourier transform of $\boldsymbol{\mathcal{X}}_d$ be denoted by $\boldsymbol{\mathcal{F}}_d$ and let Ω_d denote a frequency region only in which the Fourier components of $\boldsymbol{\mathcal{X}}_d$ are measured: Outside of the region, $\bar{\Omega}_d$,

[1] Increase of the slice thickness would degrade the spatial resolution along the through-slice direction, though the SNR of each slice image would be improved by the increase of the slice thickness because the quantity of hydrogen nuclei included in the measured slice increases and the magnitude of the signals emitted by the hydrogen nuclei also increases.

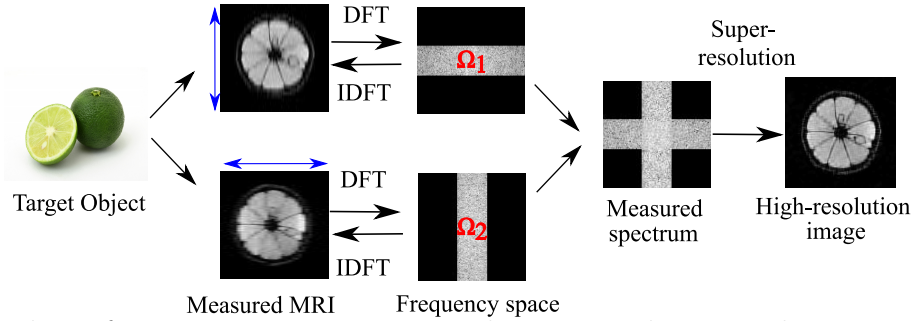


Figure 2.2 Flow of MR image acquisition and super-resolution. The specimen is measured anisotropically from different directions. The blue arrows denote the through-slice directions. MR images are obtained with a narrow bandwidth along the through-slice directions. In most MR images, the spatial resolution along the through-slice direction is more than three times lower than that in the other two directions. There are still unknown high-frequency components even when a number of measurements are available. These components are completed by super-resolution.

in the Frequency space, the frequency components are zero. As shown in Figure 2.2, it should be noted that $\Omega := \Omega_1 \cup \Omega_2$ does not cover the whole spectrum space and that diagonal high-frequency regions are not observed in any of the images. The objective here is to estimate/complete the unknown frequency components and reconstruct a high resolutional MR image.

It should be noted that there is a fundamental difference between the assumptions for our input image and that for compressed sensing MRI. For compressed sensing, it is assumed that the input is random from a sampled k-space, which includes both high and low frequency components in an incoherent manner [11, 12, 13]. By contrast, our input MR images have been taken already and only the low frequency components are given; thus, the completion approaches used for compressed sensing cannot be applied .

2.1.3 Existing Methods for Super-resolution

POCS algorithm

POCS is one of the typical frameworks for the super-resolution [19, 22]. There are various constraints which the groundtruth image must satisfy. In the image domain, some of those constraints are expressed as forms of convex sets where the reconstructed image must be included. A POCS algorithm projects an input image onto the convex sets one by one repeatedly to obtain the unique solution. The convex sets, which we refer also as *models*, vary depended on the various conventional POCS methods. For example, there are methods which employ data fidelity and nonnegativity as the models [48, 22]. We focus on the Gerchberg algorithm, which is one of the earliest POCS algorithms. The Gerchberg algorithm employs two models; The fidelity of the spectrum, and the boundary of the region where the object exists.

In the following of this section, we introduce the Gerchberg algorithm [20, 21]. The Gerchberg algorithm assumes that an image signal is spatially finite and that the outer boundary of the finite region, Γ, is known in advance. In the Gerchberg algorithm, an image is super-resolved by iteratively repeating two projections onto two convex sets: (I) setting the image signal outside of Γ, denoted by $\bar{\Gamma}$, zero; and (II) updating the spectrum within the observed region, Ω, so as to remain as the observed value. An example of the algorithm in the case of a one-dimensional signal is shown in Figure 2.3. The Gerchberg algorithm is summarized in Algorithm 1.

Let $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{N_1 \times N_2 \times N_3}$ denote an image signal and let $\boldsymbol{\mathcal{F}}$ denote its Fourier transform. In the initial state, $\boldsymbol{\mathcal{F}} = \boldsymbol{\mathcal{F}}_0$, where $\boldsymbol{\mathcal{F}}_0$ denotes the observed spectrum. Let $\boldsymbol{\mathcal{P}}_\Gamma \in \{0,1\}^{N_1 \times N_2 \times N_3}$ be a 3D binary label array such that 0 and 1 indicate the outside and inside voxels of the target object, Γ, respectively. The first step of the algorithm is given by $\boldsymbol{\mathcal{X}} \leftarrow \boldsymbol{\mathcal{P}}_\Gamma \circ \mathrm{IDFT}(\boldsymbol{\mathcal{F}})$, where $\mathrm{IDFT}(\cdot)$ denotes a linear operator that provides the inverse 3D discrete Fourier transform (DFT). This operation performs the image signal outside Γ (inside $\bar{\Gamma}$) to zero. Let $\boldsymbol{\mathcal{P}}_\Omega$ denote a 3D index array such that 0 and 1 indicate $\bar{\Omega}$ and Ω, respectively. The second step of the algorithm is then given by $\boldsymbol{\mathcal{F}} \leftarrow \boldsymbol{\mathcal{P}}_\Omega \circ \boldsymbol{\mathcal{F}}_0 + \boldsymbol{\mathcal{P}}_{\bar{\Omega}} \circ \mathrm{DFT}(\boldsymbol{\mathcal{X}})$, where $\mathrm{DFT}(\cdot)$ is a linear operator that provides the 3D-DFT. This operation replaces the calculated spectrum, $\boldsymbol{\mathcal{F}}$, in the region Ω with the observed spectrum, $\boldsymbol{\mathcal{F}}_0$. These two steps are repeatedly conducted and the resultant $\boldsymbol{\mathcal{F}}$ will converge to the unique solution.

The converged unique solution should be the true spectrum under the ideal conditions. The observed spectrum $\boldsymbol{\mathcal{F}}_0$ is interpreted as being the sum of two types of spectra: the true spectrum to be restored and the *error spectrum* that represents the difference between the true spectrum and the observed spectrum (Figure 2.3(a)(b)). It should be noted that $\mathrm{IDFT}(\boldsymbol{\mathcal{F}}_0)$ denotes a low-resolution image that is blurred because the high frequency components are not observed. The blurred image is interpreted as being the sum of the true high resolution image to be restored and the *error image* that is the IDFT of the error spectrum, as shown in Figure 2.3. In step (I), the operator $\boldsymbol{\mathcal{P}}_\Gamma$ reduces only the power of the error image by removing the blur image components in $\bar{\Gamma}$. In step (II), the operator $\boldsymbol{\mathcal{P}}_\Gamma$ has no effect on the true signal, which is zero in $\boldsymbol{\mathcal{P}}_{\bar{\Gamma}}$. Here, one can remove only the energy of the error spectrum by replacing only the spectrum components within Ω with the observed values, $\boldsymbol{\mathcal{F}}_0$ because the true spectrum is observed in the lower frequency region, Ω. Repeating the two projections (I) and (II) described above, the error spectrum converges toward zero and the resulting spectrum converges toward the true spectrum.

In practice, however, it is assumed in the Gerchberg algorithm that the observed low-frequency spectrum is strictly same as that of the groundtruth image. Thus the resultant image reaches to an invalid solution that deviate from the true spectrum when the observed image is contaminated with some noise. It is also assumed that the object exists inner Γ in the image domain. This assumption means that Γ should not invade the true region

where the object of groundtruth exists. On the other hand, if Γ is redundant from the true region, the reconstruction performance would more or less degrade though the algorithm attempt to reach towards the groundtruth. It is also known that the reconstructed image could be contaminated by ringing artifacts even under the ideal conditions [49, 50, 51]. In order to improve the performance, we introduce regularization approaches. In the next sub-section, we describe the introduced regularizers.

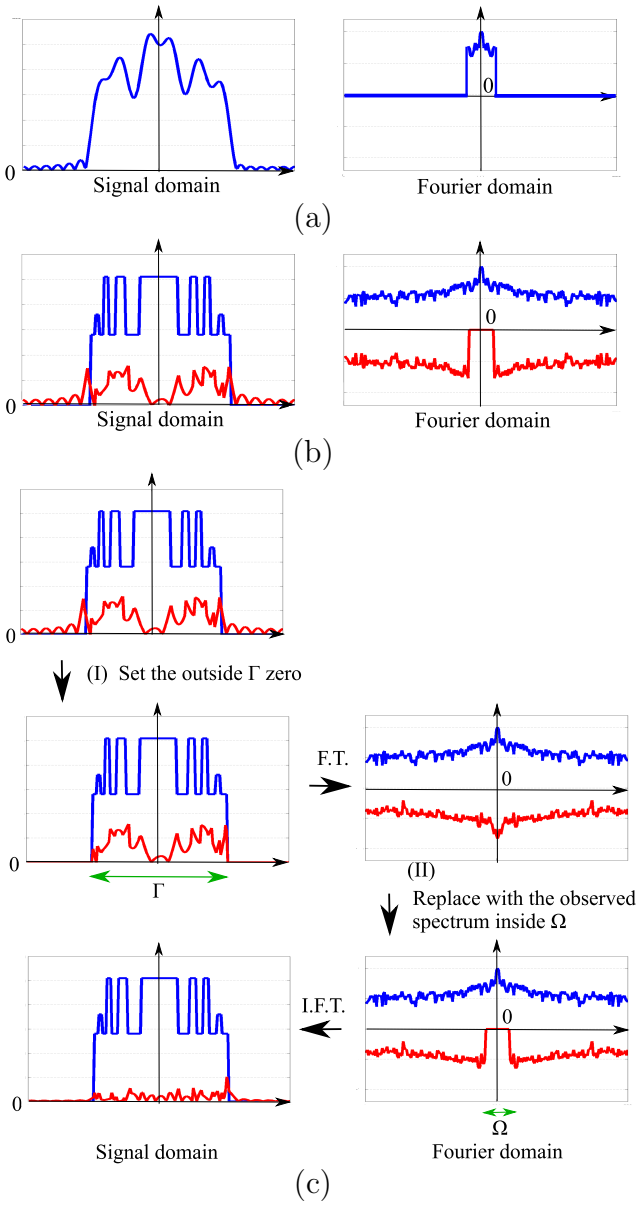


Figure 2.3 Illustration of the Gerchberg algorithm. (a) The observed image (left) and its spectrum (right). (b) The interpretation of (a). The observed image is the sum of the true signal/spectrum (blue line) and the error signal/spectrum (red line). (c) The procedure followed by the algorithm. The error spectrum is reduced by iterating in both the signal and Fourier domains.

Algorithm 1 Gerchberg Algorithm [20, 21]

1: **input:** An observed spectrum $\boldsymbol{\mathcal{F}}_0$, a boundary index array $\boldsymbol{\mathcal{P}}_\Gamma$, and a pass-band index array $\boldsymbol{\mathcal{P}}_\Omega$.
2: $\boldsymbol{\mathcal{F}} = \boldsymbol{\mathcal{F}}_0$;
3: **repeat**
4: $\quad \boldsymbol{\mathcal{X}} \leftarrow \boldsymbol{\mathcal{P}}_\Gamma \circ \mathrm{IDFT}(\boldsymbol{\mathcal{F}})$;
5: $\quad \boldsymbol{\mathcal{F}} \leftarrow \boldsymbol{\mathcal{P}}_\Omega \circ \boldsymbol{\mathcal{F}}_0 + \boldsymbol{\mathcal{P}}_{\bar{\Omega}} \circ \mathrm{DFT}(\boldsymbol{\mathcal{X}})$;
6: **until** $\boldsymbol{\mathcal{F}}$ converges
7: **output:** $\boldsymbol{\mathcal{F}}$ and $\boldsymbol{\mathcal{X}}$

Regularization-Based Methods

In this section, we review conventional regularization-based super-resolution approaches introduced in our method: TV regularization, rank regularization, and their combination. TV is an evaluation measure for the smoothness of an image and its minimization plays an important role in solving the inverse problem in signal processing, such as denoising, interpolation, deconvolution, and super-resolution [52, 53, 54, 55, 56, 57]. Using simple notations, super-resolution problem with TV regularization can be formulated as

$$\arg\min_{\boldsymbol{\mathcal{X}}} \quad D_{\mathrm{S}}(\boldsymbol{\mathcal{X}}_0, \boldsymbol{\mathcal{X}}) + \lambda ||\boldsymbol{\mathcal{X}}||_{\mathrm{TV}}, \tag{2.2}$$

where $\boldsymbol{\mathcal{X}}_0$ is the observed signal, $||\cdot||_{\mathrm{TV}}$ is the total variation, and $D_{\mathrm{S}}(\cdot)$, which is some kind of distance measure between $\boldsymbol{\mathcal{X}}_0$ and $\boldsymbol{\mathcal{X}}$, evaluates the image fidelity. TV is defined as

$$||\boldsymbol{\mathcal{X}}||_{\mathrm{TV}} := \sum_{s,t,u} \sqrt{\sum_{d=1}^{3} [\nabla_d(\boldsymbol{\mathcal{X}}) \circ \nabla_d(\boldsymbol{\mathcal{X}})]_{stu}}, \tag{2.3}$$

where s, t, u are voxel indices for an 3D tensor, and ∇_d is a partial differential operator with respect to the d-th axis of a 3D image.

In many cases, $D_{\mathrm{S}}(\cdot)$ is a linear operator such as L_2-norm for the image fidelity considering the existence of gaussian noise. Thus the problem in (2.2) is often a convex optimization problem. However, classical gradient-based and Newton-like methods cannot be used since $||\boldsymbol{\mathcal{X}}||_{\mathrm{TV}}$ is not a differentiable function. The primal-dual splitting (PDS) method [53, 58, 57], ADMM [59, 60], and the majorization-minimization (MM) algorithm [25] can solve the TV regularization problem in an efficient manner.

For the regularization term, it is also possible to use the low rank property in the image restoration. For tensor completion, regularization with rank is known to obtain superior reconstructions [61]. The rank of a matrix is not a convex function, but its approximation can be minimized as convex optimization using the trace norm [62, 63, 64]. The trace

norm of a matrix is defined as the sum of all the singular values. The rank of a tensor can also be approximated effectively as the trace norm of a tensor, which is defined as the weighted sum of all the matrix trace norms for the individual mode-matricization of a tensor [30]:

$$||\boldsymbol{\mathcal{X}}||_* = \sum_{i=1}^{\mathcal{N}} \alpha_i ||\boldsymbol{X}_{(i)}||_*, \tag{2.4}$$

where $\mathcal{N}$ is the number of tensor dimensions and $\{\alpha_i\}_{i=1}^{\mathcal{N}}$ are parameters that satisfy $\sum_{i=1}^{N} \alpha_i = 1$ and $\alpha_i >$. $\boldsymbol{X}_{(i)} \in \mathbb{R}^{N_i \times N_j N_k}$, where $i, j, k \in \{1, 2, 3\}$ $(i \neq j \neq k)$, is the matrix obtained by $\boldsymbol{X}_{(i)} = \text{unfold}_i(\boldsymbol{\mathcal{X}})$. In the following, $\alpha_i = 1/\mathcal{N}$ and $\mathcal{N} = 3$ are set because 3D MR images are 3D tensors. Then, the 3D tensor completion problem regularized by rank is configured as

$$\arg\min_{\boldsymbol{\mathcal{X}}} \quad D_{\text{S}}(\boldsymbol{\mathcal{X}}_0, \boldsymbol{\mathcal{X}}) + \lambda ||\boldsymbol{\mathcal{X}}||_*, \tag{2.5}$$

where $\boldsymbol{\mathcal{P}}_\Psi \in \{0, 1\}^{N_1 \times N_2 \times N_3}$ indicate the indices where the elements are observed. (2.5) can be optimized by ADMM using the singular value thresholding operator(2.1) [30].

In an application of regularized-based super-resolution for MR imaging, [56] also imposed rank regularization on the problem(2.2) and achieved a satisfactory improvement in performance. They configured the optimization problem as

$$\arg\min_{\boldsymbol{\mathcal{X}}} \quad D_{\text{S}}(\boldsymbol{\mathcal{X}}_0, \boldsymbol{\mathcal{X}}) + \lambda_{\text{TV}} ||\boldsymbol{\mathcal{X}}||_{\text{TV}} + \lambda_{\text{LR}} ||\boldsymbol{\mathcal{X}}||_*. \tag{2.6}$$

In practice, the tensor trace norm can be minimized by using slack variables for each dimension [30, 56].

2.2 Proposed Method

We have introduced two types of super-resolution methods: the Gerchberg algorithm [20, 21] and regularization-based approaches [65, 30, 66, 56]. The Gerchberg algorithm can be characterized by the global boundary prior and the observed spectrum maintenance. By contrast, regularization-based methods can be characterized as performing super-resolution by signal fitting with a local smoothness (low TV) or global similarity (low rank) prior, which is generally satisfied in natural images. The proposed super-resolution algorithm combines both strategies and modifies it by including signal and spectral fitting with smoothness (low TV) and global (low rank and the boundary) priors.

Outline of the Proposed Method

The proposed method is obtained by combining LRTV super-resolution [56] and the Gerchberg algorithm. As mentioned in Section 2.1.3, the Gerchberg algorithm is given in

the form of an iterative projection with $\boldsymbol{\mathcal{P}}_{\Gamma}$, $\boldsymbol{\mathcal{P}}_{\Omega}$, and $\boldsymbol{\mathcal{P}}_{\bar{\Omega}}$, and hence these two methods cannot be combined straightforwardly. Thus, in order to impose regularization technique on the Gerchberg algorithm, we first give a re-interpretation of the Gerchberg algorithm. The Gerchberg algorithm can be re-interpreted as solving the following convex optimization problem for the spectrum $\boldsymbol{\mathcal{F}}$:

$$\begin{aligned} \arg\min_{\boldsymbol{\mathcal{F}}} \quad & ||\boldsymbol{\mathcal{P}}_{\Omega} \circ (\boldsymbol{\mathcal{F}}_0 - \boldsymbol{\mathcal{F}})||_F^2, + i_{\Gamma}(\boldsymbol{\mathcal{X}}), \\ \text{s.t.} \quad & \boldsymbol{\mathcal{X}} = \text{IDFT}(\boldsymbol{\mathcal{F}}), \end{aligned} \tag{2.7}$$

where $\boldsymbol{\mathcal{F}}_0$ is the observed spectrum (See Section 2.2 in detail). In Problem (2.7), $i_{\Gamma}(\boldsymbol{\mathcal{X}})$ is the following indicator function:

$$i_{\Gamma}(\boldsymbol{\mathcal{X}}) = \begin{cases} 0 & \text{if } \boldsymbol{\mathcal{P}}_{\bar{\Gamma}} \circ \boldsymbol{\mathcal{X}} = \boldsymbol{\mathcal{O}} \\ \infty & \text{otherwise} \end{cases} \tag{2.8}$$

where $\boldsymbol{\mathcal{P}}_{\bar{\Gamma}} = 1 - \boldsymbol{\mathcal{P}}_{\Gamma}$, $\boldsymbol{\mathcal{O}} \in \{0\}^{N_1 \times N_2 \times N_3}$ is a zero 3D array. The first term in Problem (2.7) represents fitting the spectrum $\boldsymbol{\mathcal{F}}$ with $\boldsymbol{\mathcal{F}}_0$ for the pass-band, considering the existence of gaussian noise with the observation. The second term, $i_{\Gamma}(\boldsymbol{\mathcal{X}})$, implies that all the outside voxels of the image are zero. Each linear term in (2.7) corresponds to the projection onto the convex set in the signal or Fourier domains in the Gerchberg algorithm.

Based on (2.7) and LRTV regularization, we propose to solve the following convex optimization problem:

$$\begin{aligned} \arg\min_{\boldsymbol{\mathcal{F}}} \quad & \lambda_{\text{TV}}||\boldsymbol{\mathcal{X}}||_{\text{TV}} + \lambda_{\text{LR}}||\boldsymbol{\mathcal{X}}||_* \\ & + i_{\Gamma}(\boldsymbol{\mathcal{X}}) + \frac{1}{2}||\boldsymbol{\mathcal{P}}_{\Omega} \circ (\boldsymbol{\mathcal{F}}_0 - \boldsymbol{\mathcal{F}})||_F^2, \\ \text{s.t.} \quad & \boldsymbol{\mathcal{X}} = \text{IDFT}(\boldsymbol{\mathcal{F}}), \end{aligned} \tag{2.9}$$

where $\lambda_{\text{TV}}, \lambda_{\text{LR}} > 0$ are parameters that control the balance between the respective terms. In contrast to the image-fidelity-based problems in Equations(2.2) and (2.6), the error terms in the proposed method are for fitting the Fourier spectrum. The image/frequency fidelities are regularized/constrained by TV, rank, and the region Γ. The behavior of each term in (2.9) is considered in Section 2.3.6.

The following notice must be considered before the optimization of Problem (2.9). First, the Gerchberg algorithm assumes that the spectrum profile is a rectangular function. However, in clinical MR imaging, the spectrum of the slice-profile forms a Gaussian or windowed-sinc function e.g.[67, 68, 69]. With this notice, we use $\boldsymbol{\mathcal{F}}_0^{'} = \boldsymbol{\mathcal{F}}_0 \oslash \boldsymbol{\mathcal{P}}_{\Xi}$ instead of $\boldsymbol{\mathcal{F}}_0$, where $\boldsymbol{\mathcal{P}}_{\Xi}$ is the spectrum of the slice-profile along through-slice directions and $\oslash$ is the element-wise division operator. Next, a slack variable for each dimension, $\boldsymbol{\mathcal{M}}_i \in \mathbb{R}^{N_1 \times N_2 \times N_3}$ for the optimization process. Considering (2.4) and setting $i_{\Gamma}(\boldsymbol{\mathcal{X}})$ into the

constraints, (2.9) is rewritten as

$$\begin{aligned}\arg\min_{\boldsymbol{\mathcal{F}}} \quad & \lambda_{\mathrm{TV}}||\boldsymbol{\mathcal{X}}||_{\mathrm{TV}} + \frac{\lambda_{\mathrm{LR}}}{\mathcal{N}}\sum_{i=1}^{\mathcal{N}}||\boldsymbol{M}_{i(i)}||_* \\ & + \frac{1}{2}||\boldsymbol{\mathcal{F}}_0^{'} - \boldsymbol{\mathcal{P}}_{\Omega}\circ\boldsymbol{\mathcal{F}}||_F^2, \\ \text{s.t.} \quad & \boldsymbol{\mathcal{X}} = \mathrm{IDFT}(\boldsymbol{\mathcal{F}}), \quad \boldsymbol{\mathcal{O}} = \boldsymbol{\mathcal{P}}_{\bar{\Gamma}}\circ\boldsymbol{\mathcal{X}}, \\ & \boldsymbol{X}_{(i)} = \boldsymbol{M}_{i(i)}, \quad i = 1, \cdots, \mathcal{N}\end{aligned} \tag{2.10}$$

where $\boldsymbol{M}_{i(i)} = \mathrm{unfold}_i(\boldsymbol{\mathcal{M}}_i)$. The constraints can be simplified by introducing a variable $\boldsymbol{V}_i$ with respect to the third constraint. Then, the vector form of the proposed problem (2.10) with relaxation is given by

$$\begin{aligned}\arg\min_{\boldsymbol{f}} \quad & \lambda_{\mathrm{TV}}||\boldsymbol{x}||_{\mathrm{TV}} + \frac{\lambda_{\mathrm{LR}}}{\mathcal{N}}\sum_{i=1}^{\mathcal{N}}||\boldsymbol{M}_{i(i)}||_* \\ & + \frac{1}{2}||\boldsymbol{f}_0^{'} - \boldsymbol{R}_{\Omega}\boldsymbol{f}||_2^2 \\ & + \frac{\epsilon}{2}\sum_{i=1}^{\mathcal{N}}(||\boldsymbol{x} - \boldsymbol{m}_i + \boldsymbol{v}_i||_2^2 - ||\boldsymbol{v}_i||_2^2), \\ \text{s.t.} \quad & \boldsymbol{x} = \boldsymbol{G}\boldsymbol{f}, \quad \boldsymbol{0} = \boldsymbol{R}_{\bar{\Gamma}}\boldsymbol{x},\end{aligned} \tag{2.11}$$

where $\boldsymbol{x} = \mathrm{vec}(\boldsymbol{\mathcal{X}})$, $\boldsymbol{x}_0 = \mathrm{vec}(\boldsymbol{\mathcal{X}}_0)$, $\boldsymbol{f} = \mathrm{vec}(\boldsymbol{\mathcal{F}})$, $\boldsymbol{f}_0^{'} = \mathrm{vec}(\boldsymbol{\mathcal{F}}_0^{'})$, $\boldsymbol{R}_{\Omega} = \mathrm{diag}(\mathrm{vec}(\boldsymbol{\mathcal{P}}_{\Omega}))$, $\boldsymbol{m}_i = \mathrm{vec}(\boldsymbol{\mathcal{M}}_i)$, $\boldsymbol{R}_{\bar{\Gamma}} = \mathrm{diag}(\mathrm{vec}(\boldsymbol{\mathcal{P}}_{\bar{\Gamma}}))$, $\boldsymbol{0} = \mathrm{vec}(\boldsymbol{\mathcal{O}})$. $\boldsymbol{G}$ is a linear operator (matrix) that gives the inverse 3D DFT. $\epsilon > 0$ is an additional parameter for the fitting term of the slack variables. Note that all of the terms and constraints in (2.11) are convex or linear; thus, (2.11) is a convex optimization problem that can be solved using the PDS, ADMM, and MM algorithms.

An Optimization Algorithm

Several algorithms can be used to solve (2.11). In this section, we introduce a convex optimization algorithm that uses ADMM [59] as an example.

To hold proximity, we reformulate the (2.11) as

$$\begin{aligned}\arg\min_{\boldsymbol{f}} \ & \lambda_{\mathrm{TV}}||\boldsymbol{Y}||_{1,2} + \frac{\lambda_{\mathrm{LR}}}{\mathcal{N}}\sum_{i=1}^{\mathcal{N}}||\boldsymbol{M}_{i(i)}||_* \\ & + \frac{1}{2}||\boldsymbol{f}_0^{'} - \boldsymbol{R}_\Omega \boldsymbol{f}||_2^2 \\ & + \frac{\epsilon}{2}\sum_{i=1}^{\mathcal{N}}(||\boldsymbol{x} - \boldsymbol{m}_i + \boldsymbol{v}_i||_2^2 - ||\boldsymbol{v}_i||_2^2), \\ \text{s.t.} \ & \boldsymbol{Y} = [\boldsymbol{L}_1\boldsymbol{x}, \boldsymbol{L}_2\boldsymbol{x}, \boldsymbol{L}_3\boldsymbol{x}], \\ & \boldsymbol{x} = \boldsymbol{G}\boldsymbol{f}, \quad \boldsymbol{0} = \boldsymbol{R}_{\bar{\Gamma}}\boldsymbol{x},\end{aligned} \tag{2.12}$$

where $|||[\boldsymbol{z}_1, \boldsymbol{z}_2, ..., \boldsymbol{z}_N]^{\mathrm{T}}||_{1,2} := \sum_{n=1}^{N}||\boldsymbol{z}_n||_2$ is an $l_{1,2}$-norm and $\boldsymbol{L}_d$ is a partial differential operator with respect to the d-th axis. The first constraint can be rewritten as $\boldsymbol{y} := \mathrm{vec}(\boldsymbol{Y}) = [\boldsymbol{L}_1^T, \boldsymbol{L}_2^T, \boldsymbol{L}_3^T]^T\boldsymbol{x} =: \boldsymbol{L}\boldsymbol{x}$.

The augmented Lagrangian of (2.12) is given by

$$\begin{aligned}\mathcal{L}(\boldsymbol{f},\boldsymbol{x},\boldsymbol{y},\boldsymbol{z},\boldsymbol{\alpha},\boldsymbol{\gamma}) = & \lambda_{\mathrm{TV}}||\boldsymbol{Y}||_{1,2} + \frac{\lambda_{\mathrm{LR}}}{\mathcal{N}}\sum_{i=1}^{\mathcal{N}}||\boldsymbol{M}_{i(i)}||_* \\ & + \frac{1}{2}||\boldsymbol{f}_0^{'} - \boldsymbol{R}_\Omega \boldsymbol{f}||^2 \\ & + \frac{\epsilon}{2}\sum_{i=1}^{\mathcal{N}}(||\boldsymbol{x} - \boldsymbol{m}_i + \boldsymbol{v}_i||_2^2 - ||\boldsymbol{v}_i||_2^2) \\ & + \langle \boldsymbol{z}, \boldsymbol{y} - \boldsymbol{L}\boldsymbol{x}\rangle + \langle \boldsymbol{\alpha}, \boldsymbol{x} - \boldsymbol{G}\boldsymbol{f}\rangle + \langle \boldsymbol{\gamma}, \boldsymbol{R}_{\bar{\Gamma}}\boldsymbol{x}\rangle \\ & + \frac{\rho}{2}||\boldsymbol{y} - \boldsymbol{L}\boldsymbol{x}||_2^2 + \frac{\rho}{2}||\boldsymbol{x} - \boldsymbol{G}\boldsymbol{f}||_2^2 + \frac{\rho}{2}||\boldsymbol{R}_{\bar{\Gamma}}\boldsymbol{x}||_2^2,\end{aligned} \tag{2.13}$$

where $\boldsymbol{z}$, $\boldsymbol{\alpha}$, and $\boldsymbol{\gamma}$ are the Lagrange coefficients, and $\rho > 0$ is the penalty weight. By minimizing (2.13) with respect to $\boldsymbol{f}$, $\boldsymbol{x}$, $\boldsymbol{y}$, and $\boldsymbol{m}_j$, the following update rules can be obtained:

$$\boldsymbol{f}^{k+1} = (\boldsymbol{R}_\Omega + \rho\boldsymbol{I})^{-1}[\boldsymbol{R}_\Omega\boldsymbol{f}_0^{'} + \boldsymbol{G}^T(\boldsymbol{\alpha}^k + \rho\boldsymbol{x}^k)], \tag{2.14}$$

$$\begin{aligned}\boldsymbol{x}^{k+1} = & \left(\rho(\boldsymbol{I} + \boldsymbol{R}_{\bar{\Gamma}} + \boldsymbol{L}^{\mathrm{T}}\boldsymbol{L}) + N\epsilon\boldsymbol{I}\right)^{-1}\left[\epsilon\sum_{i=1}^{N}(\boldsymbol{m}_i - \boldsymbol{v}_i)\right. \\ & \left. -\boldsymbol{L}^{\mathrm{T}}\boldsymbol{z}^k + \boldsymbol{\alpha} + \boldsymbol{R}_{\bar{\Gamma}}\gamma + \rho(\boldsymbol{G}\boldsymbol{f} - \boldsymbol{L}^{\mathrm{T}}\boldsymbol{y}^k)\right],\end{aligned} \tag{2.15}$$

$$[\boldsymbol{Y}^{k+1}]_{st} = \max(1 - \lambda_{TV}\cdot(\rho||\boldsymbol{w}_s^k||_2)^{-1}, 0)w_{st}^k, \tag{2.16}$$

$$\begin{aligned}\boldsymbol{m}_i^{k+1} &= \mathrm{vec}(\boldsymbol{\mathcal{M}}_i^{k+1}) \\ &= \mathrm{vec}(\mathrm{fold}_i[\mathrm{SVT}_{\frac{\lambda_{LR}}{N\epsilon}}(\mathrm{unfold}_j(\boldsymbol{X}^{k+1} + \boldsymbol{V}_i^k))]),\end{aligned} \tag{2.17}$$

$$\boldsymbol{v}_i^{k+1} = \boldsymbol{v}_i^k + \boldsymbol{x} - \boldsymbol{m}_i \tag{2.18}$$

Algorithm 2 LRTVG algorithm

input: Observed images along multiple dimensions $\boldsymbol{\mathcal{X}}_j, j = 1, \cdots, D$, and their combined spectrum $\boldsymbol{\mathcal{F}}_0$.
$\boldsymbol{f}_0^{'} = \mathrm{vec}(\boldsymbol{\mathcal{F}}_0 \oslash \boldsymbol{\mathcal{P}}_\Xi)$, $\boldsymbol{f} = \boldsymbol{f}_0^{'}$, and $\boldsymbol{x} = \boldsymbol{G}\boldsymbol{f}$;
repeat
  Update $\boldsymbol{f}$ based on (2.14);
  Update $\boldsymbol{x}$ based on (2.15);
  Update $\boldsymbol{Y}$ based on (2.16);
  Update $\boldsymbol{m}_i$ based on (2.17);
  Update $\boldsymbol{v}_i$ based on (2.18);
  Update $\boldsymbol{z}$ based on (2.19);
  Update $\boldsymbol{\alpha}$ based on (2.20);
  Update $\boldsymbol{\gamma}$ based on (2.21);
until The cost (2.13) converges
output: $\boldsymbol{\mathcal{F}} = \mathrm{fold}(\boldsymbol{f})$ and $\boldsymbol{\mathcal{X}} = \mathrm{fold}(\boldsymbol{x})$

where $\boldsymbol{I}$ is an identity matrix, $\boldsymbol{w}_s^k = [w_{s1}^k, w_{s2}^k, w_{s3}^k]^T$, and $w_{st}^k = [\mathrm{mat}(\boldsymbol{L}\boldsymbol{x}^{k+1} - \rho^{-1}\boldsymbol{z}^k)]_{st}$. The derivations of the update rules given above are described in Appendix. The Lagrange multipliers are updated by

$$\boldsymbol{z}^{k+1} = \boldsymbol{z}^k + \rho(\boldsymbol{y}^{k+1} - \boldsymbol{L}\boldsymbol{x}^{k+1}), \tag{2.19}$$

$$\boldsymbol{\alpha}^{k+1} = \boldsymbol{\alpha}^k + \rho(\boldsymbol{x}^{k+1} - \boldsymbol{G}\boldsymbol{f}^{k+1}), \tag{2.20}$$

$$\boldsymbol{\gamma}^{k+1} = \boldsymbol{\gamma}^k + \rho(\boldsymbol{R}_{\bar{\Gamma}}\boldsymbol{x}^{k+1}). \tag{2.21}$$

For (2.15), the conjugate gradient method can be used instead of the inverse matrix, which requires a large amount of calculations. The parameters are updated by repeatedly applying (2.14)–(2.21) alternatively until convergence of the original cost function in (2.11). The proposed method with the above notations is summarized in Algorithm 2.

2.3 Results and Discussion

We examined the characteristics of the proposed method using MR images of a brain phantom and of human head portions. The experiments were performed with brain phantom images [70] and with clinical MR images.

For the phantom images, different four images, which vary together in the modality (T1 or T2 weighted) and in the pathological status (with or without lesion), were used. Each phantom image had a spatial resolution of $1 \times 1 \times 1\mathrm{mm}^3$. After setting an original phantom image as the ground truth, we simulated two anisotropic observed images by downsampling toward different orthogonal directions. The blurring kernel for downsampling was

rectangular (average) toward the downsampling direction and we assumed that the slice profile in the signal domain was rectangular along the through-slice direction. Thus, the two observed images had spatial resolutions of $1/\beta \times 1 \times 1\text{mm}^3$ and $1 \times 1/\beta \times 1\text{mm}^3$, where β is the scaling-factor. The variational settings of scaling-factors and noise-levels are simulated for the observations. 16 settings of the observations in total, which vary together in modalities, in pathological status, in scaling-factors, and in noise-levels, were simulated and are shown in Table 2.1.

As for the clinical images, 37 MR images from the OASIS [71] were used for the experiment. MR images of different subjects were randomly chosen from *disc1* in OASIS-1 dataset. For each session (subject) of OASIS, the first scanned image was chosen for the evaluation. Each image had a spatial resolution of $1 \times 1 \times 1.25\text{mm}^3$. We employed the same observation procedure described above.

Using the MR images of a brain phantom, we at first examined the sensitivity of the accuracy of the image reconstruction against the hyper parameters in Section 2.3.1 and the computational time in Section 2.3.2. Then we compared the accuracy of the images reconstructed by the proposed methods and other conventional super-resolution methods and evaluated the reconstruction stability against the change of the noise level and scaling factor in Section 2.3.3 and 2.3.4. As described, our method requires to label the outer boundary of the target in a given image. We also evaluated the sensitivity of the reconstruction accuracy with respect to the accuracy of the labeled outer boundary in Section 2.3.4. Each performance was evaluated based on the peak signal to noise ratio (PSNR) in the target region of the restored images.

The performance of the proposed method is compared with the following existing methods: nearest neighbor interpolation (NN), bicubic interpolation, zero-padding in the Fourier space (ZP) [72], the Gerchberg algorithm [20], TV regularized super-resolution [66], and LRTV [56]. In the remainder of this paper, the proposed method is denoted as LRTVG and the proposed method without the rank regularization term ($\lambda_{\text{LR}} = \epsilon = 0$) is denoted as TVG.

2.3.1 Sensitivity with respect to Hyper Parameters

First, we show behaviors of the parameters in the proposed model. Figure 2.4 shows an example of changes in the PSNR with respect to λ_{TV}, λ_{LR}, and ϵ in (2.9). The changes in the PSNR with respect to λ_{TV} are more steeper than those with respect to λ_{LR}. We can say that TV regularization must be more carefully tuned than LR regularization. ϵ should be set enough small so that the data fidelity term is retained well, as it is shown in Figure 2.4 that larger ϵ rapidly degrades the performance.

The proposed method as well as TV and LRTV needs to be tuned the hyper parameters by the input image. Figure 2.5 shows two examples of the different input images when

λ_{TV} and λ_{LR} are simultaneously changed. As shown in Figure 2.5, λ_{TV} and λ_{LR}, which actually control the regularization weights, should be varied by the image so as to exert the best performance for each image. Although the balance of the two parameters is case by case, it can be said at least that λ_{TV} and λ_{LR} would be larger when the noise-level gets higher, and that λ_{TV} and λ_{LR} would be smaller when the scale-factor gets larger.

With those considerations, λ_{TV}, λ_{LR} were manually tuned by the input image while fixing $\epsilon = 0.01$ in the following experiments.

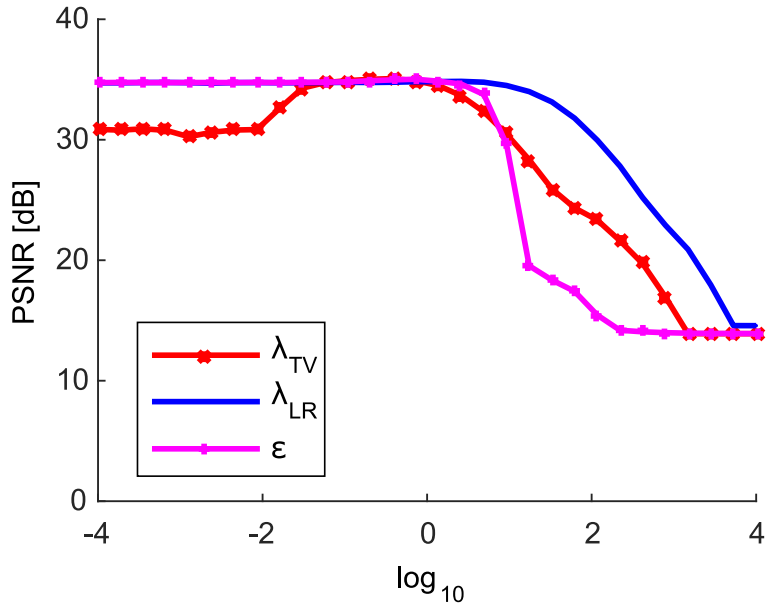


Figure 2.4 An example of PSNR results obtained using the proposed method in terms of the parameters $\lambda_{\mathrm{TV}}, \lambda_{\mathrm{LR}},$ and ϵ. The input image was the Image (a) in Table 2.1.

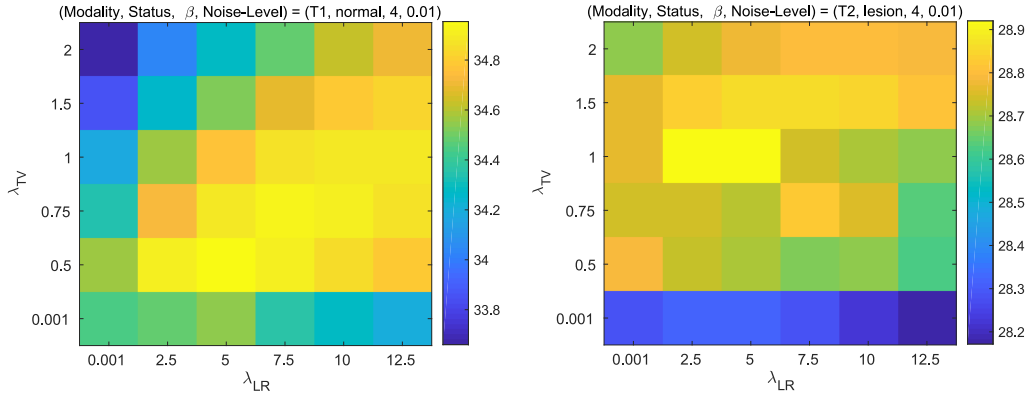


Figure 2.5 Examples of PSNR results of the proposed method when λ_{TV} and λ_{LR} are simultaneously changed. The input images were the Image (a) and (m) in Table 2.1.

2.3.2 Computational Time

Next, we show the number of iterations each method took until the convergence in Figure 2.6, and the average processing times for one iteration in Table 2.2. The number of iteration until the convergence of the proposed method was larger than that of LRTV and TV. We suppose that this is because the additional two Lagrange multipliers, $\boldsymbol{\alpha}$ and $\boldsymbol{\gamma}$, are necessary for the optimization. The Gerchberg model, which also needs $\boldsymbol{\alpha}$ and $\boldsymbol{\gamma}$ when solved with ADMM, took the much more iterations than the other methods. It can be said that regularizations accelerate the convergence speeds of the POCS methods.

Theoretically speaking, the computational orders of LRTVG/TVG for each iteration equal to those of LRTV/TV. The experimental results in Table 2.2 show that computational times of LRTVG/TVG are a little longer than those of LRTV/TV, which is because more several times of FFT are necessary for the proposed method compared with LRTV/TV. We discuss the computational complexity in detail in Section 2.3.6.

2.3.3 Comparison of Accuracy of Reconstruction

In this section, we show the reconstruction accuracy of the proposed method compared to the existing methods: NN, bicubic interpolation, ZP, the Gerchberg algorithm, TV regularized super-resolution, and LRTV super-resolution. The restored images and PSNR results of the simulational observations in Table 2.1 are shown in Figure 2.7 and Figure 2.8. All of the PSNR results were calculated in the region Γ. The parameters of TV, LRTV, and the proposed method are set by manually tuning as mentioned in Section 2.3.1

The simple interpolation methods, i.e., NN, bicubic, and ZP, generated blurred images. The Gerchberg algorithm was affected severely by ringing artifacts and noises, although sharp edges and high frequency components can be observed in the results. Although the TV-based approaches preserved their edges clearly, the results of TV and LRTV lack high frequency components in the Fourier space. The proposed method restored the high frequency components as well as clear edges, and had the best performance for all the input images in Table 2.1. All the PSNR results of T2-weighted images are clearly degraded compared to T1-weighted images. This would be because the image gradients in T2-weighted images more steeply change than those of T1 weighted images because of their modality characteristics.

We also show the reconstruction accuracy of the 37 subjects from OASIS [71]. Box-plots of the results when $\beta = 6$, and when $\beta = 12$ are shown in Figure 2.9. The proposed method performed better than the others in terms of the PSNR. We examined the statistical significance of the performance difference the proposed methods (LRTVG and TVG) and others. In case $\beta = 6$, the proposed LRTVG significantly outperformed all other methods according to the t-test. In case $\beta = 12$, both of LRTVG and TVG significantly outperformed all other methods.

Table 2.1 Correspondences of simulational settings and image indices. N/L of the **Status** row means the normal/lesional brain. Data ID is used to index in Figure 2.7 – 2.8.

Data ID	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)
Modality	T1	T1	T1	T1	T1	T1	T1	T1
Status	N	N	N	N	L	L	L	L
β	4	4	8	8	4	4	8	8
Noise level [%]	1	5	1	5	1	5	1	5
Data ID	(i)	(j)	(k)	(l)	(m)	(n)	(o)	(p)
Modality	T2	T2	T2	T2	T2	T2	T2	T2
Status	N	N	N	N	L	L	L	L
β	4	4	8	8	4	4	8	8
Noise level [%]	1	5	1	5	1	5	1	5

Table 2.2 The average processing times for one iteration. The image size was $220 \times 220 \times 220$.

Methods	LRTVG	TVG	LRTV	TV	Gerchberg
Avg. Process. Time (sec.)	5.73	4.81	5.24	4.54	4.86

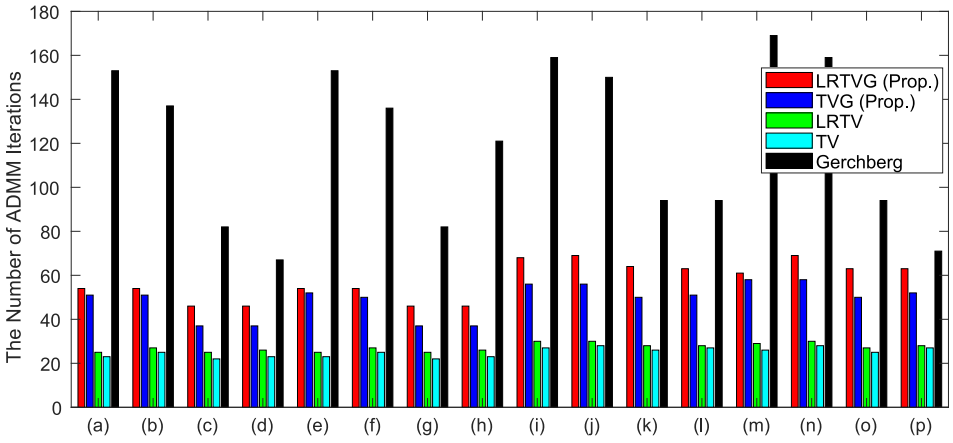


Figure 2.6 The numbers of iterations until the ADMM convergence. Each threshold of the stopping criteria was 1.0×10^{-7}.

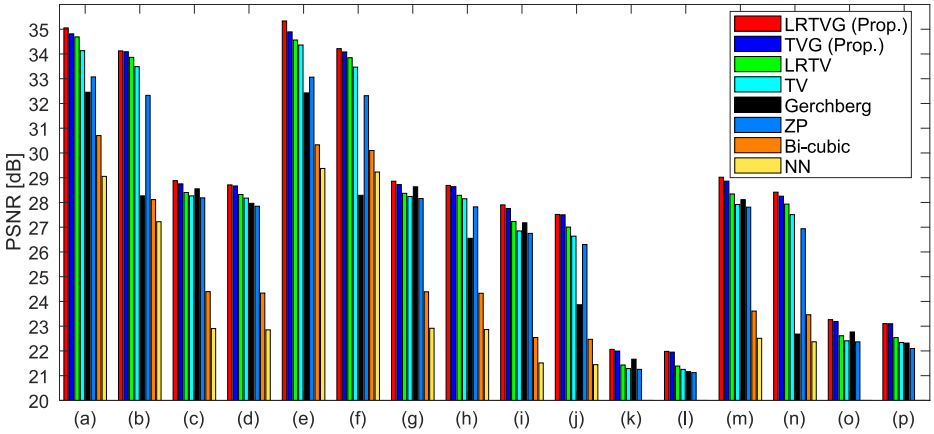


Figure 2.7 PSNR results of the variational simulations in Table 2.1

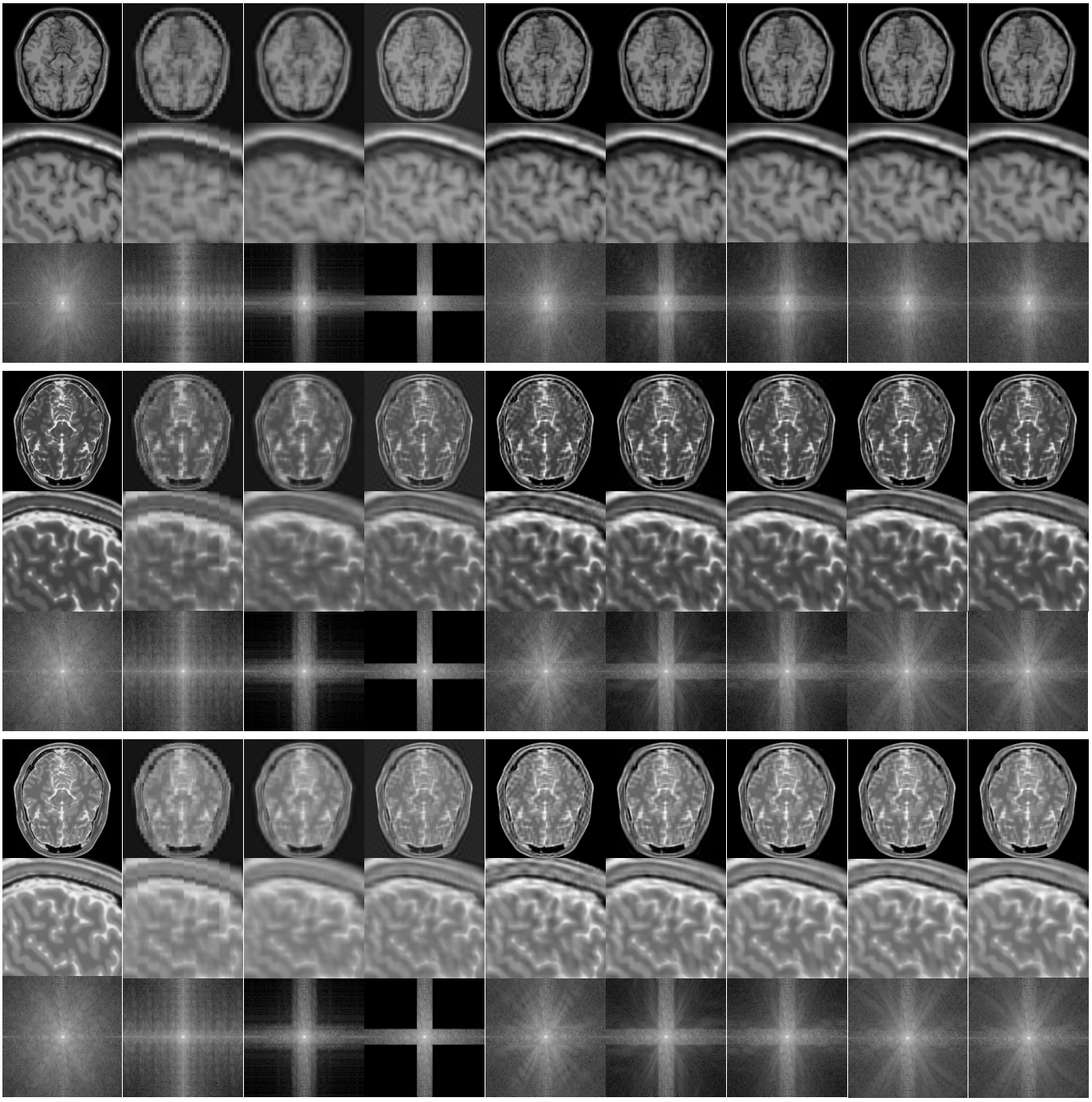
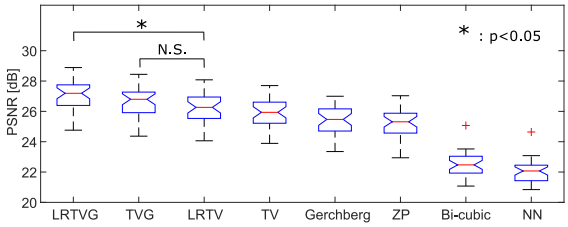
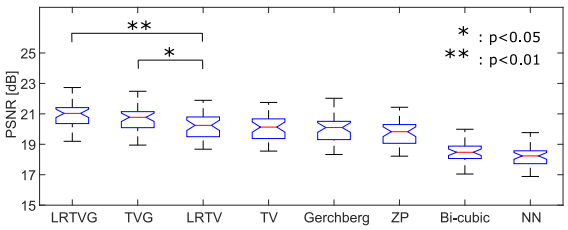


Figure 2.8 Reconstructed images of the brain phantom. Top three rows show illustrations of axial cross-sections, zoomed sagittal cross-sections, and the Fourier spectrums of the T1-weighted brain without lesion (data (c)). The next three rows show those of the T2-weighted brain without lesion (data (k)), and the bottom rows show those of the T2-weighted brain with lesion (data (o)). $\beta = 8$ and the noise-level was 1% for the observation settings.



(a)



(b)

Figure 2.9 Box-plots of the PSNR results obtained using different methods with 37 images from OASIS. The proposed method (LRTVG and TVG) performed better than the other methods. (a) $\beta = 6$, (b) $\beta = 12$.

2.3.4 Stability against Noise Level and Scaling Factor

We evaluated the change of the performance with respect to (i) noise level and (ii) scale factor. Figure 2.10 demonstrates some examples of the experimental results. Figure 2.10 (a) shows PSNR for all noise levels when $\beta = 12$. Figure 2.10 (b) shows PSNR for all scaling-factors when the observations are free of noise. The proposed method outperformed the other methods in all cases. With high noise level, the performance of the proposed method converge next to that of LRTV/TV. With the larger scaling factor, the proposed method performed significantly better compared with the other methods. These behaviors of the proposed method can also be observed in the results in Section 2.3.1, where the larger regularization weights work with high noise levels, and the smaller regularization weights work with large scaling factors.

2.3.5 Stability against Boundary Label

Finally, we focus on the boundary constraint of the proposed method. The boundary contour will differ depending on the boundary detection procedure (e.g., manual, simple thresholding, or contour detection methods), so we examined the performance with respect to the boundary by dilating/shrinking the true boundary. The true boundary was dilated/shrunk by thresholding the distance map created from the level-set function. Figure 2.11 shows the PSNR results as the distance from the true boundary changed. This distance corresponds to the difference in radius between the true boundary and the referred boundary. When the distance was positive (the referred boundary was redundant),

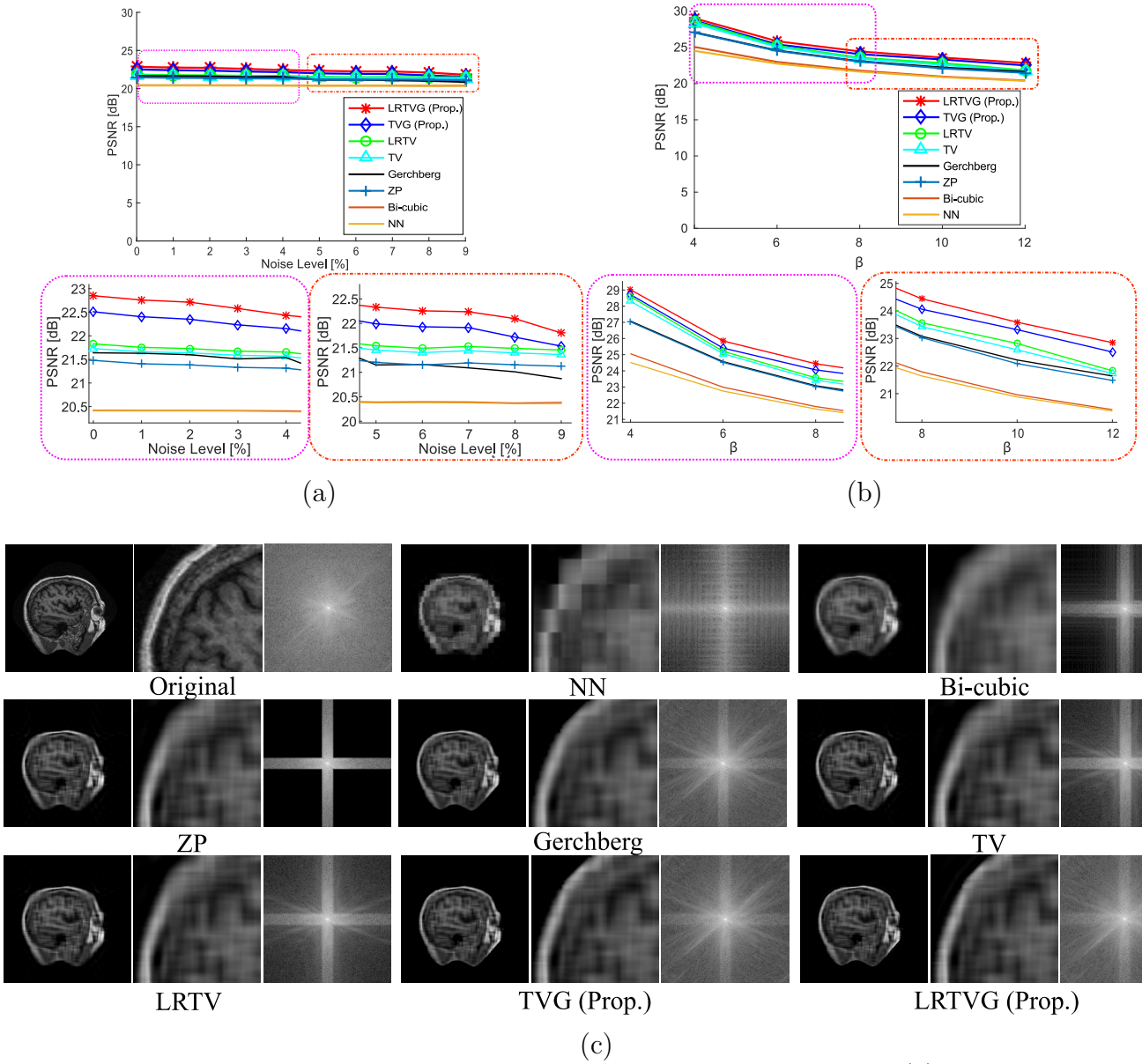


Figure 2.10 Experimental results obtained for an image from OASIS. (a) PSNR in terms of the Gaussian noise level (upper) and enlarged views (lower). (b) PSNR in terms of the scaling factor β (upper), and enlarged views (lower). (c) Reconstructed images and spectra obtained with $\beta = 12$.

the performance increased slightly toward a distance of zero where the boundary was perfectly accurate. The performance decreased steeply when the distance gets negative (the referred boundary was insufficient). This is obviously because not only the background but also the true signal is regarded as noise and the constraint is broken. Thus the target region must not to be underestimated so that the proposed method works. When the boundary was more accurate, the proposed method performed better, but we need to be careful when setting the boundary not to encroach into the true target region.

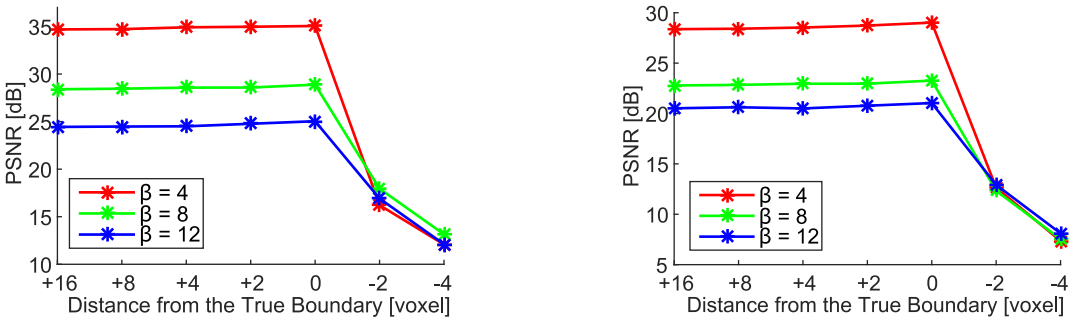


Figure 2.11 Differences in PSNR performance of the proposed method when the boundary became redundant (positive values) or insufficient (negative values) compared with the true boundary (the value is zero). The input image of the left figure was T1-weighted image of normal brain phantom, and that of the right one was T2-weighted image of lesional brain phantom.

2.3.6 Discussion

We consider the aspects of the proposed method and its other features. In terms of POCS approaches [19, 73, 22], the proposed method can be regarded as handling two convex sets: fidelity of the spectrum, and signal boundary. In the proposed method, the projections onto these convex sets can be controlled by TV and rank regularization.

Actually, general regularization-based super-resolution methods simply assume that the resultant image will be smoother or spatially more similar than a noisy input image; there is no assumption for the restoration of fine structures themselves. On the other hand, POCS approaches can retrieve fine structures theoretically as described in Section 2.1.3. However, POCS approaches strictly obey their convex sets, and sometimes they will not be able to compete with images which are out of their model. For example, the Gerchberg model cannot compete with noisy inputs theoretically, and with ringing artifacts which is caused by the discrete Fourier transform. Unlike those methods, the proposed method can allow both the restoration of fine structures, and existence of noise or artifacts.

As same as the general regularization-based super-resolution methods, influence of noise are controlled by regularization terms. The weights for regularization are decided by λ_{TV} and λ_{LR}. As λ_{TV} becomes larger, image gradients will get sparse and the restored image becomes smoother. When λ_{LR} becomes larger, the reconstructed image becomes low rank. Results in Section 2.3.1 showed the behaviors of the two parameters vary be the image. Note that the super-resolution model of the proposed method actually includes the Gerchberg model and equals to it when $\lambda_{\mathrm{TV}} = \lambda_{\mathrm{LR}} = 0$ and $\boldsymbol{\mathcal{P}}_{\Xi}$ denotes rectangular profile.

In addition to the regional constraint limitation described in Section 2.3.5, there are some implicit limitations of the proposed method considered: (i) the PSF is known in advance (the problem to be solved is not the blind deconvolution), (ii) the outer boundary

of the target can be labeled in a reasonable time, and (ii) image noises can be well represented by the normal distribution. The limitations (i) and (iii) are derived from the data fidelity term of the proposed model. TV and LRTV also have these limitations (i) and (iii). Without (i), when the PSF is unknown, the problem to be solved is the nonconvex optimization and cannot be solved easily. When the image noises do not obey the normal distribution, the error function must be corrected by the noise distribution, or some pre-processing is necessary for denoising. The limitation (ii) would be necessary for the clinical applications, and would rarely be broken. When the limitation (ii) is broken, the observed image would be contaminated severely with noises, and the target region/ background cannot be determined easily. However, in the case when the super-resolution is necessary for an input MR image, the slice-thickness is enough large and the noise-level of the observed image would be enough small so that the labeling could be easily conducted.

We also discuss the computational complexity of the proposed method. In the followings, a tensor of the size $N_c \times N_c \times N_c$ is assumed, and the total number of the voxel is N_c^3. First of all, Equations (2.16), and (2.18)–(2.21) cost $O(N_c^3)$ obviously. Equation (2.14) cost $O(N_c^3 \log N_c)$ for the FFT when it is calculated by each fibers of an 3D tensor. When the multiplication of the convolution-matrix and its inversion in (2.15) are calculated in the Fourier space, the cost of the inverse multiplication itself can be reduced to $O(N_c^3)$. Thus (2.15) also costs only $O(N_c^3 \log N_c)$ for the FFT. For (2.17), the size of a matrix, $\mathrm{unfold}_j(\boldsymbol{X}^{k+1} + \boldsymbol{V}_i^k)$, is $N_c \times N_c^2$ and its SVD costs $O(N_c^4)$. Therefore $O(N_c^4)$ for the proximal operator of rank is the worst computational cost for each iterations, which is the same as that of LRTV.

The convergence speed/ number of iterations of the proposed method are slower/ larger than that of LRTV. As mentioned in Section 2.3.2, this would be because the additional two Lagrange multipliers, $\boldsymbol{\alpha}$ and $\boldsymbol{\gamma}$, are necessary for the optimization. The convergence speed also depends on the optimization frameworks and total variation minimization algorithms employed.

There are several future works considered in this study. We proposed the super-resolution model itself and the parameters of the proposed method are hand-tuned so far as discussed above. Actually, there are several methods proposed for automatically tuning the parameters of the TV optimization ([74, e.g.]), and the proposed model also would be able to be applied these auto-tuning methods. For the clinical application, this would be necessary in order to process in shorter time. In order to achieve more accurate results, processing the segmentation of the target region and the super-resolution of the image at the same time can be considered. This can be performed by optimizing $\boldsymbol{X}$, $\boldsymbol{F}$, and $\boldsymbol{P}_{\bar{\Gamma}}$ at the same time. The fact that the fidelity term would explode when the target region is underestimated could be exploited for the simultaneous optimization. However this would lead the model to be an nonconvex optimization problem, which is

much more difficult of initialization and selecting the solvers. The other work to enforce the performance is introducing the regularization based on deep neural networks like [75] and [76], for example. [75] uses trained DCNN and its population as the regularization of the signal fidelity term. [76] exploits the structure itself of deep neural networks for the regularization. The combination of POCS optimization and deep neural networks would lead to the higher performance.

Chapter 3

PET Image Reconstruction

A temporal series of Positron Emission Tomography (PET) images is widely used to analyze the dynamics of ligands in the human brain [1]. A PET image represents a 3D spatial concentration distribution of the ligand and is reconstructed from sinogram data, which are measured by a PET scanner. A sinogram measured by a PET scanner is modelled by the Radon transformation of the spatial ligand distribution, and the Radon transformation can be represented by a linear projection. SNR of sinograms is low in general, and it is known that PET image reconstruction is an ill-posed problem in that the reconstruction is highly sensitive to measurement noises. It is still an important problem to reconstruct accurate PET images from measured sinograms stably against noise.

Most methods for PET image reconstruction solve the ill-posed problem based on regularization. For example, there are many methods where some smooth regularizers are introduced [77, 78, 79, 80, 81, 82]. The works in [83, 84, 82] introduced nonnegativity. The works in [85, 86, 87] introduced the Total Variation (TV) norm of resultant images, or both resultant images and sinograms. These methods reconstruct a PET image from a sinogram measured each time one by one. Several methods have been proposed that reconstruct 4D PET images from a temporal series of measured sinograms at the same time [88, 89, 90]. For example, Reference [90] introduced a set of basis functions for representing the temporal change of each voxel value. The objective of this study is to develop a method that reconstructs accurate 4D PET images from a temporal series of sinograms stably against noise. Analogous to the existing methods, in this study, 4D PET image reconstruction as a constrained optimization problem. Particularly, the consistency between models is focused on, which is introduced for constraining the solution space and the characteristics of the measurement noises, the dynamics of the ligand and the spatial patterns of brain PET images.

Different from typical image sensing, the measurement noises of sinograms cannot be well represented by a normal distribution, but by a Poisson distribution, which obeys

the events of gamma photon emission. This is a reason why the Kullback–Leibler (KL) divergence with a Poisson representation, and not a squared error, should be used for the distance evaluation [91, 92]. As for the noises included in reconstructed PET images, approximation with a Gaussian distribution is acceptable [93, 94] and is widely employed [95, 91]. This means that one can employ the squared error derived from the negative log likelihood when evaluating the distance between two PET images.

The nonnegativity of voxel values in PET images is consistent with the characteristics of PET images and is widely introduced for constraining PET images to be reconstructed [83, 84, 82]. A smooth regularizer such as the TV norm of reconstructed PET images is also widely introduced as a regularizer. It is reported that TV-norm regularization improves the SNR of reconstructed PET images [85, 86, 87]. Lower values of the TV norm can be obtained when the images are partially constant. Hence, TV-norm regularization can smooth out image noises. The method proposed in this study introduces the nonnegativity constraint, but not the TV-norm regularization technique because a TV regularizer can also smooth significant small spot patterns when the regularization is too strong. In order to constrain the spatial patterns in a reconstructed PET image, we constrain the outer boundary of a target body region in the image. We assume that a background region in which no radioligand exists is known. In most clinical PET images, a target body part such as the brain can be observed in a limited region surrounded by an empty background. Many artifacts and noise patterns, however, are often observed in the background in a reconstructed PET image. Thus, the proposed method restricts the voxel values in the background to be zero. One can find that some image super-resolution techniques introduce such an assumption on the empty background and that such an assumption works well [23].

The 4D or spatiotemporal reconstruction of a PET image can improve the image quality by introducing such models that represent the characteristics of the temporal changes of voxel values [89, 80, 90]. A tissue Time Activity Curve (tTAC) represents the temporal change of the ligand concentration at each location. A tTAC should be smooth, although nonsmooth tTACs are often obtained when one reconstructs each PET image captured at different times independently. For example, a method proposed in [90] reconstructs 4D PET images by representing tTACs with linear combinations of spline basis functions. Another method [96] introduces smoothness regularization in order to obtain smooth tTACs. By introducing such models of tTACs, one can improve the accuracy of reconstructed tTACs. The models employed by the existing 4D reconstruction, however, are not necessarily consistent with the dynamics of the ligand in human bodies. For example, even if no measurement noise exists, the linear combination of given spline basis functions would fail to describe tTACs perfectly. Smoothness regularization, which often cuts high-frequency components, would also fail to remove large amounts of noises while preserving the accurate shapes of tTACs. In order to accurately reconstruct tTACs, a

model that accurately represents the receptor-ligand kinetics needs to be introduced. To represent the kinetics, a compartment model, which is widely used to represent receptor-ligand kinetics, is introduced. In the compartment model, each of the compartments corresponds to some component of the human body. The transmission of the ligand between compartments is represented by a system of differential equations. Details of the compartment model will be described later. The proposed method introduces a constraint of tTACs that can be explicitly derived from the multi-compartment model. The proposed method reconstructs 4D images based on the following models:

- Noise model:
  A Poisson distribution model is employed. The distance between the given measured sinogram and the projection of the reconstructed PET image is measured by using the KL divergence.

- Spatial model:
  The nonnegativity of voxel values is assumed. In addition, it is assumed that the foreground/background regions and the reference region in an image are known in advance. The voxel values in the background are set to zero.

- Temporal model:
  A compartment model is employed. It is assumed that, given the ligand, one can identify the number of compartments appropriate for the representation of the kinetics. In this study, we assume a three-compartment model, which is appropriate for a variety of ligands, such as a [^{11}C]carfentanil and a [^{18}F]fludeoxyglucose [97, 98].

In this study, we clearly define the constrained optimization problem to be solved for the 4D PET image reconstruction and clarify a detailed algorithm that can reach the local minimum of the optimization problem.

In summary, the contributions of this study include:

1. A compartment model-based constraint is explicitly introduced in order to constrain the tTACs to the solution space in which the relationships between the tTACs are consistent with the compartment model, while retaining the KL-divergence of data fidelity as a convex function.

2. A constraint of a target region in an image is introduced to restrict the pixel values of the background to be zero.

3. The dependency of the solutions on the initial values is discussed and is experimentally shown.

In the following, first we describe basic materials of mathematics for the PET reconstruction, and then, we describe the compartment model for the PET imaging in Section

3.1. Next, we describe the proposed method in Section 3.2. In Section 3.3, simulation and clinical experimental results are presented. Some behaviors or aspects of the proposed method are discussed in Section 3.3.3.

3.1 Basic Materials

3.1.1 PET Image Reconstruction

In this chapter, let an N-dimensional vector $\boldsymbol{x}_f \in \mathbb{R}^N$ $(f = 1, 2, \cdots, F)$ denote the f-th frame of a temporal series of PET images where N denotes the number of voxels and F denotes the number of frames. Let $t_1 < t_2 < \cdots < t_F$ denote the times when the sinograms are measured. Let an M-dimensional vector $\boldsymbol{y}_f \in \mathbb{R}^M$ denote sinogram data measured by an ideal PET machine at the time $t = t_f$ where M denotes the number of bins.

The measurement model of the sinogram is mathematically represented by a linear projection of the PET image as follows:

$$\bar{\boldsymbol{y}}_f = \boldsymbol{P}\boldsymbol{x}_f \ , \tag{3.1}$$

where the projection matrix is denoted by an $M \times N$ matrix, $\boldsymbol{P}$, and is assumed to be known in advance. The projection by $\boldsymbol{P}$ corresponds to the Radon transform. Let measured sinogram data corresponding to $\bar{\boldsymbol{y}}_f$ be denoted by $\tilde{\boldsymbol{y}}_f$, and let the m-th component of $\bar{\boldsymbol{y}}_f$ and $\tilde{\boldsymbol{y}}_f$ be denoted by $\bar{y}_{f,m}$ and $\tilde{y}_{f,m}$, respectively. It is assumed that each component of the measured sinogram data obeys a Poisson distribution such that:

$$Pr(\tilde{y}_{f,m}|\bar{y}_{f,m}) = \frac{\bar{y}_{f,m}^{\tilde{y}_{f,m}}}{\tilde{y}_{f,m}!}\exp\{\bar{y}_{f,m}\} \ . \tag{3.2}$$

Conventional methods for 3D PET image reconstruction estimate $\boldsymbol{x}_f$ from a measured sinogram $\tilde{\boldsymbol{y}}_f$. Employing the Poisson model (3.2), one can estimate $\boldsymbol{x}_f$ from $\tilde{\boldsymbol{y}}_f$ by minimizing the KL divergence between $\tilde{\boldsymbol{y}}_f$ and $\boldsymbol{P}\boldsymbol{x}_f$ as follows:

$$\hat{\boldsymbol{x}}_f = \arg\min_{\boldsymbol{x}_f} \quad D_{\mathrm{KL}}(\tilde{\boldsymbol{y}}_f||\boldsymbol{P}\boldsymbol{x}_f) \ . \tag{3.3}$$

There are several methods that reconstruct 3D PET images by minimizing the KL divergence with some regularizers of the PET image, as follows:

$$\hat{\boldsymbol{x}}_f = \arg\min_{\boldsymbol{x}_f} \quad D_{\mathrm{KL}}(\tilde{\boldsymbol{y}}_f||\boldsymbol{P}\boldsymbol{x}_f) + \lambda\eta(\boldsymbol{x}_f) \ , \tag{3.4}$$

where $\eta(\cdot)$ is a function to penalize the undesirable characteristics of images and λ is a weight parameter. For example, several methods in [85, 86, 87] set $\eta(\cdot)$ as the total variation. The Expectation Maximization (EM) algorithm and the Ordered Subset EM

(OSEM) algorithms are commonly employed for the minimization. Later in this chapter, a method based on the temporal smooth regularizer is introduced and arranged for a comparison. It is described as follows:

$$\hat{\boldsymbol{X}} = \arg\min_{\boldsymbol{X}} \quad D_{\mathrm{KL}}(\tilde{\boldsymbol{Y}}||\boldsymbol{P}\boldsymbol{X}) + \frac{\lambda}{2}\sum_{f=1}^{F-1}||\boldsymbol{x}_{f+1} - \boldsymbol{x}_f||_2^2 \ , \tag{3.5}$$

where $\boldsymbol{X} = [\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_F] \in \mathbb{R}^{N\times F}$ and $\tilde{\boldsymbol{Y}} = [\tilde{\boldsymbol{y}}_1, \tilde{\boldsymbol{y}}_2, \cdots, \tilde{\boldsymbol{y}}_F] \in \mathbb{R}^{M\times F}$. Problem (3.5) is configured to make the reconstructed image smooth along the temporal change.

3.1.2 Compartment Model

There are several kinds of compartment models that represent ligand dynamics [99, 100, 1]. A three-compartment model [99] consists of three compartments that correspond to plasma, free and bound compartments. The plasma compartment corresponds to the arterial plasma. The ligand transports from the plasma into and back from the free compartment, which is represented by the other two compartments. The free compartment corresponds to the intracellular region, in which the ligand does not bond with receptors.

As shown in Figure 3.1(a), the ligand can transmit only between the plasma compartment and the free compartment. The ligand in the free compartment can also be transmitted to the bound compartment, in which the ligand binds to receptors. Let the ligand concentrations at time t in the plasma, free and bound compartments be denoted by $C_p(t)$, $C_f(t)$ and $C_b(t)$, respectively. The three-compartment models have four parameters: K_1, k_2, k_3 and k_4. The dynamics of the ligand transmission are represented by a system of differential equations with the four parameters, as follows:

$$\frac{dC_b}{dt} = k_3 C_f(t) - k_4 C_b(t), \tag{3.6}$$

$$\frac{dC_f}{dt} = K_1 C_p(t) - k_2 C_b(t) - k_3 C_f(t) + k_4 C_b(t). \tag{3.7}$$

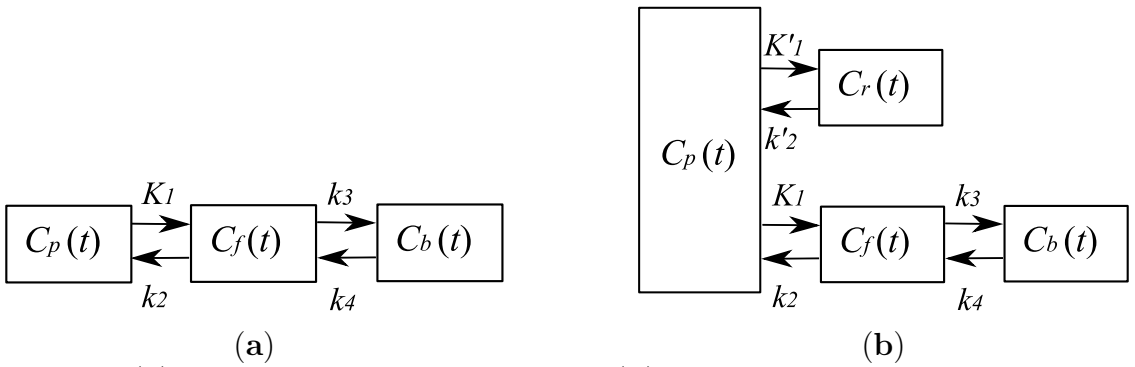


Figure 3.1 (**a**) Three-compartment model. (**b**) Simplified Reference Tissue Model (SRTM).

A PET image captured at time t describes the spatial distribution of $C_b(t) + C_f(t)$, which represents the total ligand concentration in the tissue. A PET image cannot distinguish between the ligand in the free compartment and that in the bound compartment. Let $C(t)$ denote the tTAC, where $C(t) = C_b(t) + C_f(t)$. Given $C_p(t)$, one can represent $\frac{C(t)}{dt}$ as follows:

$$\frac{dC}{dt} = K_1 C_p(t) - k_2 C(t). \tag{3.8}$$

From the above notations, $C(t)$ can be represented with the four free parameters as:

$$C(t) = \frac{K_1}{\theta_2 - \theta_1}\{(k_3 + k_4 - \theta_1)e^{-\theta_1 t} + (\theta_2 - k_3 - k_4)e^{-\theta_2 t}\} * C_p(t), \tag{3.9}$$

where $*$ denotes the convolutional operator, and $\theta_{1,2} = \frac{1}{2}\{k_2+k_3+k_4 \pm \sqrt{(k_2 + k_3 + k_4)^2 - 4k_2k_4}\}$. The above equation is widely used for estimating the values of the kinetic parameters in the brain. For example, in [99, 101], the model was fit to clearly analyze the transfer of ligands in the ROI (Region Of Interest) where tTACs suffer from severe noise. However, a significant weak point is that one needs to measure not only the tTACs $C(t)$, but also the plasma Time Activity Curves (pTACS) $C_p(t)$. The measurement of the pTACs requires the invasive treatment of patients: it requires that patient plasma is sampled for a long time. This is why a Simplified Reference Tissue Model (SRTM) is also clinically used. It is called a reference region in the tissues where no receptor for the ligand exists. Hence, no ligand is bound in the regions. In the compartment of the reference region, the values of k_3 and k_4 are zero. As shown in Figure 3.1(b), let a tTAC in the reference region be denoted by $C_r(t)$. Then, given $C_r(t)$, the tTACs observed at the i-th voxel $C_i(t)$ on the outside of the reference region can be represented as follows [100, 97]:

$$C_i(t) = R_1 C_r(t) + \left\{k_2 - \frac{R_1 k_2}{1 + BP_{ND}}\right\} C_r(t) * \exp\left\{-\frac{k_2 t}{1 + BP_{ND}}\right\}\ , \tag{3.10}$$

where $R_1 = \frac{K_1}{K_1'} = \alpha$ and $BP_{ND} = \frac{k_3}{k_4} = \beta$. The tTACs satisfy the compartment model only when Equation (3.10) is satisfied. When one analyzes a temporal series of PET images using SRTM, one manually labels the reference regions in the measured PET images [102, 103]. The proposed method introduces an SRTM shown in Equation (3.10) for the constraint with tTACs.

3.2 Proposed Method

3.2.1 4D PET Image Representation

Let a temporal series of measured sinograms be denoted by an $M \times F$ matrix, $\tilde{\boldsymbol{Y}} = [\tilde{\boldsymbol{y}}_1, \tilde{\boldsymbol{y}}_2, \cdots, \tilde{\boldsymbol{y}}_F] \in \mathbb{R}^{M \times F}$, and let the spatiotemporal PET images to be reconstructed be denoted by an $N \times F$ matrix, $\boldsymbol{X} = [\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_F] \in \mathbb{R}^{N \times F}$. Let $\mathrm{Row}_i[\boldsymbol{X}]$ denote the

i-th row of a matrix $\boldsymbol{X}$. Each row of $\boldsymbol{X}$ describes a tTAC observed at the corresponding voxel such that:

$$\mathrm{Row}_i[\boldsymbol{X}] = [C_i(t_1), C_i(t_2), \cdots, C_i(t_F)]. \tag{3.11}$$

Let a temporal series of sinograms that corresponds linearly to $\boldsymbol{X}$ be denoted by a $M \times F$ matrix, $\bar{\boldsymbol{Y}} = [\bar{\boldsymbol{y}}_1, \bar{\boldsymbol{y}}_2, \cdots, \bar{\boldsymbol{y}}_F] \in \mathbb{R}^{M \times F}$. Similar to the static case in Equation (3.1), we have:

$$\bar{\boldsymbol{Y}} = \boldsymbol{P}\boldsymbol{X} \ . \tag{3.12}$$

The measured sinograms $\tilde{\boldsymbol{Y}}$ are assumed to be derived from the Poisson distribution of which the average is $\bar{\boldsymbol{Y}}$. Given $\tilde{\boldsymbol{Y}}$, the proposed method estimates the PET images $\boldsymbol{X}$ by solving a constrained optimization problem described below.

3.2.2 Description of Proposed Method

First, we assume that the whole brain region and reference region for an SRTM in a PET image are known. Let the whole and reference regions be denoted by Ω and Γ, respectively ($\Gamma \subset \Omega$).

Let $N_\Omega = |\Omega|$ ($N_\Omega < N$), and let the j-th voxel in Ω ($j = 1, 2, \cdots, N_\Omega$) be denoted by i_j. Let a set of indexes $\mathcal{J} = \{i_j |\ j = 1, 2, \cdots,\ N_\Omega\}$, represent voxels in Ω. An $N \times N_\Omega$ matrix $\boldsymbol{\Psi}$ is defined as $\boldsymbol{\Psi} = [\boldsymbol{e}_{i_1}, \boldsymbol{e}_{i_2}, \cdots, \boldsymbol{e}_{i_{N_\Omega}}]$, where all components of the N-dimensional vector $\boldsymbol{e}_i$ are zero, except that the i-th component of $\boldsymbol{e}_i$ is one: $[\boldsymbol{e}_i]_i = 1$. Then, we can represent an entire PET image $\boldsymbol{x}_f$ using a lower-dimensional vector $\boldsymbol{z}_f$ as follows:

$$\boldsymbol{x}_f = \boldsymbol{\Psi}\boldsymbol{z}_f, \tag{3.13}$$

where $\boldsymbol{z}_f$ is an N_Ω-dimensional vector. The j-th component of $\boldsymbol{z}_f$ is a voxel value of the i_j-th voxel in $\boldsymbol{x}_f$: $[\boldsymbol{x}_f]_{i_j} = [\boldsymbol{z}_f]_j$. Let $\boldsymbol{Z} = [\boldsymbol{z}_1, \boldsymbol{z}_2, \cdots, \boldsymbol{z}_F]$. Then, we have:

$$\bar{\boldsymbol{Y}} = \boldsymbol{P}\boldsymbol{\Psi}\boldsymbol{Z} \ , \tag{3.14}$$

and we can reconstruct PET images by estimating $\boldsymbol{Z}$. By representing PET images with $\boldsymbol{Z}$, one can reduce the dimensions of the image representation and impose that the values in the empty backgrounds of PET images are zero.

Let us assume that the tTAC $C_r(t)$, which is observed in the reference region Γ, is known. Then, following Equation (3.10), a tTAC observed on the outside of Γ is consistent with the compartment model only when $C_{i_j}(t)$ ($i_j \in \mathcal{J}$) can be represented with three parameters $\alpha_j, \beta_j, \gamma_j$, as follows:

$$C_{i_j}(t|\alpha_j, \beta_j, \gamma_j) = \alpha_j C_r(t) + \beta_j \gamma_j C_r(t) * \exp\left\{-\frac{\beta_j(1-\gamma_j)}{\alpha_j}t\right\} \ . \tag{3.15}$$

Let $\boldsymbol{\alpha} = [\alpha_1, \cdots, \alpha_{N_\Omega}]^{\mathrm{T}}$, $\boldsymbol{\beta} = [\beta_1, \cdots, \beta_{N_\Omega}]^{\mathrm{T}}$, $\boldsymbol{\gamma} = [\gamma_1, \cdots, \gamma_{N_\Omega}]^{\mathrm{T}}$, and let a parametric representation of $\boldsymbol{Z}$ in (3.14) be denoted by $\boldsymbol{Z}(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma})$ such that:

$$\mathrm{Row}_j[\boldsymbol{Z}(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma})] = [C_{i_j}(t_1|\alpha_j, \beta_j, \gamma_j), C_{i_j}(t_2|\alpha_j, \beta_j, \gamma_j), \cdots, C_{i_j}(t_F|\alpha_j, \beta_j, \gamma_j)] \quad (j = 1, 2, \cdots, N_\Omega). \tag{3.16}$$

Each row of $\boldsymbol{Z}(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma})$ denotes the tTAC that is consistent with the tTAC $C_r(t)$ observed in the reference region Γ. Then, the problem to be solved here is to estimate the parameter values $\alpha_j, \beta_j, \gamma_j$ $(j = 1, 2, \cdots, N_\Omega)$ that minimize the KL divergence with the measured sinograms:

$$\hat{\boldsymbol{Z}} = \arg\min_{\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}} \; D_{\mathrm{KL}}(\tilde{\boldsymbol{Y}} || \boldsymbol{P}\boldsymbol{\Psi}\boldsymbol{Z}(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma})). \tag{3.17}$$

It should be noted that the reconstructed series of temporal images denoted by $\boldsymbol{\Psi}\hat{\boldsymbol{Z}}$ satisfies the following constraints: (1) all voxel values are nonnegative; (2) all voxel values in the backgrounds ($\bar{\Omega}$) are zero; (3) all tTACs in Ω are consistent with the tTAC in the reference region Γ from the viewpoint of the SRTM; and (4) the distance from the sinogram is minimized. A description of the reconstructed image is shown in Figure 3.2.

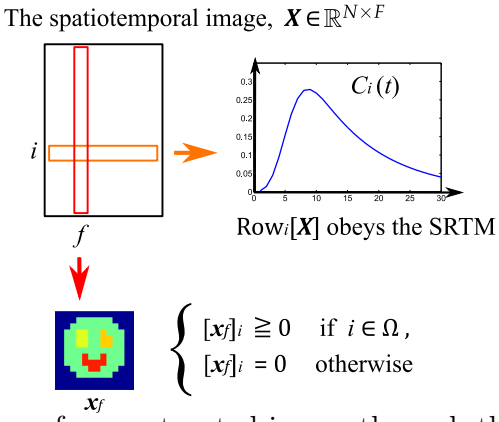


Figure 3.2 Description of reconstructed image through the proposed method.

However, it is difficult to solve the problem in (3.17) directly. The objective function $D_{\mathrm{KL}}(\tilde{\boldsymbol{Y}} || \boldsymbol{P}\boldsymbol{\Psi}\boldsymbol{Z}(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}))$ has $3 \times N_\Omega$ parameters: $\boldsymbol{\alpha} \in \mathbb{R}^{N_\Omega}$, $\boldsymbol{\beta} \in \mathbb{R}^{N_\Omega}$, and $\boldsymbol{\gamma} \in \mathbb{R}^{N_\Omega}$. The PET images $\boldsymbol{X}$ to be reconstructed have $N \times F$ voxels in total, and $3 \times N_\Omega$ is much smaller than $N \times F$. The objective function, however, is not convex in the $3 \times N_\Omega$-dimensional kinetic parameter space. Conventional techniques such as the steepest descent method can fail to obtain the appropriate parameter values.

To solve this problem, the proposed method leverages the fact that the target function $D_{\mathrm{KL}}(\tilde{\boldsymbol{Y}} || \boldsymbol{P}\boldsymbol{\Psi}\boldsymbol{Z})$ is convex with respect to $\boldsymbol{Z}$. We reformulate Problem (3.17) into a constrained optimization problem such that:

$$\begin{aligned} \hat{\boldsymbol{Z}} = \arg\min\nolimits_{\boldsymbol{Z}} \quad & D_{\mathrm{KL}}(\tilde{\boldsymbol{Y}} || \boldsymbol{P}\boldsymbol{\Psi}\boldsymbol{Z}) \;, \\ \text{s.t.} \quad & \boldsymbol{Z} \in S, \quad \boldsymbol{Z} \geq \boldsymbol{0}, \end{aligned} \tag{3.18}$$

where:

$$S = \{\boldsymbol{Z}(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma} | C_r(t))| \ \forall \alpha_j \in R_\alpha, \forall \beta_j \in R_\beta, \forall \gamma_j \in R_\gamma\}. \tag{3.19}$$

$C_r(t)$, which is assumed to be given in advance, is a tTAC in the region Γ. The minimization of the objective function $D_{\mathrm{KL}}(\tilde{\boldsymbol{Y}}||\boldsymbol{P\Psi Z})$ can be transformed as:

$$\begin{aligned} & \arg\min_{\boldsymbol{Z}} \ D_{\mathrm{KL}}(\tilde{\boldsymbol{Y}}||\boldsymbol{P\Psi Z}) = \textstyle\sum_{mf} \left(\tilde{\boldsymbol{Y}}_{mf} \log \frac{\tilde{\boldsymbol{Y}}_{mf}}{[\boldsymbol{P\Psi Z}]_{mf}} - \tilde{\boldsymbol{Y}}_{mf} + [\boldsymbol{P\Psi Z}]_{mf} \right) \\ = \ & \arg\min_{\boldsymbol{Z}} \ \textstyle\sum_{mf} \left(-\tilde{\boldsymbol{Y}}_{mf} \log [\boldsymbol{P\Psi Z}]_{mf} + [\boldsymbol{P\Psi Z}]_{mf} \right) . \end{aligned} \tag{3.20}$$

Here, $D_{\mathrm{KL}}(\tilde{\boldsymbol{Y}}||\boldsymbol{P\Psi Z})$ is convex with respect to $\boldsymbol{Z} \in \mathbb{R}^{N_\Omega \times F}$ and S is a nonconvex set that can be represented by a $3 \times N_\Omega$-dimensional manifold in the $N_\Omega \times F$-dimensional space of $\boldsymbol{Z}$. The proposed method solves the problem (3.18) by means of a projected gradient descent. That is, the constrained optimization problem is decomposed into two subproblems. One is a convex optimization problem, such that:

$$\begin{aligned} \arg\min_{\boldsymbol{Z}} \quad & D_{\mathrm{KL}}(\tilde{\boldsymbol{Y}}||\boldsymbol{P\Psi Z}) \ , \\ \text{s.t.} \quad & \boldsymbol{Z} \geq \boldsymbol{0}, \end{aligned} \tag{3.21}$$

and the other is a problem of a projection onto the manifold, which is not a convex, but a tractable optimization problem:

$$\begin{aligned} \arg\min_{\boldsymbol{Z}} \quad & ||\breve{\boldsymbol{Z}} - \boldsymbol{Z}||_F^2 \ , \\ \text{s.t.} \quad & \boldsymbol{Z} \in S, \end{aligned} \tag{3.22}$$

where $\breve{\boldsymbol{Z}}$ is a current value of $\boldsymbol{Z}$ before the projection. The proposed optimization algorithm consists of two steps, i.e., a gradient step and a projection step, as shown in Figure 3.3. These two steps are applied in order at each iterative update of $\boldsymbol{Z}$. Let k ($k = 1, 2, \cdots$) denote the number of the update iteration, and let $\boldsymbol{Z}^{(k)}$ denote the initial value of $\boldsymbol{Z}$ at the k-th iteration. The update rule at the gradient step is derived from the gradient of the subproblem (3.21) as follows:

$$\breve{\boldsymbol{Z}}^{(k)} = \mathcal{G}[\boldsymbol{Z}^{(k)}] = \boldsymbol{Z}^{(k)} - \epsilon^{(k)} \frac{\partial D_{KL}}{\partial \boldsymbol{Z}^{(k)}}, \tag{3.23}$$

where $\epsilon^{(k)}$ is the step size. The update rule at the projection step solves the problem (3.22) as follows:

$$\boldsymbol{Z}^{(k+1)} = \mathcal{P}[\breve{\boldsymbol{Z}}^{(k)}] = \arg\min_{\boldsymbol{Z} \in S} \ ||\breve{\boldsymbol{Z}}^{(k)} - \boldsymbol{Z}||_F^2 \ , \tag{3.24}$$

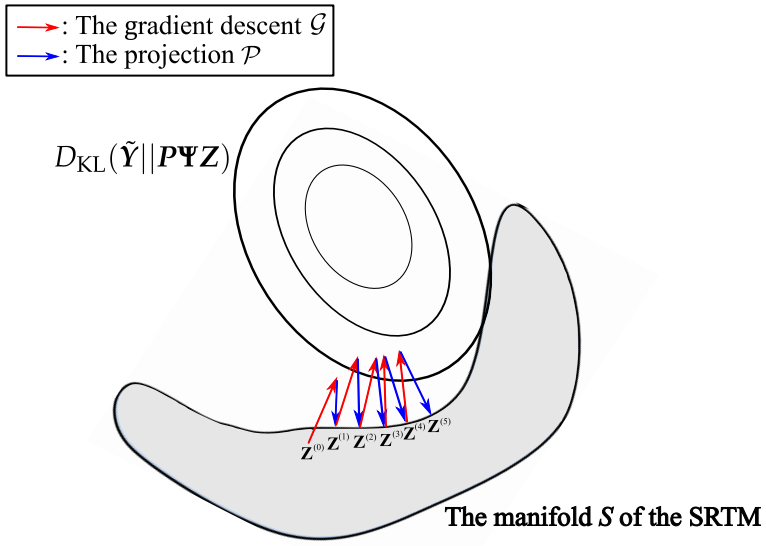


Figure 3.3 Illustration of the proposed algorithm. Red arrows indicate multiplicative update operations $\mathcal{G}$, and blue arrows indicate projections onto sets spanned by the SRTM $\mathcal{P}$.

In summary, the two update rules are applied to $\boldsymbol{Z}$ to obtain $\boldsymbol{Z}^{(k+1)}$ as follows:

$$\boldsymbol{Z}^{(k+1)} = \mathcal{P} \circ \mathcal{G}[\boldsymbol{Z}^{(k)}]. \tag{3.25}$$

The details of $\mathcal{G}$ and $\mathcal{P}$ are explained in Section 3.2.2 and 3.2.2.

The Gradient Step, $\mathcal{G}$

The gradient of the objective function in (3.21) is:

$$\frac{\partial D_{KL}}{\partial \boldsymbol{Z}} = \boldsymbol{\Psi}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\mathbf{1} - \boldsymbol{\Psi}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}(\tilde{\boldsymbol{Y}} \oslash [\boldsymbol{P}\boldsymbol{\Psi}\boldsymbol{Z}]) \ , \tag{3.26}$$

where $\mathbf{1} = \{1\}^{M \times F}$ and $\oslash$ denotes the element-wise division operator. Although the update rule in (3.23) is a basic form for a gradient descent, it is difficult to set each step size appropriately for $D_{\mathrm{KL}}(\tilde{\boldsymbol{Y}}||\boldsymbol{P}\boldsymbol{\Psi}\boldsymbol{Z})$. For $\boldsymbol{Z} \geq \mathbf{0}$, we can employ a multiplicative update rule [104] to effectively solve Problem (3.21). From (3.26), the multiplicative update rule for $\boldsymbol{Z}^{(k)}$ described by $\mathcal{G}$ is derived as follows:

$$\mathcal{G}[\boldsymbol{Z}^{(k)}] = \boldsymbol{Z}^{(k)} \circ (\boldsymbol{\Psi}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\{\tilde{\boldsymbol{Y}} \oslash [\boldsymbol{P}\boldsymbol{\Psi}\boldsymbol{Z}^{(\boldsymbol{k})}]\}) \oslash [\boldsymbol{\Psi}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\mathbf{1}] \ , \tag{3.27}$$

where $\circ$ denotes the Hadamard (element-wise) product. This update rule is analogous to that of EM [105, 106].

The Projection Step, $\mathcal{P}$

In order to obtain the projection $\mathcal{P}$ in (3.24), a tTAC $C_{i_j}(t|\alpha_j, \beta_j, \gamma_j)$, which is consistent with the compartment model as shown in (3.15), is fitted to $\mathrm{Row}_j[\breve{\boldsymbol{Z}}^{(k)}]$. Given a reference signal $C_r(t)$, $\breve{\boldsymbol{Z}}^{(k)}$ can be projected onto the manifold S by solving the following minimization problem for each voxel:

$$\underset{\alpha_j,\beta_j,\gamma_j}{\text{minimize}}\ g(\alpha_j, \beta_j, \gamma_j) = \sum_{f=1}^{F} ||z_j(t_f) - C_{i_j}(t_f|\alpha_j, \beta_j, \gamma_j)||_F^2, \tag{3.28}$$

where $z_j(t_f) = \mathrm{Row}_j[\breve{\boldsymbol{Z}}^{(k)}](t_f)$. It should be noted that each voxel in Ω has three parameters to be estimated: α_j, β_j and γ_j. Problem (3.28) can be iteratively solved voxel independently through a gradient descent method as follows:

$$\alpha_j \leftarrow \alpha_j - \epsilon_{\alpha_j} \frac{\partial g_j}{\partial \alpha_j}, \tag{3.29}$$

$$\beta_j \leftarrow \beta_j - \epsilon_{\beta_j} \frac{\partial g_j}{\partial \beta_j}, \tag{3.30}$$

$$\gamma_j \leftarrow \gamma_j - \epsilon_{\gamma_j} \frac{\partial g_j}{\partial \gamma_j}, \tag{3.31}$$

where ϵ_{α_j}, ϵ_{β_j}, ϵ_{γ_j} are step sizes. Each derivative is obtained as (3.32)–(3.37):

$$\frac{\partial g_j}{\partial \alpha_j} = \sum_{f=1}^{F} 2\{z_j(t_f) - \alpha_j C_r(t_f) - \beta_j\gamma_j h_j(t_f|\alpha_j, \beta_j, \gamma_j)\} \left\{ -C_r(t_f) - \beta_j\gamma_j \frac{\partial h_j}{\partial \alpha_j} \right\}, \tag{3.32}$$

$$\frac{\partial h_j}{\partial \alpha_j} = \frac{\beta_j(1-\gamma_j)}{{\alpha_j}^2} \left\{ C_r(t) * t\exp\left[-\frac{\beta_j(1-\gamma_j)}{\alpha_j} t \right] \right\}, \tag{3.33}$$

$$\frac{\partial g_j}{\partial \beta_j} = \sum_{f=1}^{F} 2\{z_j(t_f) - \alpha C_r(t_f) - \beta_j\gamma_j h_j(t_f|\alpha_j, \beta_j, \gamma_j)\} \left\{ -\gamma_j h_j(t_f|\alpha_j, \beta_j, \gamma_j) - \beta_j\gamma_j \frac{\partial h_j}{\partial \beta_j} \right\}, \tag{3.34}$$

$$\frac{\partial h_j}{\partial \beta_j} = -\frac{(1-\gamma_j)}{\alpha_j} \left\{ C_r(t) * t\exp\left[-\frac{\beta_j(1-\gamma_j)}{\alpha_j} t \right] \right\}, \tag{3.35}$$

$$\frac{\partial g_j}{\partial \gamma_j} = \sum_{f=1}^{F} 2\{z_j(t_f) - \alpha C_r(t_f) - \beta_j\gamma_j h_j(t_f|\alpha_j, \beta_{,j}\, \gamma_j)\} \left\{ -\beta_j h_j(t_f|\alpha_j, \beta_j, \gamma_j) - \beta_j\gamma_j \frac{\partial h_j}{\partial \gamma_j} \right\}, \tag{3.36}$$

$$\frac{\partial h_j}{\partial \gamma_j} = \frac{\beta_j}{\alpha_j} \left\{ C_r(t) * t\exp\left[-\frac{\beta_j(1-\gamma_j)}{\alpha_j} t \right] \right\} , \tag{3.37}$$

where $h_j(t|\alpha_j, \beta_j, \gamma_j) = C_r(t) * \exp\left\{ -\frac{\beta_j(1-\gamma_j)}{\alpha_j} t \right\}$. Let $\hat{\alpha}_j, \hat{\beta}_j, \hat{\gamma}_j$ denote the estimated kinetic parameters for the j-th voxel. Then, the projection operator $\mathcal{P}$ is given as follows:

$$\mathcal{P}[\breve{\boldsymbol{Z}}^{(k)}] = \boldsymbol{Z}(\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\gamma}}). \tag{3.38}$$

In practice, it should be noted that the gradient descent scheme (3.29)–(3.31) can be efficiently computed for all voxels in parallel. $C_r(t)$ can be obtained by, for example, averaging temporally-smoothed tTACs in Γ in advance.

In summary, the proposed algorithm is shown in Algorithm 3.

Algorithm 3 Description of proposed method.

1: input: An observed temporal series of sinograms $\tilde{\boldsymbol{Y}}$, the region Ω where the target object exists and the reference region Γ.

2: $\boldsymbol{\Psi}$ is configured based on Ω.

3: Initialize $\boldsymbol{X}$;

4: Initialize $\boldsymbol{Z} = \boldsymbol{\Psi}^{\mathrm{T}} \boldsymbol{X}$;

5: **repeat**

6: Update $\boldsymbol{Z}$ based on Equation (3.38);

7: Update $\boldsymbol{Z}$ based on Equation (3.27);

8: Solve the Problem (3.28) via the steepest descent method based on gradients (3.32)–(3.37);

9: **until** $\boldsymbol{Z}$ converges.

10: output: $\boldsymbol{X} = \boldsymbol{\Psi Z}$.

3.3 Results and Discussion

3.3.1 Evaluation with Simulated Data

We constructed a temporal series of 2D PET images to evaluate the performance of the proposed method. The simulated image $\boldsymbol{X}_{\mathrm{org}}$ is first designed to obey the SRTM. The image consists of four regions where the kinetic parameters $\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}$ are different from each other. Given $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$, we generated a tTAC for each pixel in each region respectively based on (3.15). In order to simulate sinograms with variational Poisson noise levels,

we conducted the following procedure: The sinogram data were first obtained from the simulated image as follows:

$$\bar{\boldsymbol{Y}}_{\mathrm{sml}} = \mu \boldsymbol{P}\boldsymbol{X}_{\mathrm{org}} \ , \tag{3.39}$$

where $\mu > 0$ is a weight parameter that affects the total counts of the sinograms. An artificial measurement of the sinogram, $\tilde{\boldsymbol{Y}}_{\mathrm{sml}}$, was derived from the Poisson distribution shown in Equation (3.2). The noise level of $\tilde{\boldsymbol{Y}}_{\mathrm{sml}}$ can be controlled by changing the value of μ.

Multiple reconstruction methods were applied to restore the original image from the simulated measured sinograms $\tilde{\boldsymbol{Y}}_{\mathrm{sml}}$. The performance was evaluated based on the Normalized Root Mean Squared Error (NRMSE) between the simulated original image $\boldsymbol{X}_{\mathrm{org}}$ and the reconstructed one. A comparison evaluation was performed with the methods of which constraints are different from each other. The comparison methods are summarized in Table 3.1. In Table 3.1, the method NC (Non-Constraint) minimizes the KL-divergence $D_{\mathrm{KL}}(\tilde{\boldsymbol{Y}}||\boldsymbol{P}\boldsymbol{X})$ without any constraints nor regularizers [106]. TR (Temporal Regularization) minimizes the KL-divergence with a regularizer that imposes temporal smoothness [96]. The option "+ spatial bases" denotes that spatial image patterns are represented by $\boldsymbol{\Psi}\boldsymbol{Z}$ as shown in (3.13) in each corresponding method in order to constrain the pixel values outside of the target to be zero. The proposed method minimizes the KL-divergence with respect to the constraint of the manifold and the optimization using the spatial bases. The solver to optimize Problem (3.5) (TR) is summarized in the Appendix. The parameter λ in Problem (3.5), which corresponds to TR in Table 3.1, was manually tuned with regard to the NRMSE performance that is best for each noise level. The variable $\boldsymbol{Z}$ of the proposed method is initialized with the resultant image of TR + spatial bases. The tTAC in the reference region $C_r(t)$ was estimated by averaging tTACs in Γ as:

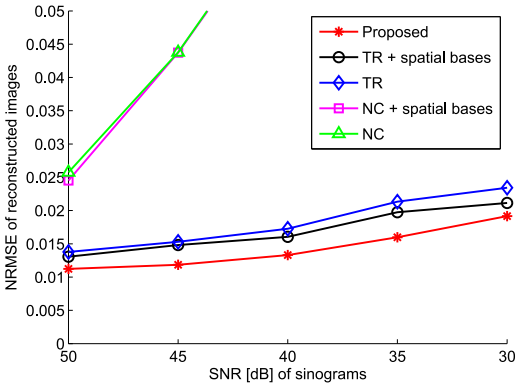
$$[C_r(t_1), C_r(t_2), \cdots, C_r(t_F)] = \frac{1}{|\Gamma|} \sum_{i_j \in \Gamma} \mathrm{Row}_{i_j}[\boldsymbol{X}], \tag{3.40}$$

from the result of TR + spatial bases.

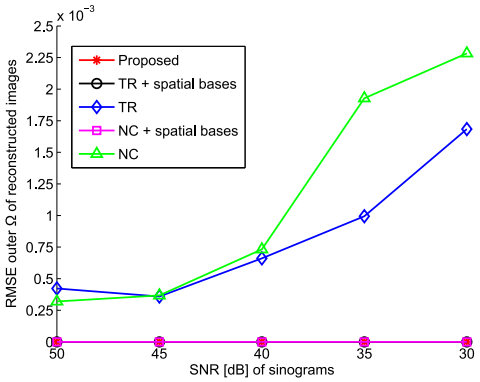
Table 3.1 Correspondence of notations of methods and their constraints for comparison. NC, Non-Constraint; TR, Temporal Regularization.

Notations of Methods	**Problems to Be Optimized**
NC	$\min_{\boldsymbol{X}} \ D_{\mathrm{KL}}(\tilde{\boldsymbol{Y}}\|\|\boldsymbol{P}\boldsymbol{X})$.
NC + spatial bases	$\min_{\boldsymbol{Z}} \ D_{\mathrm{KL}}(\tilde{\boldsymbol{Y}}\|\|\boldsymbol{P}\boldsymbol{\Psi}\boldsymbol{Z})$.
TR	$\min_{\boldsymbol{X}} \ D_{\mathrm{KL}}(\tilde{\boldsymbol{Y}}\|\|\boldsymbol{P}\boldsymbol{X}) + \frac{\lambda}{2}\sum_{f=1}^{F-1}\|\|\boldsymbol{x}_{f+1} - \boldsymbol{x}_f\|\|_2^2$.
TR + spatial bases	$\min_{\boldsymbol{Z}} \ D_{\mathrm{KL}}(\tilde{\boldsymbol{Y}}\|\|\boldsymbol{P}\boldsymbol{\Psi}\boldsymbol{Z}) + \frac{\lambda}{2}\sum_{f=1}^{F-1}\|\|\boldsymbol{\Psi}(\boldsymbol{z}_{f+1} - \boldsymbol{z}_f)\|\|_2^2$.
Proposed	$\min_{\boldsymbol{Z}} \ D_{\mathrm{KL}}(\tilde{\boldsymbol{Y}}\|\|\boldsymbol{P}\boldsymbol{\Psi}\boldsymbol{Z})$, s.t. $\boldsymbol{Z} \in S, \ \ \boldsymbol{Z} \geq \boldsymbol{0}$.

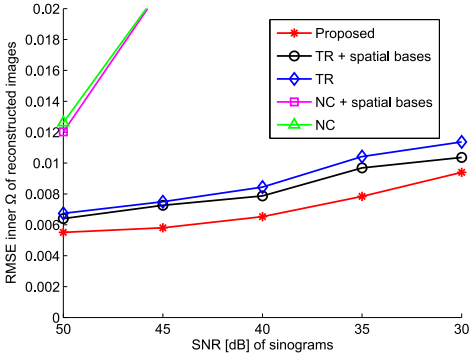
Figure 3.4(a) shows the NRMSEs for various Poisson noise levels. Figure 3.4(b) shows the Root Mean Squared Errors (RMSEs) calculated only in the backgrounds, and Figure 3.4(c) shows the RMSEs calculated only in the target region, Ω. Figures 3.5 and 3.6 show the reconstructed images and tTACs where the SNR of the measured sinograms is 50 dB (Figure 3.5) and 35 dB (Figure 3.6). The SNR of the measured sinograms $\tilde{\boldsymbol{Y}}_{\mathrm{sml}}$ is derived from the error between $\tilde{\boldsymbol{Y}}_{\mathrm{sml}}$ and $\bar{\boldsymbol{Y}}_{\mathrm{sml}}$. The NRMSE of NC deviates with a high noise level in Figure 3.4. The performance was drastically improved using the temporal regularizer. The performance was also improved especially with a high noise level when the spatial bases is introduced to TR. Furthermore, the proposed method, which uses constraints of the SRTM and spatial bases, outperformed the other methods. Not only RMSEs in the backgrounds, but also RMSEs in the target region were reduced in Figure 3.4(b), (c) when the spatial bases are used. The usefulness of the spatial bases and the model constraint was confirmed.



(**a**) The whole image



(**b**) The background $\overline{\Omega}$



(**c**) The target region Ω

Figure 3.4 NRMSE and RMSE results in terms of Poisson noise level. Each graph (**a**–**c**) is different in regions where the error is evaluated: (**a**) NRSME in whole images, (**b**) RMSE only in backgrounds and (**c**) RMSE only in the target region, Ω.

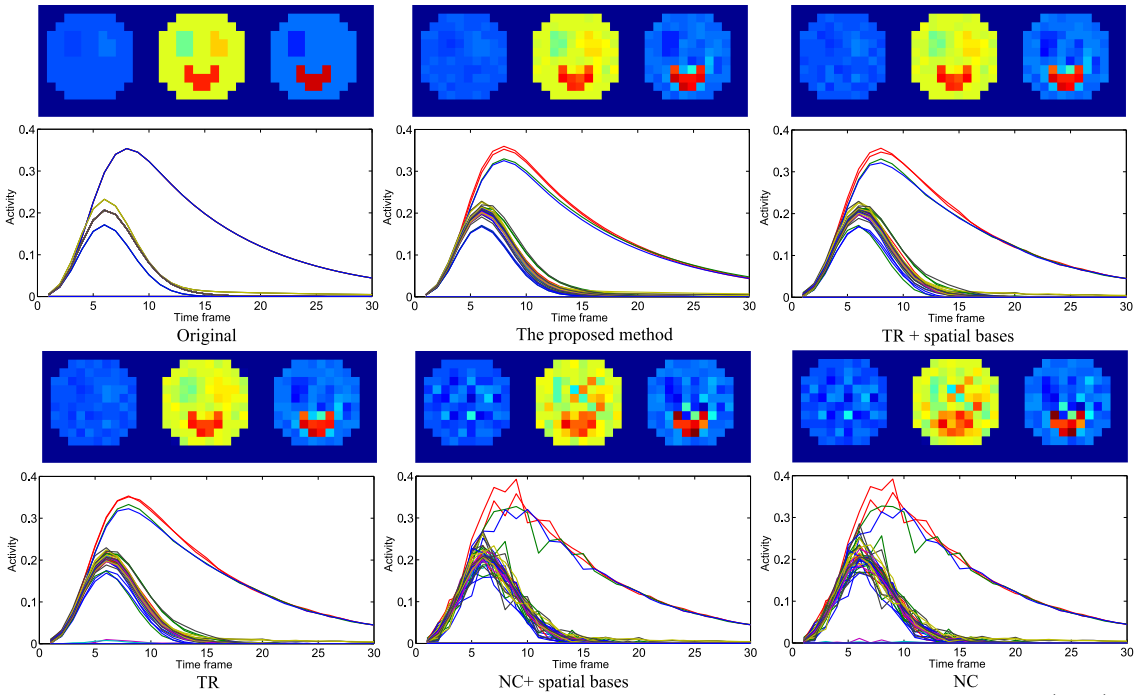


Figure 3.5 Illustrations of three time frames picked from reconstructed image (top) and some corresponding tTACs (bottom) for each method. The SNR of simulated sinograms is 50 dB.

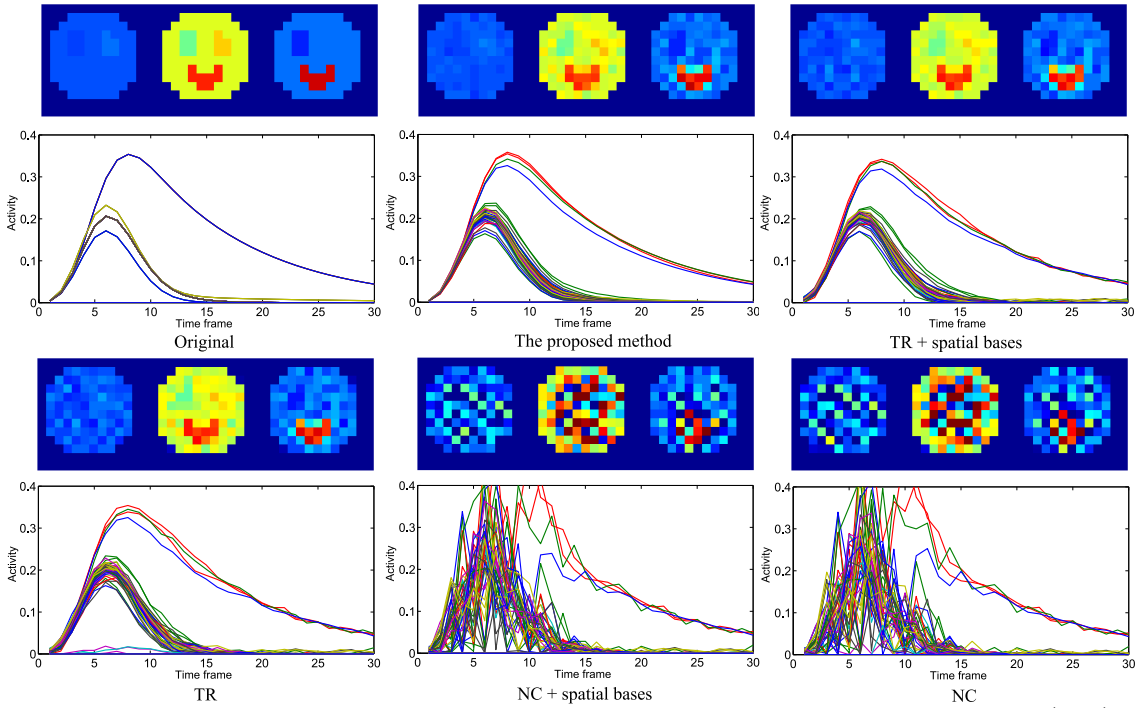


Figure 3.6 Illustrations of three time frames picked from reconstructed image (top) and some corresponding tTACs (bottom) for each method. SNR of simulated sinograms is 35 dB.

Furthermore, Figure 3.7 shows the behavior of the convergence of the proposed method

with respect to the number of the projection obtained from different initial values. For the comparison purpose, we employed four different reconstruction methods, TR, TR + spatial bases, NC, and NC + spatial bases, shown in Table 3.1, all of which reconstruct PET images via convex optimization. Each panel shows four graphs obtained from identical sinogram data with different initialization, and the graphs shown in different panels are obtained from different SNRs of the sinogram. Since the proposed method solves a nonconvex optimization problem, the performance of the proposed method depends on the initial value of $\boldsymbol{Z}$. It can be said the spatial bases are also effective for achieving the better initialization of the proposed method.

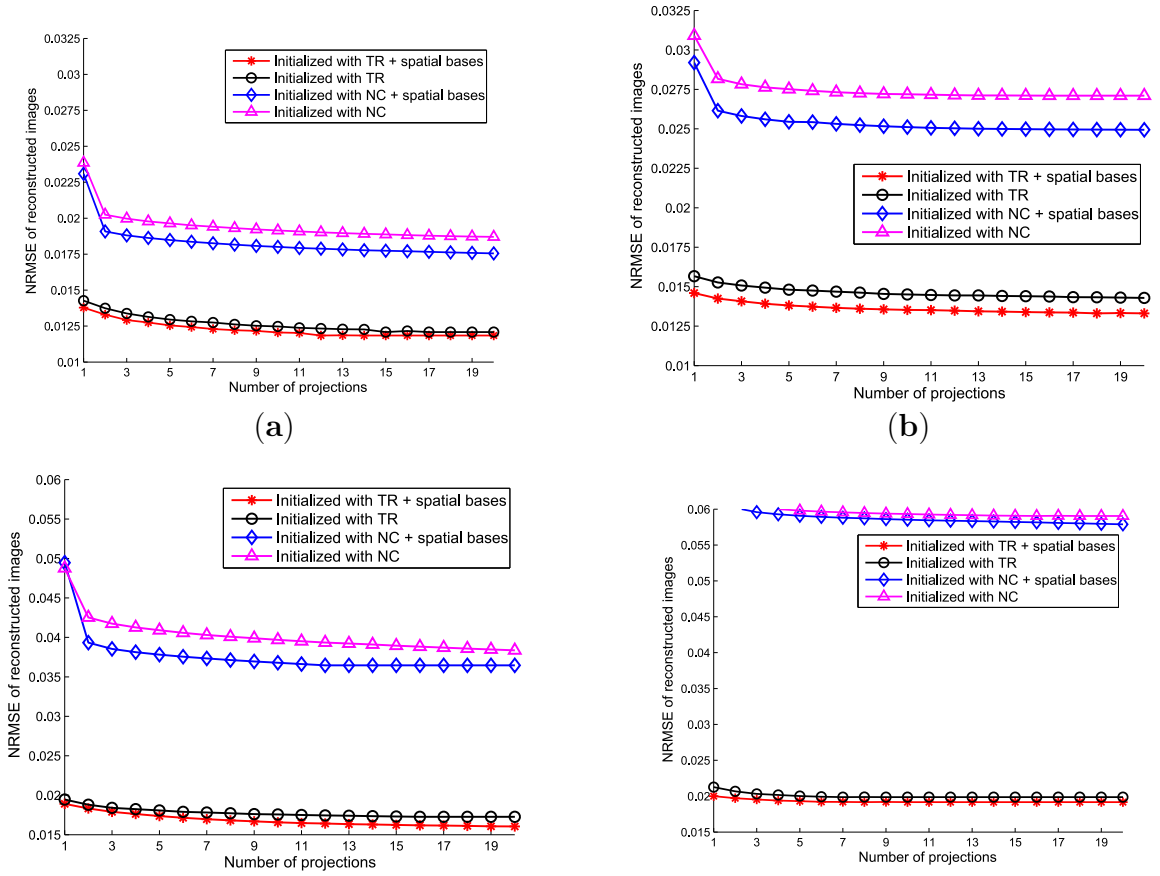


Figure 3.7 NRMSE improvement of proposed method in terms of the number of projections, k. The colors of curves vary in initialization methods, which give initial values of $\boldsymbol{Z}$ to the proposed method. The noise level of $\tilde{\boldsymbol{Y}}_{\mathrm{sml}}$ varies in each graph: (**a**) 45 dB, (**b**) 40 dB, (**c**) 35 dB and (**d**) 30 dB.

3.3.2 Practical PET Images Reconstruction from Clinical Sinograms

We also applied the proposed method to clinical data. Sinogram data were obtained by injecting [^{11}C]carfentanil into a patient. The size of the resultant 3D PET image is $96 \times 96 \times 63$ ($M = 580,608$), and the number of time frames is $F = 26$. The reference region Γ and the target region Ω were manually obtained from the resultant image of OSEM [77, 102, 103]. The tTAC in the reference region $C_r(t)$ was estimated by averaging tTACs in Γ as (3.40). The initial value for the proposed method is given from the result of TR with $\lambda = 0.0025$.

Figures 3.8–3.10 show the reconstructed images of TR with variational parameters ($\lambda = 0.005, 0.0025, 0.001$) and the proposed method. The differences between the proposed method and the others are especially clear in the first and fourth time frames. Although the resultant images of TR vary sensitively with changes in λ, the images reconstructed by the proposed method are both distinct and smooth compared to the other images.

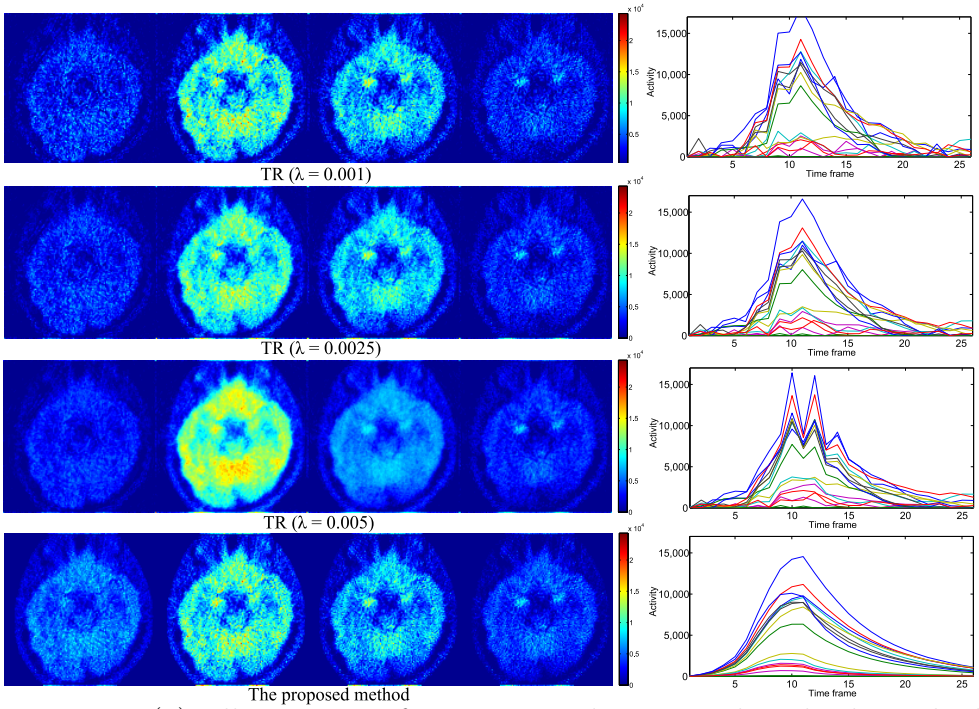


Figure 3.8 PET Data (1): Illustrations of reconstructed image with multiple methods (left) and some of their corresponding tissue Time Activity Curves (tTACs) (right). From left to right, four of 26 time frames (7, 10, 13 and 16) are described for each method (left).

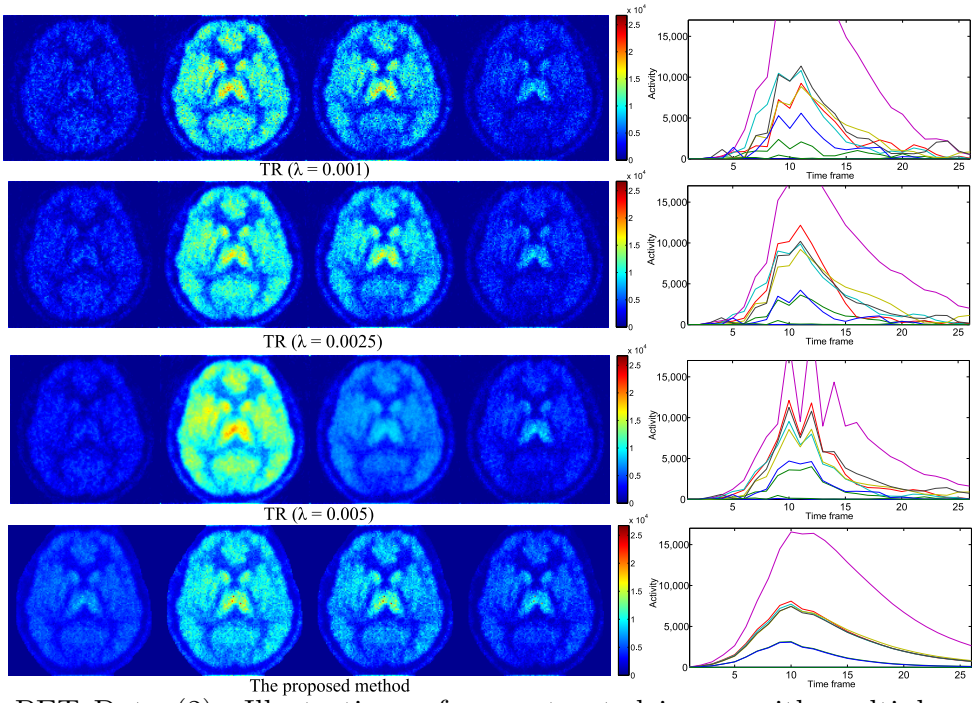


Figure 3.9 PET Data (2): Illustrations of reconstructed image with multiple methods (left) and some of their corresponding tTACs (right). From left to right, four of 26 time frames (7, 10, 13 and 16) are described for each method (left).

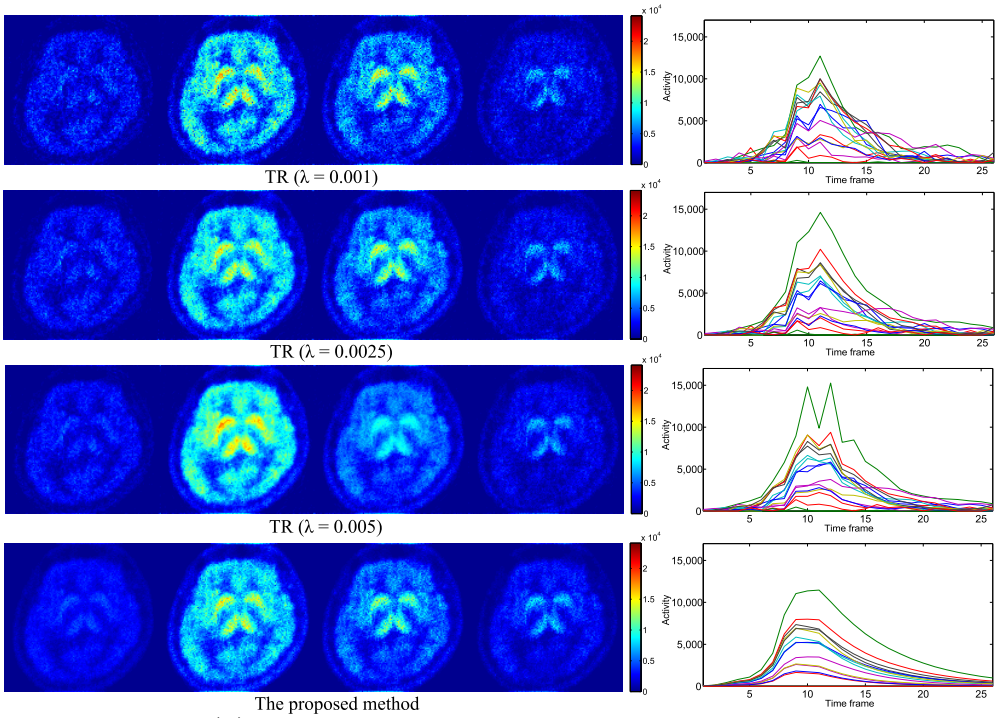


Figure 3.10 PET Data (3): Illustrations of reconstructed image with multiple methods (left) and some of their corresponding tTACs (right). From left to right, four of 26 time frames (7, 10, 13 and 16) are described for each method (left).

We employed the Coefficient of Variation (CV) [107] for the validation of the results obtained from clinical data. The CV value is defined as (3.41) and is evaluated for each ROI (Region Of Interest) that is manually labeled by an expert. In each ROI, the kinetic parameters are known to be identical, and hence, tTACs are ideally the same at every location.

$$\overline{\mathrm{CV}}(\boldsymbol{X}_{\xi}) = \frac{1}{F}\sum_{f=1}^{F} \mathrm{CV}(\boldsymbol{X}_{\xi,f}) = \frac{1}{F}\sum_{f=1}^{F} \frac{\sigma(\boldsymbol{X}_{\xi,f})}{\mu(\boldsymbol{X}_{\xi,f})}, \tag{3.41}$$

where $\boldsymbol{X}_{\xi,f}$ denotes a set of voxel values observed in the ξ-th ROI at the f-th frame and ξ and f denote the indices of the ROI and the frame, respectively. $\sigma(\boldsymbol{X}_{\xi,f})$ and $\mu(\boldsymbol{X}_{\xi,f})$ denote the standard deviation and the mean of $\boldsymbol{X}_{\xi,f}$, respectively. $\overline{\mathrm{CV}}(\boldsymbol{X}_{\xi})$ should be small if 4D PET images are accurately reconstructed. The evaluated values are shown in Table 3.2. The ROIs for which we evaluated CV values are as follows: cerebellum ($\xi = 1$), occipital lobe (#2), frontal lobe (#3), temporal lobe (#4), thalamus (#5), putamen (#6) and caudate nucleus (#7), respectively. The parameter λ for TR was manually tuned to achieve the best results. The proposed method was initialized from the results obtained by TR with $\lambda = 0.0025$. For all ROIs, the proposed method outperformed TR: the images reconstructed by the proposed method have more uniform voxel values in each ROI in each time.

Table 3.2 $\overline{\mathrm{CV}}(\boldsymbol{X}_\xi)$ calculated in each ROI w.r.t. variational methods.

Data	**Method**	**#1**	**#2**	**#3**	**#4**	**#5**	**#6**	**#7**
(1)	TR	0.507	0.663	0.515	0.645	0.660	0.486	0.647
	Prop.	0.174	0.302	0.219	0.236	0.214	0.240	0.304
(2)	TR	0.859	1.01	0.864	0.795	1.166	0.745	0.752
	Prop.	0.136	0.245	0.164	0.124	0.134	0.165	0.153
(3)	TR	0.710	0.812	0.640	0.616	0.618	0.616	0.772
	Prop.	0.240	0.295	0.251	0.199	0.240	0.221	0.351

3.3.3 Discussion

Optimization Strategy

In this study, we considered solving the minimization problem of the KL-divergence between measured sinograms and simulated sinograms, which are modeled with kinetic parameters as shown in (3.17). In addition, we incorporated a boundary prior of a brain region by using nonnegative spatial bases. Note that our objective is different from conventional penalized/regularized methods such as temporal or spatial smoothing because the compartment model strictly satisfies the theoretical consistency of 4D PET images. Furthermore, we do not need to tune any trade-off parameters unlike penalized methods. The problem here is that the cost function in (3.17) is nonconvex with respect to kinetic parameters, and it is very difficult to solve directly.

In order to solve the nonconvex optimization problem (3.17), we have rewritten (3.17) as (3.18), which consists of the convex objective function and the nonconvex constraint set. Note that the problems (3.17) and (3.18) are theoretically equivalent; however, the parameter space is expanded from $3 \times N_\Omega$ to $N_\Omega \times F$. This modification relaxed the nonconvex objective function to the convex objective function which helps to apply the projected gradient method for our problem. The proposed optimization algorithm consists of two steps: the gradient step to minimize the objective function and the projection step to put the 4D PET images on the manifold based on the SRTM.

The proposed algorithm of the projected gradient scheme is similar to those of [105, 90, 108] as also described in Section 3.3.3, but is a more generalized one. The gradient step of the proposed algorithm is based on a gradient descent that substantially includes the EM update used in [90, 108]. Furthermore, the estimation of the kinetic parameters is solved in the proposed method via a gradient descent scheme where each derivative is explicitly used. The conventional methods solve the estimation of the kinetic parameters through numerical analytic approaches where derivatives are implicit and approximated.

Sensitivity to Initialization

As the optimization problem to be solved is nonconvex, the reconstruction results depend on the initial value of the coefficient, $\boldsymbol{Z}$. From Figure 3.7, it can be said that the better the initialization adopted, the better the performance of the proposed method because of the nonconvexity of the manifold. Thus, the proposed method using the SRTM would always outperform the other regularization-based methods where the SRTM is not used as a constraint. As shown in Figure 3.7, the TR-based initialization always reconstructed images that have better NRMSE than the NC-based ones. NC-based methods reconstruct images with less constraints/regularization than TR-based ones, and hence, the images reconstructed by the NC-based methods are more consistent with the measured sinograms that are contaminated by Poisson noises: PET images reconstructed by NC-based methods would more easily overfit to the noisy data, and the NRMSE of these images rapidly increased with respect to the Poisson noise level of the sinogram data, as shown in Figure 3.7. The initialization via NC-based methods would be useful when one prefers PET images that are more consistent with measured data (often contaminated with noise), and the initialization with TR-based methods is recommended if images with less NRMSE should be reconstructed. One can of course introduce the knowledge of the target region for the initialization in order to improve NRMSE.

Related Works

There exist many methods for 4D PET image reconstruction, and our proposed method has strong relationships especially with the methods that constrain tTACs based on the compartment model [109, 110, 80, 89, 105]. First of all, we have to point out that none of these papers explicitly show a constrained optimization problem to be solved, but show only corresponding sub-problems. This fact makes it difficult to discuss the problem of 4D PET image reconstruction and its corresponding solution algorithm. As far as we know, none of the existing compartment model-based methods for the 4D PET image reconstruction except one method [108] proposed by Gravel and Reader (GR) directly use the nonlinear model derived from the compartment model, but use some linearly approximated models for the constraint, as described in their paper [108]. Only the GR method explicitly introduces the nonlinear kinetic model derived from the compartment model to constrain the solution space. The paper [108], though, does not explicitly show the main optimization problem to be solved, and no details of the solution algorithm for satisfying the nonlinear constraint are described. In addition, the sensitivity of the results to the initial values is not discussed. We, on the other hand, clearly show the problem to be solved as in (3.18), discussing the sensitivity to the initialization and experimentally demonstrating the results about the sensitivity. Furthermore, the convergence of the GR method would not be guaranteed. The GR method updates the reference signal $C_r(t)$

every other iteration. This would lead to the manifold spanned by the SRTM being reshaped every other iteration, and projections could sometimes diverge. On the other hand, $C_r(t)$ is fixed in the proposed method, which is given from the temporally-smoothed image by TR in this study. The fixed shape of the manifold would ensure the convergence of the proposed method. The constraint about the target/background regions in images is also newly incorporated in the problem of the compartment-model-based 4D PET image reconstruction.

Chapter 4

Blind-Deconvolution of 3D Pathological Images

In this chapter, we deal with the blind-deconvolution of the pathological images. Pathological images describe very micro anatomical structures such as cell cores, capillaries, and bile-ducts. Because of the quite delicate imaging process, the pathological images often have several kinds of the degradation of image qualities.

The blur, which is one of the typical degradations, is assumed to be caused not only by some kinds of filtering derived from the camera but also by the outlayers of a slice sample [111, 112, 113]. For example, when an undulating or holding slice sample is captured, the surroundings of the undulating/holding pixels would be affected by the out-of-focus blur. If parts of a pathological image are blurred, redoing the imaging process would be higher cost compared to the other modalities because mending and checking of a slice sample are necessary in the much larger image scale.

In our study, the pathological images are assumed to be captured by the following procedure. An target organ sample is first removed from the body and is separated into thin slices. Then each thin slices is stained with some chemical substances such as H&E and Ki67. Finally, the stained slices are captured one by one using a microscopic camera. The resultant pathological images are composed of a series of continuous pathological sections. Some of the slice images have blurred parts in certain regions of them.

The objective of this study is to restore the deblurred pathological image from a blurred one. This task is a blind-deconvolution problem because not only the ideal image but also the blur kernel is unknown. In general, the observation model of the blind-deconvolution is formulated as

$$y = h * x + \epsilon, \tag{4.1}$$

where y, h, x, and ϵ are the observed image, the ideal deblurred image, the observation blur kernel, and the observation noise. In the general model, the blur kernel is assumed

to be uniform in the image space. We also assume that a blur in a pathological image is uniform in a certain region of a pathological image. The blind-deconvolution problem would be one of the most complicated ill-posed problems because both of the observation blur and the ideal image are unknown and there are not enough constraints for both the blur and the image. We should constrain the solution spaces based on appropriate priors of both the blur and the image. Moreover, the problem to be solved is not only ill-posed but also nonconvex. We also have to employ an appropriate initial values so as to achieve a good local optimum.

As for the case of our problem setting, we leverage the following priors to restore blurred images.

Priors on Images

Given a blurred part of a slice image, the same locations of the adjacent slice images are assumed to be not blurred since the blurred parts of pathological images would rarely exists. Thus the adjacent slice images, which are high-resolutional compared to the blurred target image, can be used for the prior. We leverage the structural similarity between the target blurred image and its adjacent high-resolutional ones. Since there is little differences between the captured structures in the target image and those in the adjacent images, the adjacent images would be able to be used as guide images for the deconvolution . In using the adjacent slice images as the reference images, there is a trouble that there exists location/shape gaps of anatomical structures captured in the adjacent images. To tackle this problem, we model the smooth deformations of anatomical structures by Large Deformation Diffeormorphic Metric Mapping (LDDMM). With the LDDMM, the smoothly deforming sub-slices between the upper and lower high-resolutional slices are well represented.

Furthermore, it can be said that the total number of gradients in the deblurred target image is almost equals to that of the adjacent high-resolutional ones. The accurate smoothness prior from this assumption is also introduced. The high-resolutional sub-slice images and the sparse representation of the image gradients are leveraged in the proposed method.

Priors on Blur Kernels

The blur to be estimated is assumed to be represented by the uniform kernel for the blurred part of a pathological image. Since pathological slice images are statically captured by microscopic camera, we also assume that the blur kernels would have the non-negative smooth shapes rather than high-frequencial ones such as those of motion blurs and sinc-like blurs.

Similarly to the most general blind-deconvolution methods, the above priors are mod-

eled in the form of the optimization problem that is solved by the Alternating Optimization (AO) framework. Unlike those approaches, however, we employ ADMM to optimize the subproblems for the sake of the optimization stability rather than the processing speed. We also define the proximal operator for the blur kernel prior in order to explicitly constraint and appropriately optimize the kernel estimation problem. Furthermore, coarse-to-fine estimation procedure is employed so as to avoid the initialization dependency of the nonconvex optimization as well as to achieve relatively good performances.

In the remainder of this chapter, we first mention the details of some existing blind-deconvolution methods including their solvers in Section 4.1.1. Next in Section 4.1.2, we describe the brief introduction of the LDDMM [114, 115], which is the generalized method for the image deformation technique and is implicitly employed in our approach. The proposed method is stated in Section 4.2. The experimental results using both of synthetic slice images and real slice images are presented in 4.3. Finally, the discussions related to the proposed method are mentioned in Section 4.3.3.

4.1 Basic Materials

4.1.1 Existing Methods for Blind-deconvolution

In this Section, the existing blind-deconvolution methods are reviewed. The general observation model Equation 4.1 is assumed to be solved in many existing methods. The problem is highly ill-posed and is much more complicated than some other ill-posed problems because both the image and the blur are unknown: in the nonblind super-resolution as described in Chapter 2, only the ideal image is unknown. It is necessary for the blind-deconvolution to constrain the solution space based on additional priors of both the blur and the image. In fact, there are few studies for the blind-deconvolution of histopathological images [116, 111] while there are several blur detection methods of them [112, 111]. In the following, therefore, we review the typical deblurring models including the other modalities. The existing methods that are based on the iterative estimation [28, 117, 27, 118, 119, 120, 121, 122] including few methods for pathological images [116] are generalized in the form of the following optimization problem:

$$\begin{aligned} \arg\min_{\boldsymbol{X},\boldsymbol{H}} \ & \lambda_O ||\boldsymbol{H} * \boldsymbol{X} - \boldsymbol{Y}||_F^2 + R_x(\boldsymbol{X}) + R_h(\boldsymbol{H}), \\ \text{s.t.} \ & \boldsymbol{H} \in S, \end{aligned} \tag{4.2}$$

where $\boldsymbol{H} \in \mathbb{R}^{H \times W}$ and $\boldsymbol{X} \in \mathbb{R}^{H \times W \times 3}$ are the blur and the image to be restored, and $\boldsymbol{Y} \in \mathbb{R}^{H \times W \times 3}$ is the observed image. $R_x(\boldsymbol{X})$, $R_h(\boldsymbol{H})$, and S are the regularization terms of $\boldsymbol{X}$ and $\boldsymbol{H}$, and the constraint subspace of $\boldsymbol{H}$. Note that Problem 4.2 is nonconvex while the general super-resolution tasks deal with convex optimization as described in

Chapter 2. Fundamentally, most of the existing methods impose the image smoothness prior by $R(\boldsymbol{X})$, and the smoothness and sparsity of a blur by $R(\boldsymbol{H})$. There are additional regularizations and constraints vary by the methods and their respective problem settings. In Reference [28], the text image deconvolution model is proposed, which uses sparsities of text images themselves and of their gradients. Reference [117] leveraged the dark channel prior, which is an effective sparse representation for natural images. Reference [116] introduced TV and the Gaussian-based semi-parametric representation of blur kernels for pathological images. In several methods which assume blurs derived from the motion, the optimization of $\boldsymbol{H}$ is solved in the gradient space [118, 119], where the sparsity of a blur kernel is implicitly assumed. The existing methods described above introduce appropriate priors on $\boldsymbol{X}$ and $\boldsymbol{H}$ for their own modalities.

The nonconvex optimization problem is commonly solved by the AO framework, where the subproblems of $\boldsymbol{X}$ and $\boldsymbol{H}$ are iteratively solved. The AO framework is consistent for the blind-deconvolution task because the subproblems of $\boldsymbol{X}$ and $\boldsymbol{H}$ are convex. Quite a lot methods, however, optimize the subproblems just by the conjugate gradient method [28, 117, 27, 118, 119]. That is, the constraint on $\boldsymbol{H}$ is ignored during the optimization process of $\boldsymbol{H}$ and is imposed after the optimization. On the other hand, it is also reported that the global optimization is stabled when the constraint is strictly satisfied during the optimization of the subproblems [121, 120]. The stability derived from strict constraints during the convex sub-problem optimization is also reported in the tensor factorization tasks [123, 124].

The difference between the model/solver of the proposed method and those of the some existing methods are mentioned in the discussion in Section 4.3.3.

4.1.2 LDDMM

In this section, we introduce the Large Deformation Diffeormorphic Metric Mapping (LDDMM) [114, 115]. The image registration techniques is one of the central research topic applied for medical images [125, 126]. The LDDMM is the generalized method of conventional image registration techniques such as B-spline based ones [127, 126].

The LDDMM between two images, $\boldsymbol{I}_0$ and $\boldsymbol{I}_1$ is calculated by

$$\arg\min_{v} \quad \left(\int_0^1 ||v_s||_V^2 ds + \frac{1}{\sigma^2} ||\phi_1^{-1} \oplus \boldsymbol{I}_0 - \boldsymbol{I}_1||^2 \right), \tag{4.3}$$

where $\phi_t(x) = \int_0^t v_s(x) ds + \phi_0$, and $\oplus$ is the pixel-wise deformation operator. In the LDDMM, the nonrigid deformation from $\boldsymbol{I}_0$ to $\boldsymbol{I}_1$ is represented by the deformation field ϕ_t, where the time parameter $t \in [0, 1]$ continuously changes from $t = 0$ to $t = 1$. Note that the data fidelity of the second term is evaluated based on image pixel values before/after the registration in Problem (4.3). The data fidelity is regularized by the integral of the

local smoothness metric of the deformation field. In this study, the V-norm of the first term is set by the L_2-norm of the image gradient space so as to impose the smoothness of the deformation field.

The LDDMM is used in the proposed method to model the sub-slice images between the high-resolutional adjacent slice images, as later mentioned in Section 4.2.

4.1.3 Problem Statement

In this study, we assume that the pathological images are composed of a series of thin sections with a constant inter-gap. The pathological images are assumed to be captured by the following procedure. An target organ sample is first removed from the body and is separated into thin sections. Then each thin section is stained with some chemical substances such as H&E and Ag. Finally, the stained slices are captured one by one using a microscopic camera.

The respective raw images are deformed when the thin slice samples are sectioned and mounted on glass slides by man-hand. These nonrigid deformations are caused slice-independently, and the adjacent slice images have big gaps of the locations/shapes of the anatomical structures. To tackle this issue, all the slice images are smoothly registered towards the through-slice direction based on [113, 128]. The registration is processed based on the smoothed structural landmarks and the resultant images have smooth structures running across the slices as shown in Figure 4.1. The resultant images of the landmark-based registration can be regarded as a 3D reconstructed pathological image. The detail of the 3D reconstruction method is described in Appendix. After the 3D reconstruction, there would be still small location-gaps between the adjacent slice images because the registration procedure do not interpolate the inter-slice images.

Given the pathological images above, the objective of this study is to restore the blurred parts of a slice image. Therefore in fact, we need to search the blurred parts from a slice image before deblurring. There are several methods to detect or evaluate the blurred parts of pathological images [112, 111]. Actually, however, the blurred parts of a slice image can be also detected during the registration procedure. In [113, 128], blurred parts are detected not to be referred as landmarks by using template matching. Thus in this study, we assume that the blurred parts of a pathological image are already given.

In the followings of this chapter, a series of smoothly registered N pathological images is denoted by $\boldsymbol{J}_1, \boldsymbol{J}_2, \cdots \boldsymbol{J}_M$. Some of the slice images have blurred parts, and the same parts of the adjacent slice images is assumed to be not blurred as shown in Figure 4.2.

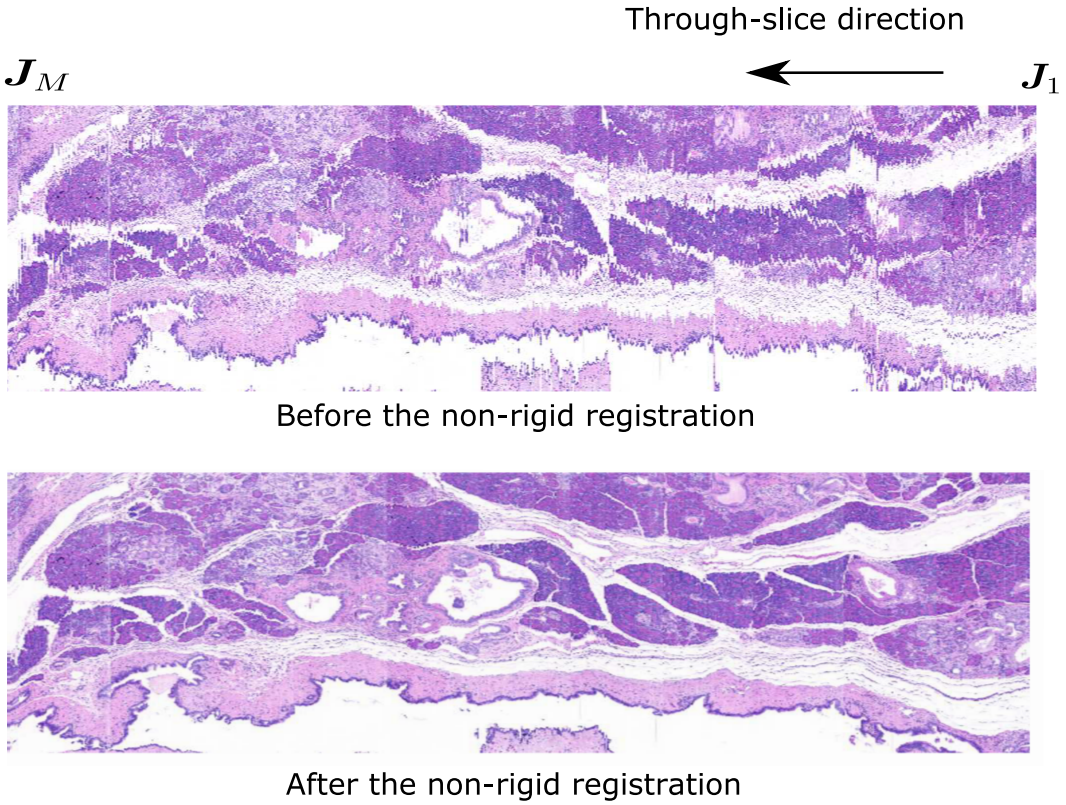


Figure 4.1 The pathological images before/after the landmark-based nonrigid registration [113, 128]. The illustrations are cross sections of 3D pathological images that are composed of piled up 2D slice images.

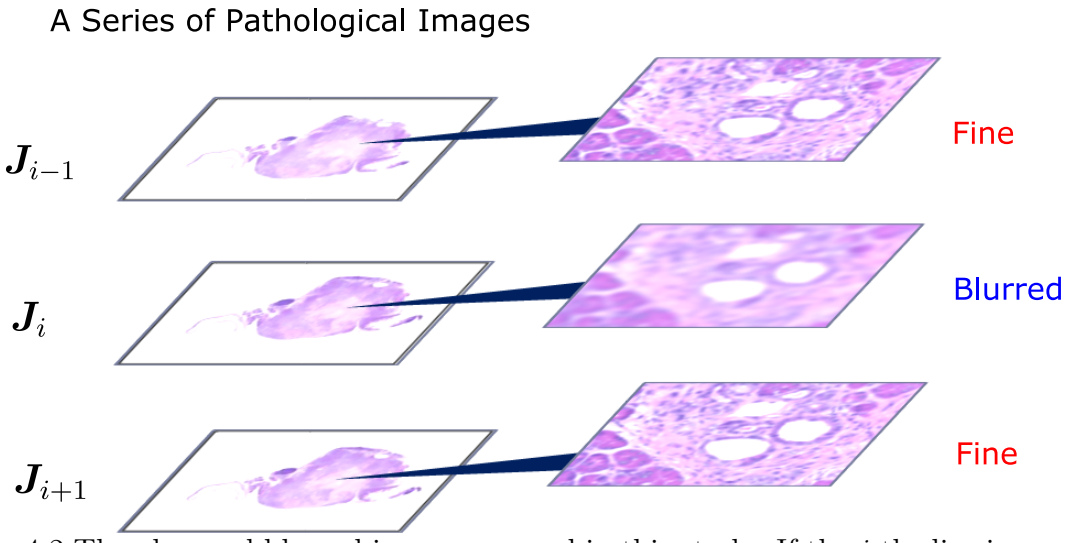


Figure 4.2 The observed blurred image assumed in this study. If the i-th slice image have a blurred part, the same location parts of the adjacent slice images are not blurred.

4.2 Proposed method

It is assumed that the target blurred image $\boldsymbol{I}_b$ and the adjacent high-resolutional images $\boldsymbol{I}_0$ and $\boldsymbol{I}_1$ are given, and they appear in the through-slice order of $\boldsymbol{I}_0$, $\boldsymbol{I}_b$ and $\boldsymbol{I}_1$. $\boldsymbol{I}_0$, $\boldsymbol{I}_b$ and $\boldsymbol{I}_1$ are the parts of continuous three slice images $\boldsymbol{J}_{i-1}, \boldsymbol{J}_i, \boldsymbol{J}_{i+1}$. With these notations, the objective of the proposed method is to deblur $\boldsymbol{I}_b$. We leverage the adjacent images $\boldsymbol{I}_0$ and $\boldsymbol{I}_1$ that have high-resolutional textures for the blind-deconvolution of $\boldsymbol{I}_b$. Note that $\boldsymbol{I}_0$ and $\boldsymbol{I}_1$ cannot directly used for the representation of high-resolutional textures of $\boldsymbol{I}_b$ because there are location/shape gaps between them even after the smooth 3D reconstruction described in Section 4.1.3. Therefore, we model the sub-slices description between high-resolutional $\boldsymbol{I}_0$ and $\boldsymbol{I}_1$ that smoothly deform from $\boldsymbol{I}_0$ and $\boldsymbol{I}_1$, and vice versa, by using the LDDMM.

In the proposed method, first of all, the LDDMM is calculated between $\boldsymbol{I}_0$ and $\boldsymbol{I}_1$ so as to achieve high-resolutional sub-slice images $\boldsymbol{R}_0^t$ and $\boldsymbol{R}_1^t$. Thus $\boldsymbol{R}_0^t = \phi_{0\to1}(t) \oplus \boldsymbol{I}_0$ and $\boldsymbol{R}_1^t = \phi_{1\to0}(t) \oplus \boldsymbol{I}_1$ are calculated, where $\phi_{0\to1}(t)$ is the vector field to deform from $\boldsymbol{I}_0$ to $\boldsymbol{I}_1$ at the time $t \in [0, 1]$. Note that, $\phi_{1\to0}(t)$ can easily calculated by inversely tracing the vector field $\phi_{0\to1}(t)$ from $\boldsymbol{I}_1$ to $\boldsymbol{I}_0$. In the resultant $\boldsymbol{R}_0^t$, the pixel coordinates are smoothly deformed from $\boldsymbol{I}_0$ to $\boldsymbol{I}_1$ in terms of $t \in [0, 1]$. $\boldsymbol{R}_j^t$ is the smoothly deforming high-resolutional slice image, which would be a good representation of sub-slices between $\boldsymbol{I}_0$ and $\boldsymbol{I}_1$. We leverage the data fidelity derived from the sub-slice description $\boldsymbol{R}_j^t$ for the blind-deconvolution. Moreover, we introduce the L_0 gradient prior derived from $\boldsymbol{R}_j^0$ because the proposed method can refer to the ideal number of image gradients from the adjacent images.

In the followings, let the image and the blur to be restored are denoted by $\boldsymbol{X}$ and $\boldsymbol{H}$, and the target blurred image by $\boldsymbol{Y} = \boldsymbol{I}_b$. In order to obtain an ideal deblurred image, the following optimization problem is solved:

$$\begin{aligned} \arg\min_{\boldsymbol{X},\boldsymbol{H},\boldsymbol{R}_j^t} \quad & \lambda_O ||\boldsymbol{H} * \boldsymbol{X} - \boldsymbol{Y}||_F^2 + \lambda_R ||\boldsymbol{H} * \boldsymbol{R}_j^t - \boldsymbol{Y}||_F^2 + \lambda_H ||\boldsymbol{H}||_F^2, \\ \text{s.t.} \quad & ||\nabla \boldsymbol{X}||_0 \leq ||\nabla \boldsymbol{R}_j^0||_0, \\ & \boldsymbol{H} \in S_p, \end{aligned} \tag{4.4}$$

which is a nonconvex blind-deconvolution problem. The first constraint imposes the sparsity of the image gradient based on the adjacent high-resolutional image. The second constraint imposes the sparsity of a kernel and the nonnegativety to $\boldsymbol{H}$. Namely, S_p denotes the convex subspace where at most $d \times d$ ($d << W, H$) nonzero components of $\boldsymbol{H}$ reside in the very narrow space of $W \times H$ domain, and the other components are entirely zero. λ_0, λ_R and λ_H are the parameters to be tuned by the input image. The data fidelity is evaluated only in the image domains although many existing methods, which deal with motion blur as well as the other smooth blurs, evaluate the data fidelity in the gradient

domain.

In the proposed method, Problem (4.4) is solved by the Alternating Optimization ADMM (AO-ADMM) [123, 124]. By using the AO-ADMM, we can efficiently solve Problem (4.4) by separating a nonconvex matrices-factorization problem into convex sub-problems. Problem (4.4) is separated into three convex sub-problems, each of which is respectively dependent on one of the three variables $\boldsymbol{X}, \boldsymbol{H}$, and $\boldsymbol{R}_j^t$. In each sub-problem, there is a single variable of $\boldsymbol{X}$, $\boldsymbol{H}$, and $\boldsymbol{R}_j^t$ to be respected, and the other two variables are fixed. The three sub-problems are alternatively optimized until the convergence.

In alternatively solving Prmblem (4.4), λ_R is reduced per alternating iteration because the second term is not necessarily satisfied with the ideal $\boldsymbol{H}$ and should behave as the weak regularization. The reduction of λ_R is performed by $\lambda_R \leftarrow \eta\lambda_R$, where $\eta \in (0, 1)$ is the reduction parameter. We also empirically found that reducing the parameter η per iterations exponentially makes both the performance and the stability better.

In the followings, let $\boldsymbol{x} = \mathrm{vec}(\boldsymbol{X})$ and $\boldsymbol{h} = \mathrm{vec}(\boldsymbol{H})$ to describe the optimization sub-problems.

The optimization of the image, $\boldsymbol{X}$

The convex sub-problem of Problem (4.4) with respect to $\boldsymbol{X}$ can be formulated as

$$\begin{aligned} \arg\min_{\boldsymbol{x}} \quad & \lambda_O ||\boldsymbol{H}_{\mathrm{cnv}}\boldsymbol{x} - \boldsymbol{y}||_F^2, \\ \text{s.t.} \quad & ||\nabla \boldsymbol{x}||_0 \leq \alpha, \end{aligned} \tag{4.5}$$

where $\boldsymbol{H}_{\mathrm{cnv}}$ denotes the cyclic convolutional matrix of $\boldsymbol{h}$, and $\alpha = ||\nabla \boldsymbol{R}_j^0||_0$.

In order to solve Problem (4.5) with the ADMM, it is reformulated to hold proximity as

$$\begin{aligned} \arg\min_{\boldsymbol{x}} \quad & \lambda_O ||\boldsymbol{H}_{\mathrm{cnv}}\boldsymbol{x} - \boldsymbol{y}||_F^2 + i_G(\boldsymbol{u}), \\ \text{s.t.} \quad & \boldsymbol{u} = \boldsymbol{D}\boldsymbol{x}, \end{aligned} \tag{4.6}$$

where $i_G(\boldsymbol{u})$ is the indicator function such as

$$i_G(\boldsymbol{u}) = \begin{cases} 0 & \text{if } ||\boldsymbol{u}||_0 \leq \alpha \\ \infty & \text{otherwise.} \end{cases} \tag{4.7}$$

In [29], it is shown in the ADMM framework that the L_0-norm constraint can be approximately solved by choosing top α components in terms of their L_2-norms, and setting the remaining $K - \alpha$ components to zero. The procedure can be formulated in the format of the proximal operator of $i_G(\boldsymbol{u})$ as

$$\mathrm{prox}_{i_G}(\boldsymbol{u}) = [\tilde{\boldsymbol{u}}_1^{\mathrm{T}}, \tilde{\boldsymbol{u}}_2^{\mathrm{T}}, \cdots, \tilde{\boldsymbol{u}}_{WH}^{\mathrm{T}}]^{\mathrm{T}}, \tag{4.8}$$

where

$$\tilde{\boldsymbol{u}}_k = \begin{cases} \boldsymbol{u}_k & \text{if } k \in \{(1), \cdots, (\alpha)\} \\ 0 & \text{if } k \in \{(\alpha+1), \cdots, (WH)\}, \end{cases} \tag{4.9}$$

where (k) is the projected index of the pixel coordinate index k. The index (k) is obtained by sorting subvectors $\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_{WH}$ in descending order in terms of their L_2-norms as $\boldsymbol{u}_{(1)}, \boldsymbol{u}_{(2)}, \cdots, \boldsymbol{u}_{(WH)}$, where $||\boldsymbol{u}_{(1)}|| \geq ||\boldsymbol{u}_{(2)}|| \geq \cdots \geq ||\boldsymbol{u}_{(WH)}||$. In fact, we should use mixture L_{10}-norm instead of L_0-norm [29] in order to apply the proximal operator to color images. The extension from L_0-gradients to L_{10}-gradients can be straightforwardly done by summing up the pixel-wise gradients in terms of the color channel.

With the above formulations, Problem (4.5) is solved by the following updating rules of the ADMM:

$$\boldsymbol{x}^{k+1} = [\lambda_O \boldsymbol{H}_{\text{cnv}}^{\mathrm{T}} \boldsymbol{H}_{\text{cnv}} + \rho \boldsymbol{D}^{\mathrm{T}} \boldsymbol{D}]^{-1} (\lambda_O \boldsymbol{H}_{\text{cnv}}^{\mathrm{T}} \boldsymbol{y} - \boldsymbol{D}^{\mathrm{T}} \boldsymbol{\beta}^k + \rho \boldsymbol{D}^{\mathrm{T}} \boldsymbol{u}^k), \tag{4.10}$$

$$\boldsymbol{u}^{k+1} = \text{prox}_{i_G}(\boldsymbol{D}\boldsymbol{x}^{k+1} + \rho^{-1}\boldsymbol{\beta}^k), \tag{4.11}$$

$$\boldsymbol{\beta}^{k+1} = \boldsymbol{\beta}^k + \rho(\boldsymbol{D}\boldsymbol{x}^{k+1} - \boldsymbol{u}^{k+1}), \tag{4.12}$$

where $\boldsymbol{u}$ is the dual variable of the image gradients, and $\boldsymbol{\beta}$ is the Lagrange multiplier.

The optimization of the blur, $\boldsymbol{H}$

The convex sub-problem of Problem (4.4) with respect to $\boldsymbol{H}$ can be formulated as

$$\begin{aligned} \arg\min_{\boldsymbol{h}} \quad & \lambda_O ||\boldsymbol{X}_{\text{cnv}} \boldsymbol{h} - \boldsymbol{y}||_F^2 + \lambda_R ||\boldsymbol{R}_{\text{cnv}} \boldsymbol{h} - \boldsymbol{y}||_F^2 + \lambda_H ||\boldsymbol{h}||_F^2, \\ \text{s.t.} \quad & \boldsymbol{h} \in S_p, \end{aligned} \tag{4.13}$$

where $\boldsymbol{X}_{\text{cnv}}$ and $\boldsymbol{R}_{\text{cnv}}$ denote the cyclic convolutional matrices of $\boldsymbol{x}$ and $\boldsymbol{r}_j^t$.

In order to solve Problem (4.13) with the ADMM, it is reformulated to hold proximity as

$$\begin{aligned} \arg\min_{\boldsymbol{h}} \quad & \lambda_O ||\boldsymbol{X}_{\text{cnv}} \boldsymbol{h} - \boldsymbol{y}||_F^2 + \lambda_R ||\boldsymbol{R}_{\text{cnv}} \boldsymbol{h} - \boldsymbol{y}||_F^2 + \lambda_H ||\boldsymbol{h}||_F^2 + i_{S_p}(\boldsymbol{v}), \\ \text{s.t.} \quad & \boldsymbol{v} = \boldsymbol{h}, \end{aligned} \tag{4.14}$$

where $i_S(\boldsymbol{v})$ is the indicator function such as

$$i_{S_p}(\boldsymbol{v}) = \begin{cases} 0 & \text{if } \boldsymbol{v} \in S_p \\ \infty & \text{otherwise.} \end{cases} \tag{4.15}$$

For reference, the proximal operator to impose nonnegativity on $\boldsymbol{v}$ is given by

$$\text{prox}_{i_{\cdot \geq 0}}(\boldsymbol{v}) = \max(\boldsymbol{v}, 0), \tag{4.16}$$

which is well-known. Similarly to this, the proximal operator of $i_S(\boldsymbol{v})$ can be defined as

$$\mathrm{prox}_{i_{S_p}}(\boldsymbol{v}) = \frac{1}{\mathcal{Z}}[\max(\boldsymbol{v}, 0) \circ \boldsymbol{m}], \tag{4.17}$$

where $\boldsymbol{m} \in \{0,1\}^{WH}$ is the indicator vector which indicates the narrow domain of $\boldsymbol{v}$ where the blur kernel should reside: only $d \times d$ components of $\boldsymbol{m}$ have the values 1 and the others have 0.

With the above formulations, Problem (4.13) is solved by the following updating rules of the ADMM:

$$\boldsymbol{h}^{k+1} = [\lambda_O \boldsymbol{X}_{\mathrm{cnv}}^{\mathrm{T}} \boldsymbol{X}_{\mathrm{cnv}} + \lambda_R \boldsymbol{R}_{\mathrm{cnv}}^{\mathrm{T}} \boldsymbol{R}_{\mathrm{cnv}} + \lambda_H + \rho]^{-1} (\lambda_O \boldsymbol{X}_{\mathrm{cnv}}^{\mathrm{T}} \boldsymbol{y} + \lambda_R \boldsymbol{R}_{\mathrm{cnv}}^{\mathrm{T}} \boldsymbol{y} - \boldsymbol{\gamma}^k + \rho \boldsymbol{v}^k), \tag{4.18}$$

$$\boldsymbol{v}^{k+1} = \mathrm{prox}_{i_{S_p}}(\boldsymbol{h}^{k+1} + \rho^{-1} \boldsymbol{\gamma}^k), \tag{4.19}$$

$$\boldsymbol{\gamma}^{k+1} = \boldsymbol{\gamma}^k + \rho(\boldsymbol{h}^{k+1} - \boldsymbol{v}^{k+1}), \tag{4.20}$$

where $\boldsymbol{v}$ is the dual variable of the blur domain, and $\boldsymbol{\gamma}$ is the Lagrange multiplier.

The optimization of the reference image, $\boldsymbol{R}_j^t$

The sub-problem of Equation (4.4) with respect to $\boldsymbol{R}_j^t$ can be formulated as

$$\arg\min_{\boldsymbol{R}_j^t} \ ||\boldsymbol{H} * \boldsymbol{R}_j^t - \boldsymbol{Y}||_F^2. \tag{4.21}$$

In fact, Equation (4.21) would be not necessarily convex because the solution space of $\boldsymbol{R}_j^t$ is restricted by the deformation while the global optimum can be obtained. Since $\boldsymbol{R}_j^t$ is given by LDDMM in advance, the optimization can be calculated by searching $\boldsymbol{R}_j^t$ that minimizes the data term, which costs at most $O(2T)$ times evaluations of the objective function.

The initialization of $\boldsymbol{H}$

Because of its strong nonconvexity, the solution of Problem (4.4) is heavily dependent on the initial values. By the way, there are several methods proposed where the coarse-to-fine estimation schemes are employed for the blur estimation [118, 119]. We also employ the multi-resolutional scheme for the initialization of $\boldsymbol{H}$, which is first proposed in [118]. The initialization procedure is described in Figure 4.3. The principal reason is that the coarse-to-fine estimation of $\boldsymbol{H}$ would improve the performance as mentioned in the above references. By the coarse-to-fine estimation, the large structures and the detail structures in the image are separately and gradually respected during the estimation of the blur [118, 119]. There is another reason that the optimization would get more independent of the initial values $\boldsymbol{H}$ when the upsampling of the kernel starts from the Dirac's delta kernel at the appropriate coarse scale. From these reasons, the coarse-to-fine estimation can be regarded as a good initialization for solving Problem (4.4).

Finally, the proposed method is summarized in Algorithm 4.

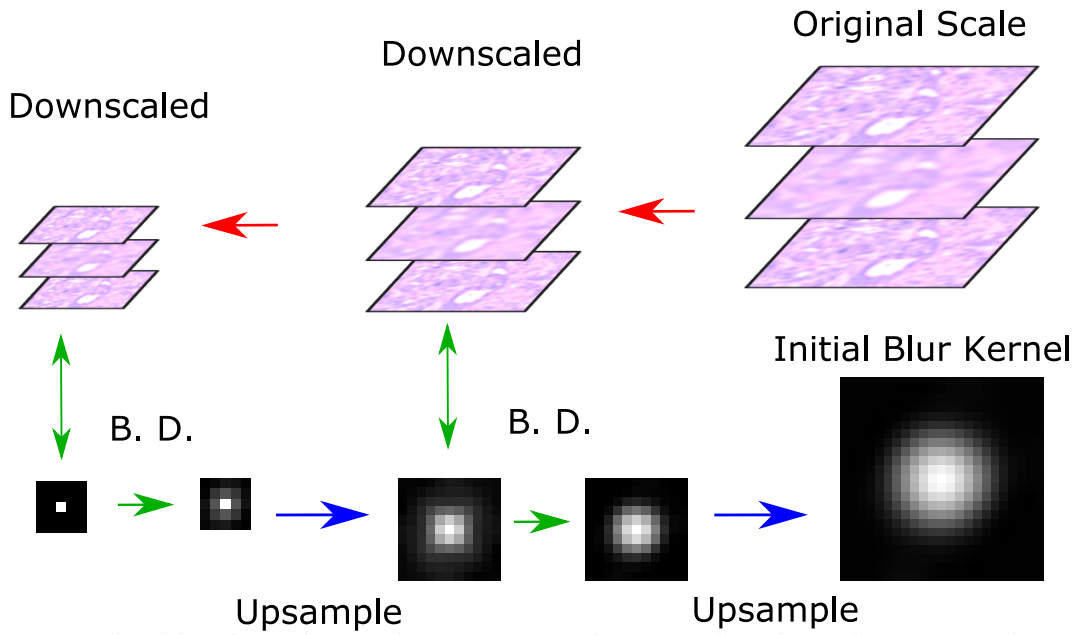


Figure 4.3 The blur kernel initialization procedure employed in the proposed method. First of all, the observed image and the referential sub-slice images are downsampled (red arrows). The 1/2-downscaling is continued until the sufficiently small scale which is enough for the blur to start the estimation from the delta function kernel. Then in the minimum image scale, the blind-deconvolution is processed (green arrows). The estimated blur kernel is upsampled so as to use it as the initial value in the next scale (blue arrows). The procedure is continued until the blur kernel size get to that of the original scale.

4.3 Results and Discussion

In our experiments, the target series of pathological images was of the pancreas of KPC-mouse [129], which were stained with H&E. There were 810 slice images and some of them have blurred parts. The image size was roughly $70K \times 100K$ and the spatial resolution set by the imaging process was $0.22\mu\text{m} \times 0.22\mu\text{m}$. Those pathological images were prepared and processed as mentioned in Section 4.1.3.

In Section 4.3.1, we first show the simulational experiments using synthetic blurs and real pathological images. Next in Section 4.3.2, the experimental results using the real blurred pathological images are shown. In the following experiments, the proposed model is denoted by L0GP+SSM, and the proposed model with $\lambda_H = 0$ is denoted by L0GP. The model of relaxed L_0 gradient regularization [27] is denoted by L0GR. The proposed model L0GP+SSM uses both the L_0 gradient constraint and the high-resolutional sub-slice images derived from adjacent slice images for the image priors. L0GP uses the L_0 gradient constraint, and L0GR [27] uses the L_0 gradient regularization. All the models have the same blur priors mentioned in Section 4.2.

For each model, the optimization problems were optimized using the AO-ADMM and

Algorithm 4 Proposed algorithm

input: The target blurred image $\boldsymbol{I}_b$,
and the adjacent high-resolutional images $\boldsymbol{I}_0$ and $\boldsymbol{I}_1$.
$\boldsymbol{R}_0^t = \phi_{0\to1}(t) \oplus \boldsymbol{I}_0$, $\boldsymbol{R}_1^t = \phi_{1\to0}(t) \oplus \boldsymbol{I}_1$, and $\boldsymbol{X} = \boldsymbol{I}_b$;
repeat
  Search $\boldsymbol{R}_j^t$ that minimizes (4.21);
  repeat
    Update $\boldsymbol{x}$ based on (4.10);
    Update $\boldsymbol{u}$ based on (4.11);
    Update $\boldsymbol{\beta}$ based on (4.12);
  until The optimization cost of (4.5) converges
  repeat
    Update $\boldsymbol{h}$ based on (4.18);
    Update $\boldsymbol{v}$ based on (4.19);
    Update $\boldsymbol{\beta}$ based on (4.20);
  until The optimization cost of (4.13) converges
  $\lambda_R \leftarrow \eta\lambda_R$;
  $\eta \leftarrow \eta^2$;
until The cost function converges and the constraints are satisfied
output: $\boldsymbol{X} = \text{fold}(\boldsymbol{x})$ and $\boldsymbol{H} = \text{fold}(\boldsymbol{h})$

our initialization procedure. The minimum kernel size for the coarse-to-fine initialization was 5×5, and was upsampled to 11×11 and 23×23, which was at the original scale. The parameters were manually tuned to perform the best at PSNR scores. The image domain sizes of $\boldsymbol{I}_0$, $\boldsymbol{I}_1$, and $\boldsymbol{I}_b$ were set to 200×200 before the LDDMM, and 180×180 before the blind-deconvolution so as to remove the irrelevant stretched boundaries caused by LDDMM.

4.3.1 Simulational Experiments

In the simulational experiments, given $\boldsymbol{I}_0$, $\boldsymbol{I}_1$, and $\boldsymbol{I}_b$ which are high-resolutional and not blurred, the synthetic blur kernels were convolved to $\boldsymbol{I}_b$. The blurred image was deblurred with several deconvolution models. The performance was evaluated both in the image domain and in the blur kernel domain. In the image domain, PSNR and SSIM [130] were employed for the evaluation. As for the blur kernel domain, the Blur Kernel Similarity (BKS) [131] was employed, which is given by the maximum response of the normalized cross-correlation calculated between the groundtruth and the estimated kernel.

Figure 4.4 − 4.7 show the illustrations of restored images and blurs (Figure 4.4 and 4.6

), and the corresponding evaluation results with respect to the AO iteration (Figure 4.5 and Figure 4.7). L0GR performed next to L0GP while not only did the cost functions form zig-zag shapes but also they continued to increase and totally not stabled. Although the objective function of L0GP converged both stably and quickly, the performance continued to improve and resulted in the slow convergence. This is because the constraint is not satisfied even if the objective function converges, and the performance of L0GP model is highly dependent on the constraints rather than the objective function. L0GP+SSM outperformed the other models. It converged much faster than the others and the performance also converged at almost the same time. In L0GP+SSM, the regularization data term of sub-slice images drastically accelerated the convergence as well as improved the accuracy performance. The result images of L0GR had distinct edges although quite a few detail structures were over-smoothed. The result image of L0GP had relatively blurred edges compared to those of L0GR while detail structures were retained. L0GP+SSM was able to restore the several detail structures and relatively deblurred image. It is also interesting that the performance was greatly improved with L0GP+SSM while the optimal $\boldsymbol{R}_j^t$ looks similar to the groundtruth but has several structural differences.

Furthermore, the results of other combinations of simulated blurs and images are shown in Figure 4.8, 4.9, and Table 4.1 − 4.3. Those results of L0GR were obtained by stopping the AO iterations at 1000 times because the cost function of L0GR was not stabled by the image as shown in Figure 4.4 − 4.7. L0GP and L0GR occasionally performed better than L0GP+SSM in terms of SSIM. This is because the oversmoothed images based on L_0-gradient could be a good representation of exactly smooth pathological structures in those cases. While the results of L0GP+SSM were sometimes contaminated with artifacts in the case of the blur kernel is large, it outperformed the other models on the whole in terms of PSNR and BKS.

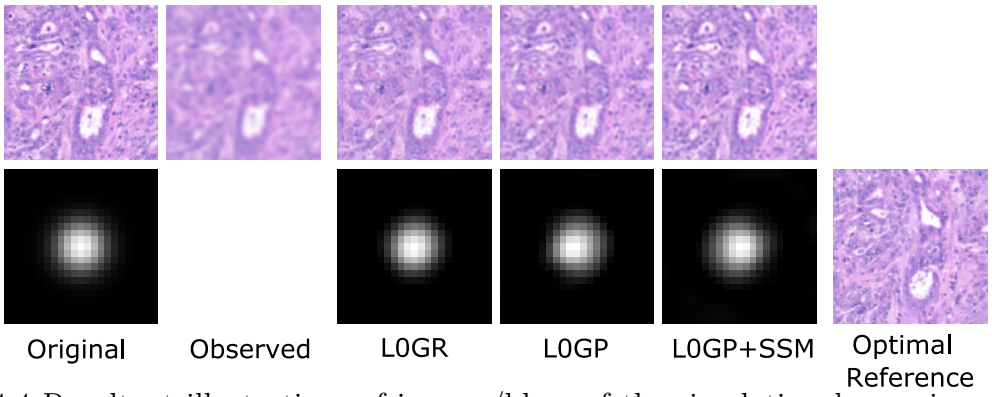


Figure 4.4 Resultant illustrations of images/blurs of the simulational experiments. The groundtruth blur is the Gaussian with $\sigma = 2.5$. The results correspond to those of Figure 4.5.

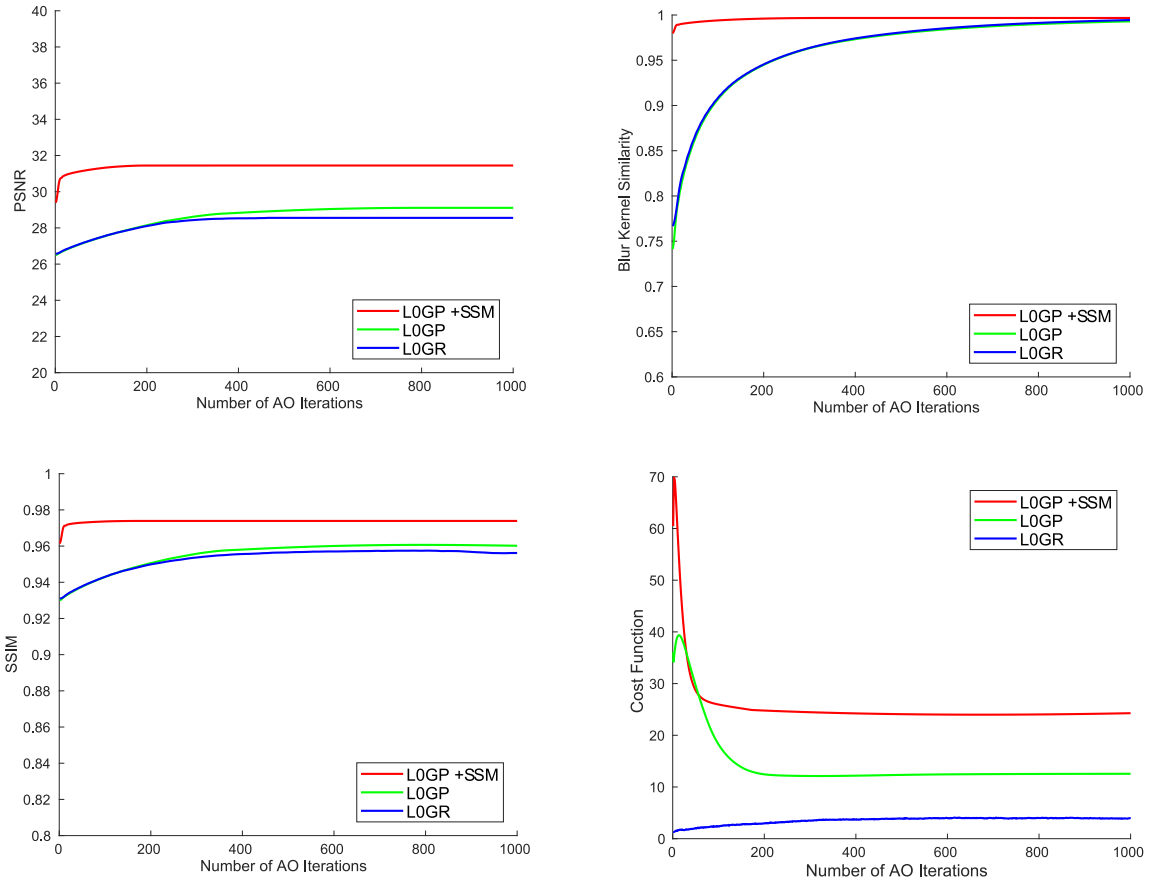


Figure 4.5 Performances of the deconvolution models with respect to AO iterations. The results correspond to those of Figure 4.4.

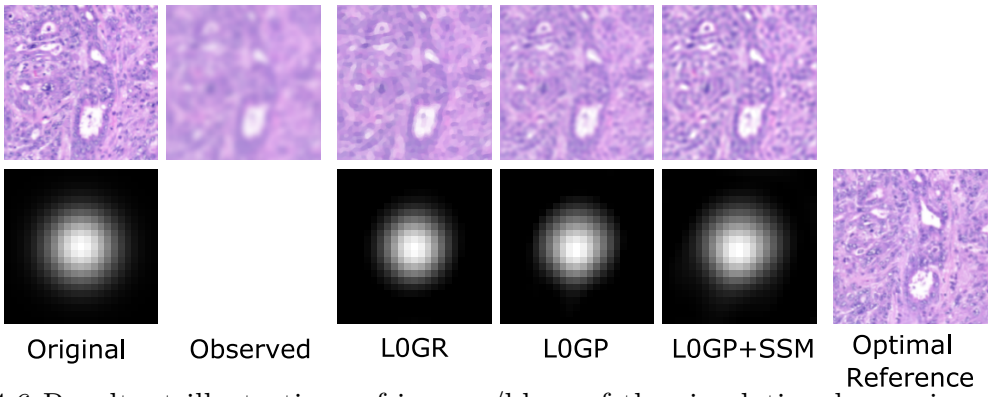


Figure 4.6 Resultant illustrations of images/blurs of the simulational experiments. The groundtruth blur is the Gaussian with $\sigma = 3.5$. The results correspond to those of Figure 4.7.

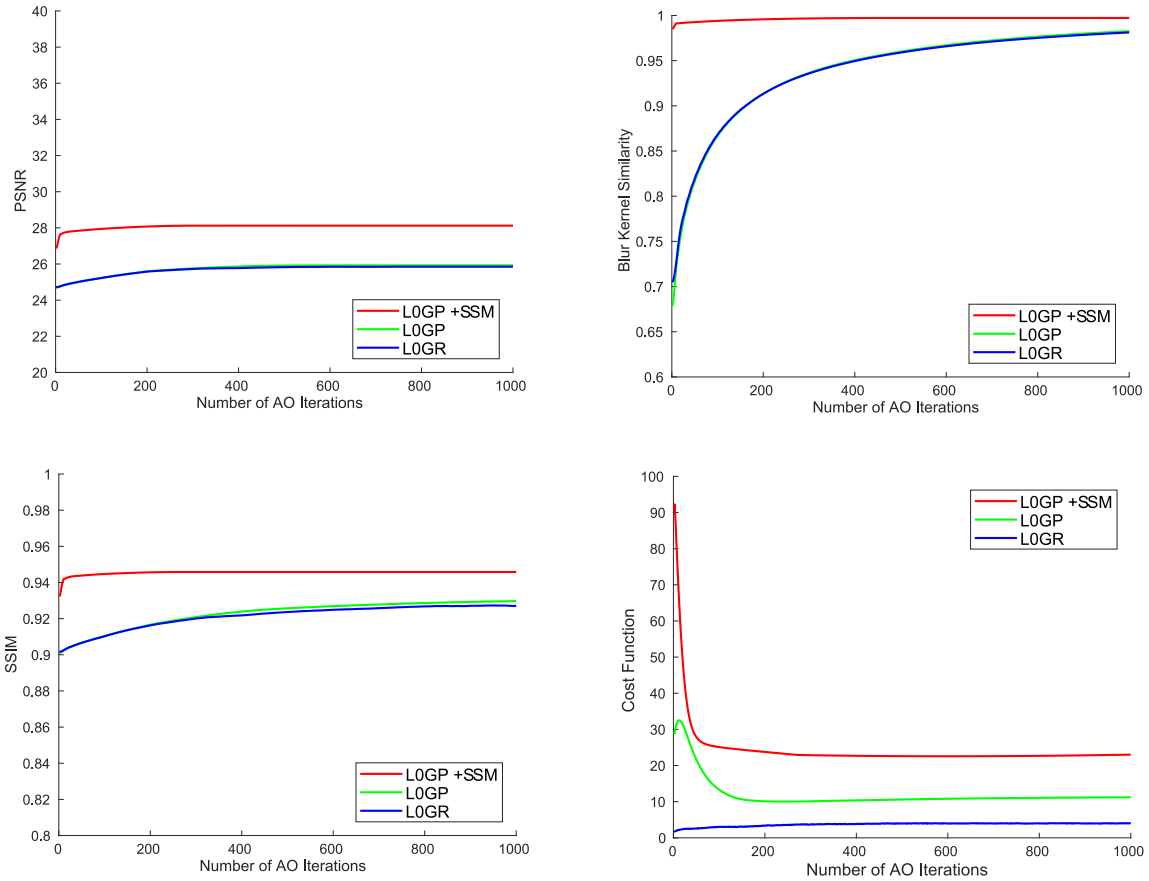


Figure 4.7 Performances of the deconvolution models with respect to AO iterations. The results correspond to those of Figure 4.6.

Table 4.1 PSNR results corresponding to images in Figure 4.8 and 4.9.

Data ID	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)
L0GP+SSM	**37.91**	**35.60**	**29.28**	**25.22**	**36.07**	**29.29**	**28.52**	**24.97**
L0GP	36.35	33.78	27.38	24.39	35.53	29.18	27.71	24.79
L0GR	33.36	30.68	27.91	24.82	30.38	26.69	28.20	24.52

Table 4.2 SSIM results corresponding to images in Figure 4.8 and 4.9.

Data ID	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)
L0GP+SSM	**0.994**	**0.991**	**0.959**	0.898	0.986	**0.944**	0.946	0.876
L0GP	0.992	0.986	0.943	0.896	**0.988**	0.939	0.939	**0.885**
L0GR	0.984	0.972	0.948	**0.901**	0.956	0.905	**0.947**	0.873

Table 4.3 BKS results corresponding to images in Figure 4.8 and 4.9.

Data ID	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)
L0GP+SSM	**0.997**	**0.998**	**0.998**	**0.998**	**0.998**	**0.998**	**0.996**	**0.993**
L0GP	0.995	0.995	0.992	0.983	**0.998**	**0.998**	0.935	0.983
L0GR	**0.997**	**0.998**	0.994	0.981	0.988	0.996	0.994	0.969

4.3.2 Real Data Experiments

Next, the results of the proposed method applied to the real blurred images are shown. Figure 4.10 shows the illustrations of the restored images. We could also found that the detailed image textures and blurs are appropriately restored with the real blurs of clinical data. However, there were several artifacts, or outlayers newly found after the deconvolution in the restored images. We consider that the artifacts were derived not from the deconvolution procedure but from the imaging process because the uniform blur kernel cannot represent those peaky artifacts. We also consider that the blur was caused accompanied with those outlayes during the imaging process.

4.3.3 Discussion

Reviews on Priors

In the proposed method, the transformations of anatomical locations/shapes across the through-slice direction are well represented by the LDDMM. In addition to the differences in the structural locations and shapes between continuous three slice images, there actually exist slice-independent anatomical structures: Some of anatomical structures, which appear in the image $\boldsymbol{I}_0$, disappear in the adjacent images $\boldsymbol{I}_1$, and vice versa. The high-resolutional sub-slice description was able to be effectively leveraged by gradually

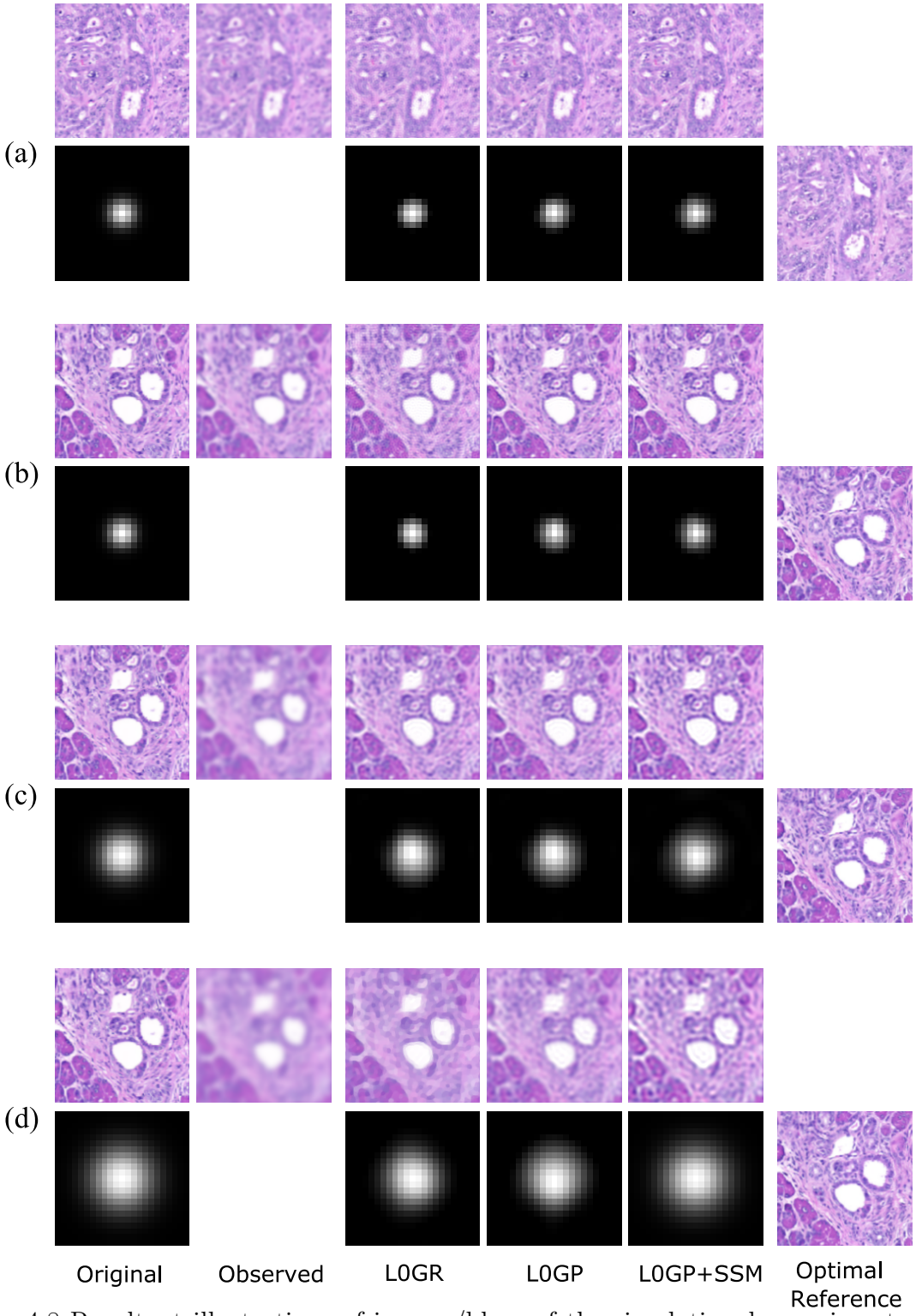


Figure 4.8 Resultant illustrations of images/blurs of the simulational experiments. The results correspond to those of (a) − (d) in Table 4.1 − 4.3.

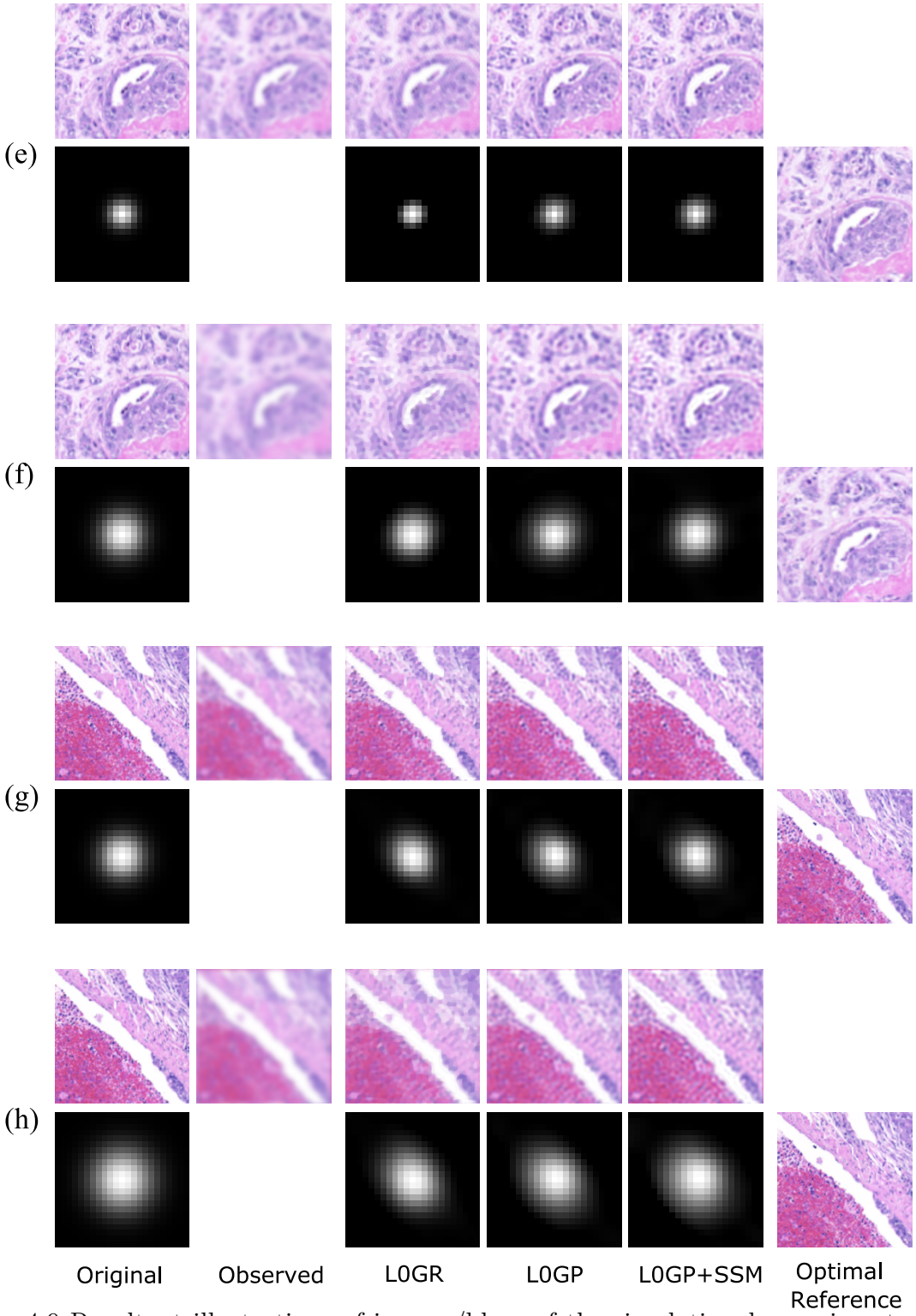


Figure 4.9 Resultant illustrations of images/blurs of the simulational experiments. The results correspond to those of (e) − (h) in Table 4.1 − 4.3.

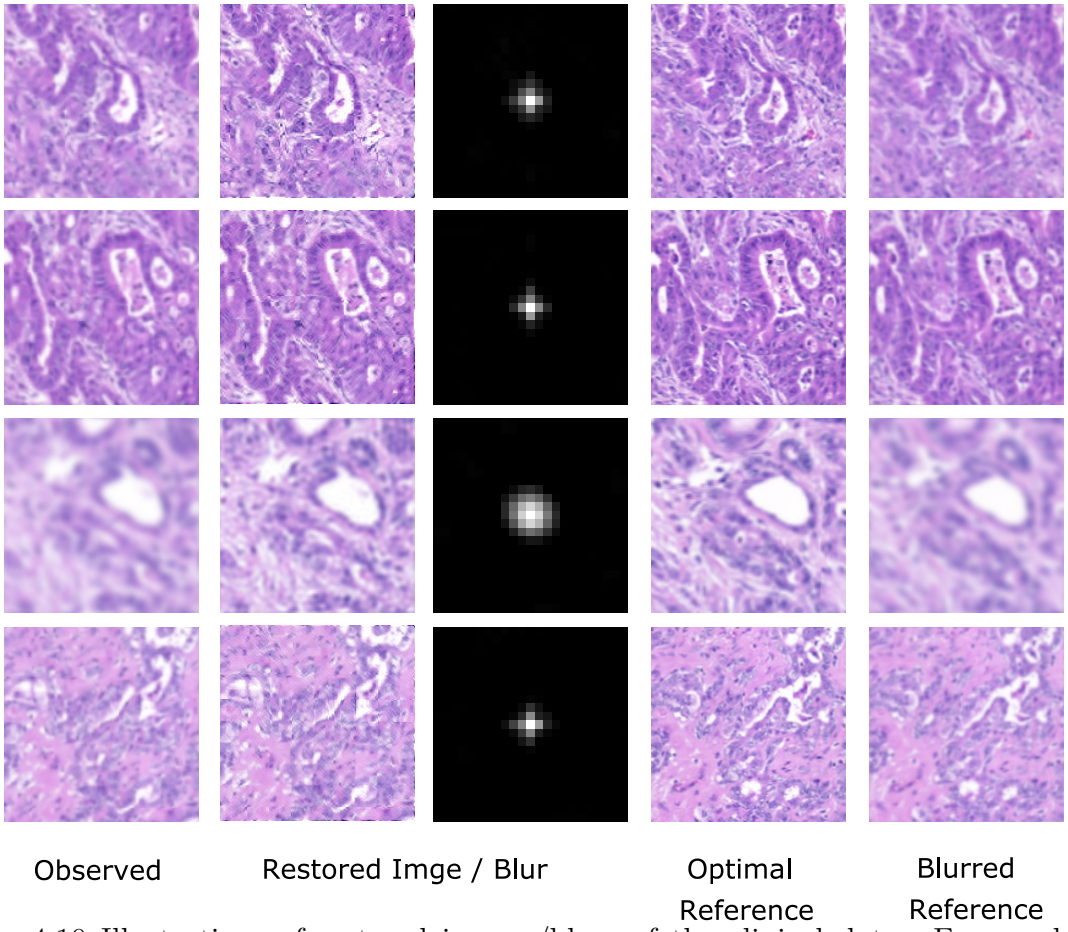


Figure 4.10 Illustrations of restored images/blurs of the clinical data. From left to right, each column denotes the observed image, the restored image, the restored blur, the optimal reference image: $\boldsymbol{R}_j^t$, and the blurred $\boldsymbol{R}_j^t$: $\boldsymbol{H} * \boldsymbol{R}_j^t$.

reducing the regularization weight λ_R even if there are structural differences between the target slice and the referred sub-slice image.

L0GP and L0GR would be essentially the same in their adopting L_0-norm components in terms of that the components with larger L_2-norm is adopted after the projection/regularization. Thus they perform similarly together under the ideal condition while L0GR needs highly complicated parameter-tuning for both of the stability and the accuracy performance.

In [116], they successfully introduced the smoothness (TV) prior for the blind-deconvolution of pathological images of muscles. In fact, however, the pathological images used in this study have much more complicated textures compared to theirs. Therefore the smoothness looks weak prior for the blind-deconvolution of our target pathological images. In spite of those properties of the target images, L0GP and L0GR achieved non-negligible deconvolution performances to some extent.

In our case, it would be also challenging to apply some priors used for natural images to the pathological images because there are many differences in image properties. There is a different property that blurs of common pathological images tend to have spread shapes such as the Gaussian and out-of-focus distributions while those of natural images tend to be motion blurs [117, 27, 118, 111, 116]. It must be also noticed that learning-based approaches should be carefully discussed on being employed especially for pathological images [111]. The usability of learning-based image reconstruction methods is often restricted because of the delicate clinical aspect. From these factors above, the appropriate priors used for the blind-deconvolution of the pathological images should be cautiously considered.

Sensitivity to initialization

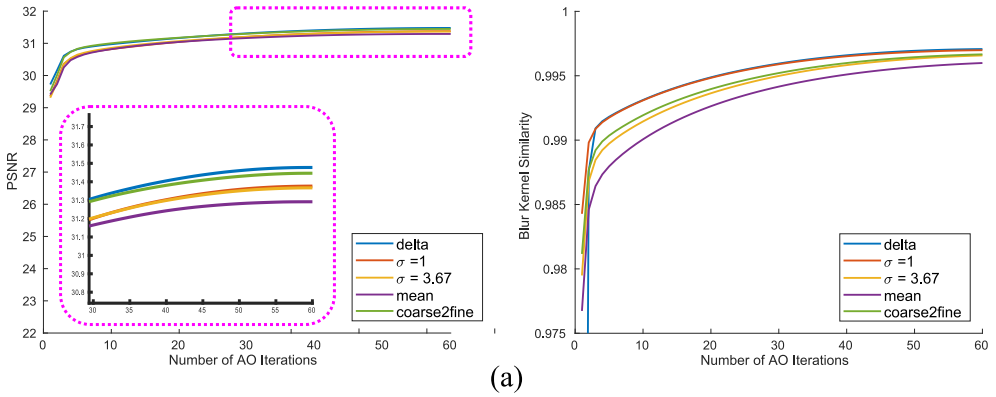
Next, we discuss the initialization of $\boldsymbol{H}$. We consider that the strict convergence at the coarse resolution makes the blur kernel overfit at the coarse kernel size, and it fall into a bad solution at the fine resolution. In order to achieve a good local optimum, we also found that the small number of the alternating optimization iterations at the coarse resolution leads to better results, as also mentioned in [118, 119]. The supplemental results in Figure 4.11 show the differences of the performances in terms of the initial value of $\boldsymbol{H}$. The number of iterations for the blur estimation in the coarser scale was set to 20. The initial value derived from coarse-to-fine procedure retained both of image and blur accuracy performances and their stabilities compared to the other initial values.

Limitation of the proposed method

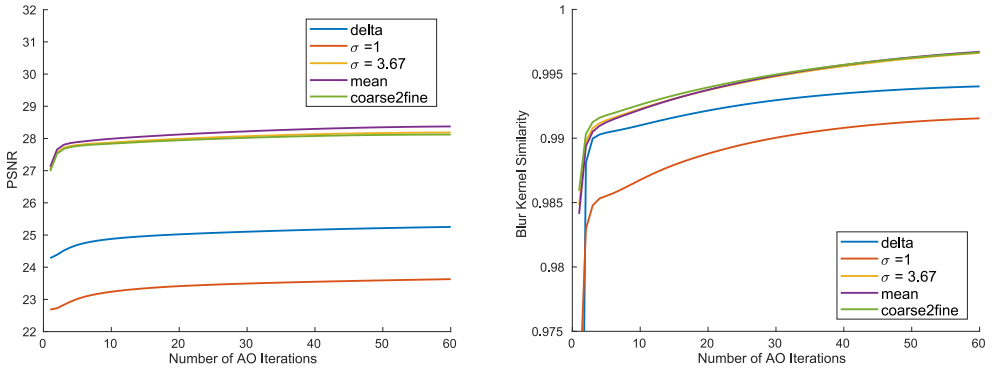
The proposed method has the following principal limitations. First of all, in order to model the high-resolutional sub-slice images, a pair of the over/under adjacent slice images of the target image is necessary in the proposed method. As mentioned in Section 4.2, roughly the same anatomical structures across the slices must be captured in those images while they can vary their shapes and locations by the slice. The landmark based registration of the slice images must be pre-processed with enough high resolution [113, 128] for the slice images to satisfy the above condition. There are also possible limitations accompanied with the blind-deconvolution model. Those limitations are discussed in the next paragraphs, comparing with the other existing methods.

Convergence stability

In many existing methods [117, 27, 118, 119], the resultant images are chosen by users during the iterations: The convergence and the best performance could not be satisfied at the same time. Compared to those existing methods, in the proposed method, the convergence can be stabled at next to the best performance. We consider that the stability is derived from the following framework for solving and the deconvolution model designed.



(a)



(b)

Figure 4.11 PSNR and BKS performances of L0GP+SSM model with variational initialization of $\boldsymbol{H}$. 'delta' is the delta function, and $\sigma = 1$ and $\sigma = 3.67$ are the Gaussian kernel with the standard deviation. 'mean' is the mean smoothing kernel, and 'coarse2fine' is the initialization described in Section 4.2.

- ADMM is employed for solving subproblems so that the variable can be exactly constrained during the optimization
- The regularization weight parameter is smoothly reduced per AO iteration because the regularization is unnecessary next to the groundtruth
- Independence of the initialization because of the coarse-to-fine estimation of $\boldsymbol{H}$ starting from the Dirac's delta kernel

Note that, however, the parameters still need to be tuned for the proposed method to perform the best image accuracy. Also, the processing time of the proposed method would be slower in exchange for the optimization stability.

Robustness against noise and artifact

In fact, the general blind-deconvolution model Prblem (4.2), that the most of all the state-of-the-art method belong to, does not assume the existence of observation noise while

the observation model Equation (4.1) includes the noise. Because both of $\boldsymbol{H}$ and $\boldsymbol{X}$ are unknown in the first term of Problem (4.2), one of the pair could try to make the other represent the noise, and vice versa. Not only the nonconvexity and the ill-posedness but also this factor makes the blind-deconvolution problems more complicated. Fortunately, pathological images are seldom contaminated with observation noise because they are statically imaged with appropriate amount of light by microscopic camera. Therefore, the affects of observation noise could be ignored in this study although the proposed method also belongs to the model Prblem (4.2).

As mentioned in Section 4.3.1, the resultant images of L0GP+SSM were sometimes contaminated with reconstruction artifacts in the case of the blur kernel is large. We expect that the artifacts could also be removed by simply enhance the threshold of the L_0 gradient constraint, i. e., by replacing the first constraint of Problem (4.4) with $||\nabla \boldsymbol{X}||_0 \leq \tau||\nabla \boldsymbol{R}_j^0||_0$, where τ is a parameter to enhance the threshold. In that case, L0GP would behave similarly to L0GR because the extent of the regularization is controlled through τ.

The computational complexity

In the followings, the total number of image pixel is denoted by $N = WH$. The single evaluation of the objective function of Problem (4.21) costs $O(N)$ and the optimization of Problem (4.21) costs $2T$ iterations at most, which would be less or equal to those of Problem (4.5) and (4.13). Therefore we consider Problem (4.5) and (4.13) to describe the computational complexity of the proposed method. The update equations (4.12), (4.19), and (4.20) cost $O(N)$ obviously. When Equation (4.10) and (4.10) are evaluated in the Fourier space, the convolutions and their inversions in the image domain can be substituted by the element-wise multiplications and divisions, which cost only $O(N)$. Thus Equation (4.10) and (4.10) cost $O(N\log\sqrt{N})$ for FFT and IFFT. The update equation Equation (4.11) costs $O(N\log N)$ for the quick sort of L_2-norm of the image gradients, which is the highest computational cost in the proposed algorithm.

Chapter 5

Conclusions

In this paper, we dealt with the image reconstructions and super-resolutions referred to some medical image modalities. These tasks can be formulated as solving the ill-posed inverse problems while the nature/extent of the ill-posedness is varied depending on the tasks and the modalities. In order to solve an ill-posed problem, additional prior models to constrain the solution space are necessary. For the respective tasks dealt with in this study, the appropriate and effective priors are introduced/modeled in the form of the optimization problems to be solved. We also introduced the more appropriate solver for each task especially in terms of the stability. In the followings of this chapter, first, the medical image inverse problems dealt with in this paper is reviewed and concluded. Then the possibility of application of other state-of-the-art priors to medical images, and possible futureworks are discussed.

Chapter 2

In this study, we dealt with the super-resolution of MR images. The objective of the super-resolution is to extrapolate the high-frequency components so as to improve the spatial resolution. The super-resolution of MR images is necessary to obtain both high-resolutional and high SNR profiles of MR images. In order to solve the super-resolution problem of MR images, the following priors were introduced:

1. The outside the object boundary of MR images have exactly zero values
2. The smoothness and the low-rankness of MR images

The Gerchberg algorithm was employed to introduce the prior (1), which is known as a strong constraint for the super-resolution. The Gerchberg algorithm, however, does not assume the existences of observation noise and ringing artifacts. Therefore in the proposed model, the Gerchberg algorithm was regularized by rank and TV in order to hold the robustness to noise/artifacts. There was a fundamental question how to introduce the regularization techniques to the Gerchberg algorithm, that is one of the POCS approaches.

In order to impose the regularization to Gerchberg algorithm, we first re-interpreted and reformulated the Gerchberg algorithm as a convex optimization problem. The proposed method was represented by the form of the convex optimization problem where the Gerchberg algorithm, TV and rank regularization terms are comprehensively described. The convex optimization problem was solved by ADMM, which is the relatively stable framework for the convex optimization. The proximal operators were also introduced to exactly minimize the TV and the nuclear norm. The proposed method was applied to MR images in order to obtain high resolution with a high SNR by using anisotropic measurements from the axial, coronal, or sagittal directions. The experimental results showed that our super-resolution technique dramatically reduced noise and ringing caused by the Gerchberg method and it also performed better than LRTV super-resolution and the other methods considered.

Chapter 3

In this study, we proposed a new PET reconstruction model where the spatial and temporal priors that PET images should obey are introduced. The PET images often suffer from terrible observation-noises and the accurate reconstruction techniques are desired. The PET image reconstruction is an ill-posed problem to reconstruct the PET image from an observed noisy sinogram. In the proposed method, following priors were introduced for the accurate PET image reconstruction:

1. The nonnegativity and the boundary constraints
2. The temporal changes of PET images obey the compartmental model

The prior (1) was explicitly introduced through the basis functions representations. The compartmental model of prior (2) is a temporal model of dynamics of medical ligand reactions, and the PET image should theoretically obey the compartmental model. In order to introduce the prior (2), first, the partial derivatives of the parameters of the compartmental model were derived. Further, the projection of an image to the nonconvex manifold covered by the parametric compartmental models was defined. The optimization problem was configured by introducing these prior models, where the data fidelity is evaluated based on KL-divergence. The nonconvex optimization problem was solved by iterative projections of spatial and temporal spaces. The experimental results showed that the effectiveness of each prior, and clinical data experiments also showed the usable performance of the proposed method.

Chapter 4

In this study, we dealt with the blind-deconvolution of a series of pathological images. The series of pathological images were captured from the continuous slice samples of a

pancreas of KPC mouse. Some of slice images included blurred parts that we do not know the blur profiles, and the objective is to deblur the blurred parts. Hence, the problem is a blind-deconvolution, which is not only highly ill-posed but also strongly nonconvex because both an image and a blur is unknown and we lack the constraint equations of both the image and the blur at the observation. The priors used for the pathological images should be carefully discussed while the blind-deconvolution itself is the severely complicated ill-posed problem. In our case, the adjacent slice images of a target blurred image can be referred to for the blind-deconvolution, which is assumed to be high-resolutional, and to be not blurred. The following prior models were introduced for both the image and the blur domain in the proposed method.

1. The smoothness of images described by the L_0-norm of the image gradients
2. The high-resolutional textures derived from the adjacent slice images
3. The smoothness and the general constraints of blurs

In introducing the prior (1), the L_0 gradient projection model was leveraged because the model can explicitly constraint the solution space by referring to the L_0 gradient norm of adjacent images. As for the prior (2), the LDDMM between the upper/lower adjacent images were first calculated in order to describe the high-resolutional sub-slice images between them. This procedure is necessary because there are location/shape gaps of anatomical structures between continuous three slice images. The sub-slice images were leveraged to evaluate the fidelity between the latent blur and the observed image in the form of a weak regularization. The prior (3) included the smoothness based on L_2-norm, and the convex constraints of the sparsity, the nonnegativity, and the normalized property. In solving the proposed blind-deconvolution model, the AO-ADMM framework, which alternatively solving the subproblems of respective variables using the ADMM, was employed to notably stable the optimization. We also employed the close-to-fine estimation procedure so as to avoid the dependence on the initial values for the nonconvex optimization. The proposed method outperformed drastically the existing methods in terms of both the stability and the accuracy while the usability of the sub-slice model is restricted to the continuous series of slice images.

Futureworks

For the futureworks of medical image inverse problems, we cannot ignore the deep learning. As mentioned in Chapter 1, the bias derived from the learned manifold should be under the consideration in order to apply the deep learning for several medical image modalities and processing tasks. Although the most of existing deep learning based methods do not argue the bias caused by learning, there are several possible approaches considered to correct or avoid the affects of the bias.

The deep image prior [76], which does not use learning, performed as good as the state-of-the art learning-based approaches such as SRResNet and LapSRN. It could be applied as the bias-free regularization to medical images if the appropriate architecture and optimizer that can well represent the modality is found. In applying it to medical images, some of modalities such as PET and f-MRI may have a trouble because they may have complicated and textures at the groundtruth. Because the deep image prior optimize the data fidelity starting from the noise representation, the fidelity distance would be much more faraway than those of smooth images. More architectural constraints would be necessary in that case.

The content loss, which is a data term used in several networks to make images more photorealistic [40], could be employed without learning if backpropagation and ADAM, or numerical differentiation are available. For instance, it could behave as the balanced regularization between the MSE-based data term and the photorealistic term. Note that, however, the tuning of the VGG network for the medical images would be necessary because it is tuned for natural images. It should be also discussed if the *photorealisticness* is desired in the medical images because the contents loss itself does not retain image colors, textures, and structural locations after the reconstruction.

The trained network can also be used as long as it behaves as just the regularization. The usage of SRCNN as the regularization term is reported for medical images [75]. As for the regularization usage of the trained network, one could apply the method with the technique to reduce the regularization weight per iteration, which is mentioned in Chapter 4, although the iterations are stopped by users in [75],

To the contrary, we could add explicit or obvious priors to deep neural networks. For example, we can indicate exactly that the backgrounds of MR images and, doctors can also get more detailed information such that the image object is lesion or not. These priors could communicate with and improve the deep neural networks by using conditional normalization techniques [132, 133]. Using the conditional layer normalization, it could be possible to keep the solution space close to the ideal manifold that is still not well represented by the training data.

These theoretical improvements would extend the flexibility of the model design. The model design that is supported by appropriate priors of the ideal image is still an important factor for the image inverse problems.

Appendices

ADMM optimization for solving Problem (2.12)

In this appendix, we explain the procedure to derive the update rules (2.14)–(2.17) for (2.12). Each update rule can be obtained by minimizing (2.12) with respect to each variable one by one.

First, (2.14) can be obtained by the following partial derivative:

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial \boldsymbol{f}} &= \boldsymbol{R}_\Omega^{\mathrm{T}}\boldsymbol{R}_\Omega\boldsymbol{f} - \boldsymbol{R}_\Omega\boldsymbol{f}_0^{'} - \boldsymbol{G}^{\mathrm{T}}\boldsymbol{\alpha} + \rho\boldsymbol{G}^{\mathrm{T}}\boldsymbol{G}\boldsymbol{f} - \rho\boldsymbol{G}^{\mathrm{T}}\boldsymbol{x} \\ &= \boldsymbol{R}_\Omega\boldsymbol{f} - \boldsymbol{f}_0^{'} - \boldsymbol{G}^{\mathrm{T}}\boldsymbol{\alpha} + \rho\boldsymbol{I}\boldsymbol{f} - \rho\boldsymbol{G}^{\mathrm{T}}\boldsymbol{x} \to 0. \end{aligned} \tag{5.1}$$

Thus, by solving (5.1) with respect to $\boldsymbol{f}$, we obtain

$$\begin{aligned}&(\boldsymbol{R}_\Omega + \rho\boldsymbol{I})\boldsymbol{f} = \boldsymbol{f}_0^{'} - \boldsymbol{G}^{\mathrm{T}}\boldsymbol{\alpha} + \rho\boldsymbol{G}^{\mathrm{T}}\boldsymbol{x}, \\ &\boldsymbol{f} = (\boldsymbol{R}_\Omega + \rho\boldsymbol{I})^{-1}[\boldsymbol{R}_\Omega\boldsymbol{f}_0^{'} + \boldsymbol{G}^{T}(\boldsymbol{\alpha} + \rho\boldsymbol{x})]. \end{aligned} \tag{5.2}$$

Similar to the case with $\boldsymbol{f}$, the update rule (2.15) for $\boldsymbol{x}$ can be obtained from

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial \boldsymbol{x}} \quad &= \epsilon\sum_{i=1}^{\mathcal{N}}(\boldsymbol{x} - \boldsymbol{m}_i + \boldsymbol{v}_i) - \boldsymbol{L}^{\mathrm{T}}\boldsymbol{z} + \alpha + \boldsymbol{R}_{\bar{\Gamma}}^{\mathrm{T}}\boldsymbol{\gamma} \\ &\quad + \rho(\boldsymbol{L}^{\mathrm{T}}\boldsymbol{L}\boldsymbol{x} - \boldsymbol{L}^{\mathrm{T}}\boldsymbol{y}) + \rho(\boldsymbol{x} - \boldsymbol{G}\boldsymbol{f}) + \rho\boldsymbol{R}_{\bar{\Gamma}}\boldsymbol{x} \\ &\to 0. \end{aligned} \tag{5.3}$$

For (2.16), the following problem with respect to $\boldsymbol{Y} = \mathrm{mat}(\boldsymbol{y})$ is solved:

$$\arg\min_{\boldsymbol{Y}} \ \lambda_{\mathrm{TV}}||\boldsymbol{Y}||_{1,2} + \frac{\rho}{2}||\boldsymbol{Y} - \boldsymbol{L}\boldsymbol{x} + \frac{1}{\rho}\mathrm{mat}(\boldsymbol{z})||_F^2. \tag{5.4}$$

The update rules of $\boldsymbol{y}$ and $\boldsymbol{z}$ for solving (5.4) are given by ADMM as

$$\begin{aligned}&\boldsymbol{y}^{k+1} = \mathrm{prox}_{\rho^{-1}\lambda_{\mathrm{TV}}||\cdot||_{1,2}}(\boldsymbol{L}\boldsymbol{x}^{k+1} - \rho^{-1}\boldsymbol{z}^k) \\ &\boldsymbol{z}^{k+1} = \boldsymbol{z}^k + \rho(\boldsymbol{y}^{k+1} - \boldsymbol{L}\boldsymbol{x}^{k+1}). \end{aligned} \tag{5.5}$$

The proximal function for the $l_{1,2}$-norm is defined as

$$[\mathrm{prox}_{\nu||\cdot||_{1,2}}(\boldsymbol{a})]_\iota = \boldsymbol{a}_\iota \circ \max\left(1 - \frac{\nu}{||\boldsymbol{a}_\iota||_2}, 0\right), \tag{5.6}$$

where $\boldsymbol{a} = [\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_I]^{\mathrm{T}}$. By applying $\boldsymbol{w}^k = \boldsymbol{L}\boldsymbol{x}^{k+1} - \rho^{-1}\boldsymbol{z}^k$ to the notations above, the update rule(2.16) can be obtained by

$$\begin{aligned}[\boldsymbol{y}^{k+1}]_s &= \mathrm{prox}_{\rho^{-1}\lambda_{\mathrm{TV}}||\cdot||_{1,2}}(\boldsymbol{w}_s^k) \\ &= \max(1 - \lambda_{\mathrm{TV}}(\rho||\boldsymbol{w}_s^k||_2)^{-1}, 0)\boldsymbol{w}_s^k,\end{aligned} \tag{5.7}$$

where $\boldsymbol{w}_s^k = [w_{s1}^k, w_{s2}^k, w_{s3}^k]^{\mathrm{T}}$.

Finally, the sub-problem with respect to $\boldsymbol{m}_i = \mathrm{vec}(\boldsymbol{\mathcal{M}}_i)$ denoted by

$$\arg\min_{\boldsymbol{m}_i} \ \frac{\lambda_{\mathrm{LR}}}{\mathcal{N}}||\boldsymbol{M}_{i(i)}||_* + \frac{\epsilon}{2}||\boldsymbol{x} - \boldsymbol{m}_i + \boldsymbol{v}_i||_2^2, \tag{5.8}$$

is solved to obtain (2.17). The optimal solution is given by singular value thresholding [31, 30]:

$$\boldsymbol{\mathcal{M}}_i^{k+1} = \mathrm{fold}_i[\mathrm{SVT}_{\frac{\lambda_{\mathrm{LR}}}{\mathcal{N}\epsilon}}(\mathrm{unfold}_j(\boldsymbol{X}^{k+1} + \boldsymbol{V}_i^k))]. \tag{5.9}$$

Optimization for solving Problem (3.5)

In this Appendix, we explain an optimization algorithm to solve Problem (3.5) using a set of spatial bases $\boldsymbol{\Psi}$. Problem (3.5) can be transformed as:

$$\hat{\boldsymbol{Z}} = \arg\min_{\boldsymbol{Z}} \quad F(\boldsymbol{Z}) = D_{\mathrm{KL}}(\tilde{\boldsymbol{Y}}||\boldsymbol{P}\boldsymbol{\Psi}\boldsymbol{Z}) + \frac{\lambda}{2}||\boldsymbol{\Psi}\boldsymbol{Z}\boldsymbol{L}||_F^2 \ , \tag{5.10}$$

where $\boldsymbol{L}$ is a given matrix to differentiate matrix $\boldsymbol{\Psi}\boldsymbol{Z}$ along columns. The derivative of Problem (5.10) is:

$$\frac{\partial F}{\partial \boldsymbol{Z}} = \boldsymbol{\Psi}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\mathbf{1} - \boldsymbol{\Psi}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}(\tilde{\boldsymbol{Y}} \oslash [\boldsymbol{P}\boldsymbol{\Psi}\boldsymbol{Z}]) + \lambda\boldsymbol{\Psi}^{\mathrm{T}}\boldsymbol{\Psi}\boldsymbol{Z}\boldsymbol{L}\boldsymbol{L}^{\mathrm{T}}. \tag{5.11}$$

Then, the multiplicative update rule to solve Problem (5.10) is given by:

$$\boldsymbol{Z}^{(k+1)} \leftarrow \boldsymbol{Z}^{(k)} \circ (\boldsymbol{\Psi}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\{\tilde{\boldsymbol{Y}} \oslash [\boldsymbol{P}\boldsymbol{\Psi}\boldsymbol{Z}^{(\boldsymbol{k})}]\} + [\lambda\boldsymbol{\Psi}^{\mathrm{T}}\boldsymbol{\Psi}\boldsymbol{Z}^{(k)}\boldsymbol{L}\boldsymbol{L}^{\mathrm{T}}]_{-}) \oslash (\boldsymbol{\Psi}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\mathbf{1} + [\lambda\boldsymbol{\Psi}^{\mathrm{T}}\boldsymbol{\Psi}\boldsymbol{Z}^{(k)}\boldsymbol{L}\boldsymbol{L}^{\mathrm{T}}]_{+}), \tag{5.12}$$

where $[\cdot]_{-}$ and $[\cdot]_{+}$ are projection operators, such that $[[\boldsymbol{A}]_{-}]_{vw} = |\min(\boldsymbol{A}_{vw}, 0)|$ and $[[\boldsymbol{A}]_{+}]_{vw} = \max(\boldsymbol{A}_{vw}, 0)$ for an arbitrary matrix $\boldsymbol{A}$.

3D reconstruction of pathological images

In this appendix, we describe the 3D reconstruction method of a series of 2D pathological images [113, 128], which is mentioned in Section 4.1.3.

First of all, a series of M pathological images, $\{I_1,\ I_2,\ \cdots,\ I_M\}$, is assumed to be obtained and given by the following procedure. A series of thin sections with a constant inter-gap is first obtained from a chemically fixed tissue, and their microscopic images are captured after the chemical staining. One can reconstruct a 3D image from the series of the pathological images by piling up the images that are nonrigidly registered together. The nonrigid registration is necessary because of the independent translation, rotation, and deformation of the histological sections caused by the sectioning from the tissue and by the mounting of the sections on glass slides. Before the nonrigid registration of the slice images, respective images are rigidly registered towards through-slice direction to obtain roughly registered pathological images $\{I_1,\ I_2,\ \cdots,\ I_M\}$.

Given a series of the pathological images described above, the reconstruction method [113, 128] performs the following procedure. The method [113, 128] first detects landmarks corresponded among several slice images. The corresponding landmarks across the slice images is assumed to be derived from a part of the same anatomical structures. There are many sets of the corresponding landmarks that describe different parts of different anatomical structures by the set of points. These sets of points would form zig-zag shaped trajectory across the slices because the pathological images are not sufficiently registered yet. Next, each zig-zag trajectory of the corresponding landmarks across the slices is smoothed. The trajectory smoothing is performed by minimizing the total variation of pixel coordinates across the slice images. Finally, the nonrigid registration towards the through-slice direction is processed with the clue of the smoothed sets of the landmarks. The details of respective three processes are described in the next paragraphs.

Landmark Detection and Tracking

The method first iteratively samples a set of keypoints, P_i^j $(j = 1, 2, \dots)$ from $I_i(\cdot)$. P_i^j, which is the j-th keypoints in i-th slice image $I_i(\cdot)$, is sampled with a probability, p_i^j, which is given by

$$p_i^j(\boldsymbol{x}) = \begin{cases} \|\nabla I_i(\boldsymbol{x})\|/Z, & \text{if } (\|\nabla I_i(\boldsymbol{x})\| > T) \wedge \left(\boldsymbol{x} \in \Pi_i^j\right) \\ 0, & \text{otherwise,} \end{cases} \tag{5.13}$$

where Z is a normalization term and T is a threshold. The region from which the j-th keypoint is sampled is restricted to a region, Π_i^j, in order to make the spatial distribution of the keypoints uniform and to avoid an over-sampling.

Once keypoints are detected from a slice image, the respective sets of corresponding points across the slice images are traced by using the template matching. One can employ

conventional techniques, e. g. the normalized cross-correlation (NCC) based method or the SIFT based method, for the detection of the corresponding landmarks. In this study, the NCC-based template matching is employed. In order to suppress false matching, the method at first detects landmarks $P_{i+1}^j (j = 1, 2, \dots)$ in I_{i+1} that correspond to the landmarks P_i^j detected in I_i and then applies backward template matching: Landmarks that correspond to P_{i+1}^j are detected in I_i and the locations of these detected landmarks are compared with the locations of P_i^j. If the locations are largely different, then the method rejects the landmarks found in P_{i+1}^j.

Trajectory Smoothing

The method smooths the trajectories and obtains the destination points, Q_i^j, to which P_i^j should be transported. The coordinates of the destination points, $\boldsymbol{y}_i^j$, are calculated by minimizing the square of the total variation of each trajectory and the square errors with a trade-off parameter λ as follows:

$$\{\boldsymbol{y}_{s^j}^j, \dots, \boldsymbol{y}_{t^j}^j\} = \arg \min_{\tilde{\boldsymbol{y}}_{s^j}^j, \dots, \tilde{\boldsymbol{y}}_{t^j}^j} \left(\sum_{i=s_i}^{t_i - 1} \|\tilde{\boldsymbol{y}}_i^j - \tilde{\boldsymbol{y}}_{i+1}^j\|^2 + \lambda \|\tilde{\boldsymbol{y}}_i^j - \boldsymbol{x}_i^j\|^2 \right). \tag{5.14}$$

The cost function shown in Equation (5.14) is convex and has a unique optimum.

Image Deformation

Once the destination points, $\boldsymbol{y}_i^j$, is obtained for each landmark coordinate, $\boldsymbol{x}_i^j$, then the diffeomorphic mapping, ϕ_i, is computed. Here, one can employ any nonrigid registration technique whose data fidelity term is based on the landmark positions. In [113], a classical B-spline-based deformation method [134] is employed for the computation of the mapping. On the other hand, the LDDMM is employed in [128]. Both of the methods are represented as solving the following optimization problem:

$$\arg \min_{\phi_{i+1}} \; g(\phi_{i+1}) + \sum_j ||\phi_{i+1} \oplus \boldsymbol{y}_{i+1}^j - \boldsymbol{y}_i^j||^2, \tag{5.15}$$

where $g(\cdot)$ is the appropriate smoothness regularizer for the deformation field.

Acknowledgement

First of all, I wish to thank Prof. Hidekata Hontani and Prof. Tatsuya Yokota with great sincerity for helpful detailed guidance. Owing a lot to Prof. Hontani for six years since my bachelor degrees, I have realized that my mathematical and research thinking faculty grown up year by year. Prof. Yokota also really helped me out for my growth on values, knowhow, and coding techniques as a researcher. I also really thank Prof. Mauricio Kugler for several co-researches and helpful code/data preparations. Next, I sincerely thank Prof. Jun Sato and Prof. Taizo Umezaki for reviewing and judging this paper, and for several examinations. I also would like to thank Prof. Jun Sato and Prof. Fumihiko Sakaue, for several chances to perceive the different thinking values and for a lot of help/advices. I wish to thank Ms. Aya Hioki for many office procedures. I with to thank Prof. Hiroshi Matsuzoe for many mathematical advices given when co-researching and when employing me as a teaching assistant. I also thank Prof. Yoshitaka Masutani, Prof. Yuichi Kimura, Prof. Muneyuki Sakata, Dr. Chika Iwamoto, Dr. Kenoki Ohucida, and Prof. Makoto Hashizume for several co-researches and the data offers. I thank all of my colleagues of the laboratory for six years. Finally, I would like to thank my families for allowing and supporting my doctoral degrees cheerfully.

The experimental dataset used in Chapter 2 was provided from the OASIS project, which was funded by NIH Grant Number: P50 AG05681, P01 AG03991, R01 AG021910, P50 MH071616, U24 RR021382, and R01 MH56584. This study was supported by JSPS KAKENHI Grant Number: 26108003, and 15K16067.

Achievements

Journal Papers

- N. Kawamura, T.Yokota, H.Hontani, “Super-Resolution of Magnetic Resonance Images via Convex Optimization with Local and Global Prior Regularization and Spectrum Fitting”, *International Journal of Biomedical Imaging*, 17 pages, 2018.
- N. Kawamura, T.Yokota, H.Hontani, M. Sakata, Y. Kimura,“Parametric PET Image Reconstruction via Regional Spatial Bases and Pharmacokinetic Time Activity Model”, *Entropy*, 20 pages, 2017.

Conference Proceedings

- N. Kawamura, H. Kobayashi, T. Yokota, H. Hontani, C. Iwamoto, K. Ohuchida, and M. Hashizume, ”Landmark”, in Proceesings of SPIE Medical Imaging 2018 : Digital Pathology, 7pages (**Poster Award Cum Laude**).
- 河村直輝，横田達也，本谷秀堅,“Regularized POCS Super-Resolution Based on Global Priors of MR images”, 日本医用画像工学会大会，4 pages, 2017 (**大会奨励賞受賞**).
- 河村直輝，横田達也，本谷秀堅,“空間の低ランク性と平滑性を考慮したフーリエ係数最適化による MR 超解像”, 医用画像研究会 (MI)，2 pages, 2017.
- 河村直輝，横田達也，本谷秀堅, 岩本千佳, 大内田研宙, 橋爪誠,“TV 最小化と周波数適合による MR 画像の高解像度化”, 医用画像研究会 (MI)，6 pages, 2016.
- 河村直輝，本谷秀堅，岩本千佳，大内田研宙，橋爪誠．” 膵臓病理画像の位置合わせを想定した輪郭構造の記述の試み”．医用画像研究会 (MI)，4 pages, 2015.
- M. Kugler, Y. Goto, N. Kawamura, H. Kobayashi, T. Yokota, C. Iwamoto, K. Ohuchida, M. Hashizume, H. Hontani, “Accurate 3D Reconstruction of a Whole Pancreatic Cancer Tumor from Pathology Images with Different Stains”, OMIA 2018, COMPAY 2018, pp 35-43.

- 本谷秀堅，河村直輝，横田達也，松添博，“ q-正規分布における最尤推定法について ”， 医用画像研究会 (MI)，2 pages
- 出浦祐視，河村直輝，山田光統，本谷秀堅，“臓器領域セグメンテーションのための前処理に関する一検討 ～ Superpixel と Saliency 検出 ～ ”， 医用画像研究会 (MI)，5 pages.
- 神谷叔季，河村直輝，本谷秀堅，“ 解剖学的ランドマークを利用する部分医用画像に対する臓器位置合わせ ”， MI 研究会，4 pages.

The Other Awards

- IEICE Tokai Section Student Award 2017.

References

[1] H. Watabe, Y. Ikoma, Y. Kimura, M. Naganawa, and M. Shidahara, "Pet kinetic analysis-compartmental model," *Ann. Nucl. Med.*, vol. 20, no. 9, pp. 583–588, 2006.

[2] I. Bankman, *Handbook of medical image processing and analysis*. Academic Press, 2008.

[3] J. Jan, *Medical image processing, reconstruction and restoration: concepts and methods*. Crc Press, 2005.

[4] T. Sakamoto, "Medical image processing apparatus, medical image processing method, and medical image processing system," June 19 2018. US Patent App. 10/002,423.

[5] S. Li, H. Yin, and L. Fang, "Group-sparse representation with dictionary learning for medical image denoising and fusion," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 12, pp. 3450–3459, 2012.

[6] G. Wang and J. Qi, "Pet image reconstruction using kernel method," *IEEE Trans. Med. Imaging*, vol. 34, no. 1, pp. 61–71, 2015.

[7] J. Mohan, V. Krishnaveni, and Y. Guo, "A survey on the magnetic resonance image denoising methods," *Biomed. Signal Process. Control*, vol. 9, pp. 56–69, 2014.

[8] K. Nasrollahi and T. B. Moeslund, "Super-resolution: a comprehensive survey," *Machine vision and applications*, vol. 25, no. 6, pp. 1423–1468, 2014.

[9] W. Shi, J. Caballero, C. Ledig, X. Zhuang, W. Bai, K. Bhatia, A. M. S. M. de Marvao, T. Dawes, D. O' Regan, and D. Rueckert, "Cardiac image super-resolution with global correspondence using multi-atlas patchmatch," in *Int Conf on Med Image Comput and Comput-Assisted Intervention*, pp. 9–16, Springer, 2013.

[10] J. Yang, J. Wright, T. S. Huang, and Y. Ma, "Image super-resolution via sparse representation," *IEEE Trans. Image Process.*, vol. 19, no. 11, pp. 2861–2873, 2010.

[11] M. Lustig, D. Donoho, and J. M. Pauly, "Sparse MRI: The application of compressed sensing for rapid MR imaging," *Magn. Reson. Med.*, vol. 58, no. 6, pp. 1182–1195, 2007.

[12] M. Lustig, D. L. Donoho, J. M. Santos, and J. M. Pauly, "Compressed sensing MRI," *IEEE Signal Processing Mag*, vol. 25, no. 2, pp. 72–82, 2008.

[13] A. Hirabayashi, N. Inamuro, K. Mimura, T. Kurihara, and T. Homma, "Compressed sensing MRI using sparsity induced from adjacent slice similarity," in *Int Conf Sampling Theory and App*, pp. 287–291, IEEE, 2015.

[14] A. Tikhonov, "Numerical methods for the solution of ill-posed problems,"

[15] T. Yokota and H. Hontani, "Simultaneous visual data completion and denoising based on tensor rank and total variation minimization and its primal-dual splitting algorithm," in *IEEE Conf Compt Vision and Pattern Recog*, pp. 3732–3740, 2017.

[16] F. Cao, M. Cai, and Y. Tan, "Image interpolation via low-rank matrix completion and recovery," *IEEE Trans on Circuits and Syst for Video Technol*, vol. 25, no. 8, pp. 1261–1270, 2015.

[17] L. L. Geyer, U. J. Schoepf, F. G. Meinel, J. W. Nance Jr, G. Bastarrika, J. A. Leipsic, N. S. Paul, M. Rengo, A. Laghi, and C. N. De Cecco, "State of the art: iterative ct reconstruction techniques," *Radiology*, vol. 276, no. 2, pp. 339–357, 2015.

[18] K. Choi, J. Wang, L. Zhu, T.-S. Suh, S. Boyd, and L. Xing, "Compressed sensing based cone-beam computed tomography reconstruction with a first-order method a," *Med. Phys.*, vol. 37, no. 9, pp. 5113–5125, 2010.

[19] D. Youla, "Generalized image restoration by the method of alternating orthogonal projections," *IEEE Trans on Circuits and Syst*, vol. 25, no. 9, pp. 694–702, 1978.

[20] R. W. Gerchberg, "A practical algorithm for the determination of the phase from image and diffraction plane pictures," *Optik*, vol. 35, pp. 237–246, 1972.

[21] R. W. Gerchberg, "Super-resolution through error energy reduction," *Journal of Mod Opt*, vol. 21, no. 9, pp. 709–720, 1974.

[22] M. I. Sezan and H. Stark, "Image restoration by the method of convex projections: Part 2-applications and numerical results," *IEEE Trans on Med Imag*, vol. 1, no. 2, pp. 95–101, 1982.

[23] R. Gerchberg, "Super-resolution through error energy reduction," *J. Mod. Opt.*, vol. 21, no. 9, pp. 709–720, 1974.

[24] A. N. Tikhonov, "Solution of incorrectly formaulated problems and the regularization method," in *Dokl. Akad. Nauk.*, vol. 151, pp. 1035–1038, 1963.

[25] M. A. T. Figueiredo, J. B. Dias, J. P. Oliveira, and R. D. Nowak, "On total variation denoising: A new majorization-minimization algorithm and an experimental comparisonwith wavalet denoising," in *IEEE Conf Image Process*, pp. 2633–2636, IEEE, 2006.

[26] N. Parikh, S. Boyd, *et al.*, "Proximal algorithms," *Foundations and Trends® in Optimization*, vol. 1, no. 3, pp. 127–239, 2014.

[27] L. Xu, S. Zheng, and J. Jia, "Unnatural l0 sparse representation for natural image deblurring," in *IEEE Conf Comput Vision and Pattern Recog*, pp. 1107–1114, 2013.

[28] J. Pan, Z. Hu, Z. Su, and M.-H. Yang, "Deblurring text images via l0-regularized intensity and gradient prior," in *IEEE Conf Comput Vision and Pattern Recog*, pp. 2901–2908, 2014.

[29] S. Ono, "*l*_{0} gradient projection," *IEEE Trans. Image Process.*, vol. 26, no. 4, pp. 1554–1564, 2017.

[30] J. Liu, P. Musialski, P. Wonka, and J. Ye, "Tensor completion for estimating missing values in visual data," *IEEE Trans on Pattern Anal Mach Intell*, vol. 35, no. 1, pp. 208–220, 2013.

[31] J. F. Cai, E. J. Candès, and Z. Shen, "A singular value thresholding algorithm for matrix completion," *SIAM Journal on Optim*, vol. 20, no. 4, pp. 1956–1982, 2010.

[32] J.-B. Huang, A. Singh, and N. Ahuja, "Single image super-resolution from transformed self-exemplars," in *IEEE Conf Compt Vision and Pattern Recog*, pp. 5197–5206, 2015.

[33] C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 2, pp. 295–307, 2016.

[34] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," *arXiv preprint arXiv:1502.03167*, 2015.

[35] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

[36] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. Van Der Laak, B. Van Ginneken, and C. I. Sánchez, "A survey on deep learning in medical image analysis," *Med. Image Anal.*, vol. 42, pp. 60–88, 2017.

[37] J. Johnson, A. Alahi, and L. Fei-Fei, "Perceptual losses for real-time style transfer and super-resolution," in *European Conference on Computer Vision*, pp. 694–711, Springer, 2016.

[38] P. Vincent, H. Larochelle, Y. Bengio, and P.-A. Manzagol, "Extracting and composing robust features with denoising autoencoders," in *25th Int Conf on Machine learning*, pp. 1096–1103, ACM, 2008.

[39] K. H. Jin, M. T. McCann, E. Froustey, and M. Unser, "Deep convolutional neural network for inverse problems in imaging," *IEEE Trans. Image Process.*, vol. 26, no. 9, pp. 4509–4522, 2017.

[40] C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, *et al.*, "Photo-realistic single image super-resolution using a generative adversarial network," *arXiv preprint arXiv:1609.04802*, 2016.

[41] E. L. Denton, S. Chintala, a. szlam, and R. Fergus, "Deep generative image models using a laplacian pyramid of adversarial networks," pp. 1486–1494, Curran Associates, Inc, 2015.

[42] W.-S. Lai, J.-B. Huang, N. Ahuja, and M.-H. Yang, "Deep laplacian pyramid networks for fast and accurate superresolution," in *IEEE Conf Comput Vision and Pattern Recog*, vol. 2, p. 5, 2017.

[43] H. Greenspan, B. Van Ginneken, and R. M. Summers, "Guest editorial deep learning in medical imaging: Overview and future promise of an exciting new technique," *IEEE Trans. Med. Imaging*, vol. 35, no. 5, pp. 1153–1159, 2016.

[44] F. Milletari, N. Navab, and S.-A. Ahmadi, "V-net: Fully convolutional neural networks for volumetric medical image segmentation," in *4th Int Conf on 3D Vision*, pp. 565–571, IEEE, 2016.

[45] A. Qayyum, S. M. Anwar, M. Awais, and M. Majid, "Medical image retrieval using deep convolutional neural network," *Neurocomputing*, vol. 266, pp. 8–20, 2017.

[46] H. Greenspan, "Super-resolution in medical imaging," *The Comput Journal*, vol. 52, no. 1, pp. 43–63, 2009.

[47] J. L. Prince and J. M. Links, *Medical Imaging Signals and Systems*. Pearson Prentice Hall Upper Saddle River, New Jersey, 2006.

[48] E. Y. Sidky, C.-M. Kao, and X. Pan, "Accurate image reconstruction from few-views and limited-angle data in divergent-beam CT," *Journal of X-ray Sci and Technol*, vol. 14, no. 2, pp. 119–139, 2006.

[49] P. Chatterjee, S. Mukherjee, S. Chaudhuri, and G. Seetharaman, "Application of Papoulis–Gerchberg method in image super-resolution and inpainting," *The Comput Journal*, vol. 52, no. 1, pp. 80–89, 2009.

[50] C.-Y. Hsu and T.-M. Lo, "Improved Papoulis-Gerchberg algorithm for restoring lost samples," in *IEEE Symp Signal Process Info Technol*, pp. 717–721, IEEE, 2005.

[51] L. M. Kani and J. C. Dainty, "Super-resolution using the Gerchberg algorithm," *Opt. Commun.*, vol. 68, no. 1, pp. 11–17, 1988.

[52] S. D. Babacan, R. Molina, and A. K. Katsaggelos, "Total variation super resolution using a variational approach," in *IEEE Conf Image Process*, pp. 641–644, IEEE, 2008.

[53] A. Chambolle and T. Pock, "A first-order primal-dual algorithm for convex problems with applications to imaging," *Journal of Math Imag and Vision*, vol. 40, no. 1, pp. 120–145, 2011.

[54] M. K. Ng, H. Shen, E. Y. Lam, and L. Zhang, "A total variation regularization based super-resolution reconstruction algorithm for digital video," *EURASIP Journal on Adv in Signal Process*, vol. 2007, pp. 1–16, 2007.

[55] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys D: Nonlinear Phenom*, vol. 60, no. 1-4, pp. 259–268, 1992.

[56] F. Shi, J. Cheng, L. Wang, P.-T. Yap, and D. Shen, "LRTV: MR image super-resolution with low-rank and total variation regularizations," *IEEE Trans on Med Imag*, vol. 34, no. 12, pp. 2459–2466, 2015.

[57] T. Yokota and H. Hontani, "An efficient method for adapting step-size parameters of primal-dual hybrid gradient method in application to total variation regularization," in *Asia-Pacific Signal and Info Process Assoc Annu Summit and Conf (APSIPA ASC)*, pp. 973–979, IEEE, 2017.

[58] M. Zhu and T. Chan, "An efficient primal-dual hybrid gradient algorithm for total variation image restoration," *UCLA CAM Rep*, pp. 08–34, 2008.

[59] D. Gabay and B. Mercier, "A dual algorithm for the solution of nonlinear variational problems via finite element approximation," *Comput & Math with Appl*, vol. 2, no. 1, pp. 17–40, 1976.

[60] S. Yang, J. Wang, W. Fan, X. Zhang, P. Wonka, and J. Ye, "An efficient ADMM algorithm for multidimensional anisotropic total variation regularization problems,"

in *ACM SIGKDD Conf Knowledge Discov and Data Mining*, pp. 641–649, ACM, 2013.

[61] T. Yokota, Q. Zhao, and A. Cichocki, "Smooth PARAFAC decomposition for tensor completion," *IEEE Trans on Signal Processing*, vol. 64, no. 20, pp. 5423–5436, 2015.

[62] E. Candès and B. Recht, "Exact matrix completion via convex optimization," *Commun of the ACM*, vol. 55, no. 6, pp. 111–119, 2012.

[63] E. J. Candès and T. Tao, "The power of convex relaxation: Near-optimal matrix completion," *IEEE Trans on Inf Theory*, vol. 56, no. 5, pp. 2053–2080, 2010.

[64] B. Recht, M. Fazel, and P. A. Parrilo, "Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization," *SIAM Rev.*, vol. 52, no. 3, pp. 471–501, 2010.

[65] F. Guichard and F. Malgouyres, "Total variation based interpolation," in *European Signal Process Conf*, pp. 1–4, IEEE, 1998.

[66] A. Marquina and S. J. Osher, "Image super-resolution by TV-regularization and Bregman iteration," *Journal of Sci Comput*, vol. 37, no. 3, pp. 367–382, 2008.

[67] H. Greenspan, G. Oz, N. Kiryati, and S. Peled, "MRI inter-slice reconstruction using super-resolution," *Magn. Reson. Imaging*, vol. 20, no. 5, pp. 437–446, 2002.

[68] S. Jiang, H. Xue, A. Glover, M. Rutherford, D. Rueckert, and J. V. Hajnal, "MRI of moving subjects using multislice snapshot images with volume reconstruction (SVR): application to fetal, neonatal, and adult brain studies," *IEEE Trans on Med Imag*, vol. 26, no. 7, pp. 967–980, 2007.

[69] V. S. Lee, *Cardiovascular MRI: physical principles to practical protocols*. Lippincott Williams & Wilkins, 2006.

[70] C. A. Cocosco, V. Kollokian, R. K.-S. Kwan, G. B. Pike, and A. C. Evans, "Brainweb: Online interface to a 3d MRI simulated brain database," in *NeuroImage*, Citeseer, 1997.

[71] D. S. Marcus, T. H. Wang, J. Parker, J. G. Csernansky, J. C. Morris, and R. L. Buckner, "Open access series of imaging studies (OASIS): cross-sectional mri data in young, middle aged, nondemented, and demented older adults," *Journal of Cogn Neurosci*, vol. 19, no. 9, pp. 1498–1507, 2007.

[72] T. M. Lehmann, C. Gonner, and K. Spitzer, "Survey: Interpolation methods in medical image processing," *IEEE Trans on Med Imag*, vol. 18, no. 11, pp. 1049–1075, 1999.

[73] D. C. Youla and H. Webb, "Image restoration by the method of convex projections: Part 1-theory," *IEEE Trans on Med Imag*, vol. 1, no. 2, pp. 81–94, 1982.

[74] C. He, C. Hu, W. Zhang, and B. Shi, "A fast adaptive parameter estimation for total variation image restoration," *IEEE Trans. Image Process.*, vol. 23, no. 12, pp. 4954–4967, 2014.

[75] D. Ayyagari, N. Ramesh, D. Yatsenko, T. Tasdizen, and C. Atria, "Image reconstruction using priors from deep learning," in *Med Imag 2018: Image Process*, vol. 10574, p. 105740H, International Society for Optics and Photonics, 2018.

[76] D. Ulyanov, A. Vedaldi, and V. Lempitsky, "Deep image prior," *arXiv preprint arXiv:1711.10925*, 2017.

[77] H. M. Hudson and R. S. Larkin, "Accelerated image reconstruction using ordered subsets of projection data," *IEEE Trans. Med. Imaging*, vol. 13, no. 4, pp. 601–609, 1994.

[78] J. A. Fessler and W. L. Rogers, "Spatial resolution properties of penalized-likelihood image reconstruction: space-invariant tomographs," *IEEE Trans. Image Process.*, vol. 5, no. 9, pp. 1346–1358, 1996.

[79] J. Qi, R. M. Leahy, S. R. Cherry, A. Chatziioannou, and T. H. Farquhar, "High-resolution 3d bayesian image reconstruction using the micropet small-animal scanner," *Phys. Med. Biol.*, vol. 43, no. 4, p. 1001, 1998.

[80] G. Wang and J. Qi, "Generalized algorithms for direct reconstruction of parametric images from dynamic pet data," *IEEE Trans. Med. Imaging*, vol. 28, no. 11, pp. 1717–1726, 2009.

[81] S. Ross, "Q. clear," *GE Healthcare, White Paper*, pp. 1–9, 2014.

[82] E. U. Mumcuoglu, R. Leahy, S. R. Cherry, and Z. Zhou, "Fast gradient-based methods for bayesian reconstruction of transmission and emission pet images," *IEEE Trans. Med. Imaging*, vol. 13, no. 4, pp. 687–701, 1994.

[83] J. S. Lee, D. D. Lee, S. Choi, K. S. Park, and D. S. Lee, "Non-negative matrix factorization of dynamic images in nuclear medicine," in *Nuclear Science Symposium Conference Record*, vol. 4, pp. 2027–2030, IEEE, 2001.

[84] J. S. Lee, D. D. Lee, S. Choi, and D. S. Lee, "Application of nonnegative matrix factorization to dynamic positron emission tomography," in *The 3rd Int Conf on Independent Component Analysis and Blind Signal Separation*, pp. 629–632, 2001.

[85] N. Pustelnik, C. Chaux, J.-C. Pesquet, and C. Comtat, "Parallel algorithm and hybrid regularization for dynamic pet reconstruction," in *Nuclear Science Symposium Conference Record*, pp. 2423–2427, IEEE, 2010.

[86] S. Ahn, S. M. Kim, J. Son, D. S. Lee, and J. Sung Lee, "Gap compensation during pet image reconstruction by constrained, total variation minimization," *Med. Phys.*, vol. 39, no. 2, pp. 589–602, 2012.

[87] M. Burger, J. Müller, E. Papoutsellis, and C.-B. Schönlieb, "Total variation regularization in measurement and image space for pet reconstruction," *Inverse Prob.*, vol. 30, no. 10, p. 105003, 2014.

[88] T. Li, B. Thorndyke, E. Schreibmann, Y. Yang, and L. Xing, "Model-based image reconstruction for four-dimensional pet," *Med. Phys.*, vol. 33, no. 5, pp. 1288–1298, 2006.

[89] M. E. Kamasak, C. A. Bouman, E. D. Morris, and K. Sauer, "Direct reconstruction of kinetic parameter images from dynamic pet data," *IEEE Trans. Med. Imaging*, vol. 24, no. 5, pp. 636–650, 2005.

[90] J. C. Matthews, G. I. Angelis, F. A. Kotasidis, P. J. Markiewicz, and A. J. Reader, "Direct reconstruction of parametric images using any spatiotemporal 4d image based model and maximum likelihood expectation maximisation," in *Nuclear Science Symp Conf Record*, pp. 2435–2441, IEEE, 2010.

[91] A. J. Reader and H. Zaidi, "Advances in pet image reconstruction," *PET clinics*, vol. 2, no. 2, pp. 173–190, 2007.

[92] Y. Vardi, L. Shepp, and L. Kaufman, "A statistical model for positron emission tomography," *J. Am. Stat. Assoc.*, vol. 80, no. 389, pp. 8–20, 1985.

[93] R. E. Carson, Y. Yan, M. E. Daube-Witherspoon, N. Freedman, S. L. Bacharach, and P. Herscovitch, "An approximation formula for the variance of pet region-of-interest values," *IEEE Trans. Med. Imaging*, vol. 12, no. 2, pp. 240–250, 1993.

[94] H. H. Barrett, D. W. Wilson, and B. M. Tsui, "Noise properties of the em algorithm. i. theory," *Phys. Med. Biol.*, vol. 39, no. 5, p. 833, 1994.

[95] J. A. Fessler, "Penalized weighted least-squares image reconstruction for positron emission tomography," *IEEE Trans. Med. Imaging*, vol. 13, no. 2, pp. 290–300, 1994.

[96] M. N. Wernick, E. J. Infusino, and M. Milosevic, "Fast spatio-temporal image reconstruction for dynamic pet," *IEEE Trans. Med. Imaging*, vol. 18, no. 3, pp. 185–195, 1999.

[97] R. N. Gunn, A. A. Lammertsma, S. P. Hume, and V. J. Cunningham, "Parametric imaging of ligand-receptor binding in pet using a simplified reference region model," *Neuroimage*, vol. 6, no. 4, pp. 279–287, 1997.

[98] F. E. Turkheimer, P. Edison, N. Pavese, F. Roncaroli, A. N. Anderson, A. Hammers, A. Gerhard, R. Hinz, Y. F. Tai, and D. J. Brooks, "Reference and target region modeling of [11c]-(r)-pk11195 brain studies," *J. Nucl. Med.*, vol. 48, no. 1, pp. 158–167, 2007.

[99] J. Logan, J. S. Fowler, N. D. Volkow, A. P. Wolf, S. L. Dewey, D. J. Schlyer, R. R. MacGregor, R. Hitzemann, B. Bendriem, S. J. Gatley, *et al.*, "Graphical analysis of reversible radioligand binding from time-activity measurements applied to [n-11c-methyl]-(-)-cocaine pet studies in human subjects," *J. Cereb. Blood Flow Metab.*, vol. 10, no. 5, pp. 740–747, 1990.

[100] A. A. Lammertsma and S. P. Hume, "Simplified reference tissue model for pet receptor studies," *Neuroimage*, vol. 4, no. 3, pp. 153–158, 1996.

[101] J. Logan, "Graphical analysis of pet data applied to reversible and irreversible tracers," *Nucl. Med. Biol.*, vol. 27, no. 7, pp. 661–670, 2000.

[102] M. Yaqub, B. N. Van Berckel, A. Schuitemaker, R. Hinz, F. E. Turkheimer, G. Tomasi, A. A. Lammertsma, and R. Boellaard, "Optimization of supervised cluster analysis for extracting reference tissue input curves in (r)-[11c] pk11195 brain pet studies," *J. Cereb. Blood Flow Metab.*, vol. 32, no. 8, pp. 1600–1608, 2012.

[103] Y. Wu and R. E. Carson, "Noise reduction in the simplified reference tissue model for neuroreceptor functional imaging," *J. Cereb. Blood Flow Metab.*, vol. 22, no. 12, pp. 1440–1452, 2002.

[104] D. D. Lee and H. S. Seung, "Algorithms for non-negative matrix factorization," in *Advances in Neural Info Process Syst*, pp. 556–562, 2001.

[105] A. J. Reader, J. C. Matthews, F. C. Sureau, C. Comtat, R. Trebossen, and I. Buvat, "Iterative kinetic parameter estimation within fully 4d pet image reconstruction," in *Nuclear Science Symposium Conference Record*, vol. 3, pp. 1752–1756, IEEE, 2006.

[106] L. A. Shepp and Y. Vardi, "Maximum likelihood reconstruction for emission tomography," *IEEE Trans. Med. Imaging*, vol. 1, no. 2, pp. 113–122, 1982.

[107] J. Logan, J. S. Fowler, N. D. Volkow, G.-J. Wang, Y.-S. Ding, and D. L. Alexoff, "Distribution volume ratios without blood sampling from graphical analysis of pet data," *J. Cereb. Blood Flow Metab.*, vol. 16, no. 5, pp. 834–840, 1996.

[108] P. Gravel and A. J. Reader, "Direct 4d pet mlem reconstruction of parametric images using the simplified reference tissue model with the basis function method for [11c] raclopride," *Phys. Med. Biol.*, vol. 60, no. 11, p. 4533, 2015.

[109] G. Wang and J. Qi, "Direct estimation of kinetic parametric images for dynamic pet," *Theranostics*, vol. 3, no. 10, p. 802, 2013.

[110] C. Tsoumpas, F. E. Turkheimer, and K. Thielemans, "A survey of approaches for direct parametric image reconstruction in emission tomography," *Med. Phys.*, vol. 35, no. 9, pp. 3963–3971, 2008.

[111] D. Komura and S. Ishikawa, "Machine learning methods for histopathological image analysis," *Comput. Struct. Biotechnol. J.*, vol. 16, pp. 34–42, 2018.

[112] H. Wu, J. H. Phan, A. K. Bhatia, B. Shehata, and M. D. Wang, "Detection of blur artifacts in histopathological whole-slide images of endomyocardial biopsies," in *IEEE Engineering in Medicine and Biology Society. Annu. Conf.*, vol. 2015, p. 727, NIH Public Access, 2015.

[113] N. Kawamura, H. Kobayashi, T. Yokota, H. Hontani, C. Iwamoto, K. Ohuchida, and M. Hashizume, "Landmark-based reconstruction of 3d smooth structures from serial histological sections," in *Medical Imaging 2018: Digital Pathology*, vol. 10581, p. 105811E, International Society for Optics and Photonics, 2018.

[114] M. F. Beg, M. I. Miller, A. Trouvé, and L. Younes, "Computing large deformation metric mappings via geodesic flows of diffeomorphisms," *Int. J. Comput. Vision*, vol. 61, no. 2, pp. 139–157, 2005.

[115] M. I. Miller, L. Younes, and A. Trouvé, "Diffeomorphometry and geodesic positioning systems for human anatomy," *Technology*, vol. 2, no. 01, pp. 36–43, 2014.

[116] X. Kong and L. Yang, "Microscopic muscle image enhancement," *arXiv preprint arXiv:1612.05719*, 2016.

[117] J. Pan, D. Sun, H. Pfister, and M.-H. Yang, "Blind image deblurring using dark channel prior," in *IEEE Conf on Comput Vision and Pattern Recog*, pp. 1628–1636, 2016.

[118] S. Cho and S. Lee, "Fast motion deblurring," in *ACM Transactions on graphics (TOG)*, vol. 28, p. 145, ACM, 2009.

[119] S. Anwar, C. P. Huynh, and F. Porikli, "Image deblurring with a class-specific prior," *IEEE Trans. Pattern Anal. Mach. Intell.*, 2018.

[120] S. Anwar, C. Phuoc Huynh, and F. Porikli, "Class-specific image deblurring," in *IEEE Conf on Comput Vision*, pp. 495–503, 2015.

[121] M. S. Almeida and M. A. Figueiredo, "Blind image deblurring with unknown boundaries using the alternating direction method of multipliers," in *IEEE Conf on Image Process*, pp. 586–590, IEEE, 2013.

[122] P. Campisi and K. Egiazarian, *Blind image deconvolution: theory and applications*. CRC press, 2016.

[123] S. Smith, A. Beri, and G. Karypis, "Constrained tensor factorization with accelerated ao-admm," in *46th Int Conf on Parallel Process (ICPP)*, pp. 111–120, IEEE, 2017.

[124] K. Huang, N. D. Sidiropoulos, and A. P. Liavas, "A flexible and efficient algorithmic framework for constrained matrix and tensor factorization," *IEEE Trans. Signal Process.*, vol. 64, no. 19, pp. 5052–5065, 2016.

[125] F. P. Oliveira and J. M. R. Tavares, "Medical image registration: a review," *Comput. Methods Biomech. Biomed. Engin.*, vol. 17, no. 2, pp. 73–93, 2014.

[126] J. A. Maintz and M. A. Viergever, "A survey of medical image registration," *Medical imageanalysis*, vol. 2, no. 1, pp. 1–36, 1998.

[127] B. Zitova and J. Flusser, "Image registration methods: a survey," *Image Vision Comput.*, vol. 21, no. 11, pp. 977–1000, 2003.

[128] M. Kugler, Y. Goto, N. Kawamura, H. Kobayashi, T. Yokota, C. Iwamoto, K. Ohuchida, M. Hashizume, and H. Hontani, "Accurate 3d reconstruction of a whole pancreatic cancer tumor from pathology images with different stains," in *Comput Pathology and Ophthalmic Med Image Anal*, pp. 35–43, Springer, 2018.

[129] K. P. Olive, M. A. Jacobetz, C. J. Davidson, A. Gopinathan, D. McIntyre, D. Honess, B. Madhu, M. A. Goldgraben, M. E. Caldwell, D. Allard, *et al.*, "Inhibition of hedgehog signaling enhances delivery of chemotherapy in a mouse model of pancreatic cancer," *Science*, 2009.

[130] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, 2004.

[131] Z. Hu and M.-H. Yang, "Good regions to deblur," in *European Conference on Computer Vision*, pp. 59–72, Springer, 2012.

[132] X. Wang, K. Yu, C. Dong, and C. C. Loy, "Recovering realistic texture in image super-resolution by deep spatial feature transform," *arXiv preprint arXiv:1804.02815*, 2018.

[133] X. Huang and S. J. Belongie, "Arbitrary style transfer in real-time with adaptive instance normalization.," in *ICCV*, pp. 1510–1519, 2017.

[134] D. Rueckert, P. Aljabar, R. A. Heckemann, J. V. Hajnal, and A. Hammers, "Diffeomorphic registration using b-splines," in *Int Conf on Med Image Comput and Comput-Assisted Intervention*, pp. 702–709, Springer, 2006.