

分类号 _____

密级 _____

UDC^{注 1} _____



南京理工大学
NANJING UNIVERSITY OF SCIENCE & TECHNOLOGY

博士 学位 论 文

基于变分和稀疏表示的定量 MR 图像 快速重建模型和加速算法

(题名和副题名)

王冬

(作者姓名)

指导教师姓名 _____ 杨孝平 教授

学 位 类 别 _____ 工学博士

学 科 名 称 _____ 控制科学与工程

研 究 方 向 _____ 图像处理的数学建模、理论和算法

论文提交日期 _____ 2019.11

注 1: 注明《国际十进分类法 UDC》的类号

基于变分和稀疏表示的定量MR图像快速重建模型和加速算法

南京理工大学

博士 学位 论文

基于变分和稀疏表示的定量 MR 图像快速重建
模型和加速算法

作 者：王冬

指导教师：杨孝平 教授

南京理工大学

2019 年 11 月

Ph.D. Dissertation

**Variational Methods and Sparse
Representation Based Fast
Reconstruction Models and Acceleration
Algorithms for Quantitative MRI**

By

Dong Wang

Supervised by Prof. Xiaoping Yang

Nanjing University of Science & Technology

November, 2019

声 明

本学位论文是我在导师的指导下取得的研究成果, 尽我所知, 在本学位论文中, 除了加以标注和致谢的部分外, 不包含其他人已经发表或公布过的研究成果, 也不包含我为获得任何教育机构的学位或学历而使用过的材料。与我一同工作的同事对本学位论文做出的贡献均已在论文中作了明确的说明。

研究生签名: _____ 年 月 日

学位论文使用授权说明

南京理工大学有权保存本学位论文的电子和纸质文档, 可以借阅或上网公布本学位论文的部分或全部内容, 可以向有关部门或机构送交并授权其保存、借阅或上网公布本学位论文的部分或全部内容。对于保密论文, 按保密的有关规定和程序处理。

研究生签名: _____ 年 月 日

摘要

核磁共振成像 (magnetic resonance imaging, 简称 MRI) 是医学临床和研究中最常用的成像方式, 可以非侵入式地获取人体内组织的信息, 用于指导病灶检测与医学临床诊断。但是由于 MR 成像受到物理和生理上的限制, 其成像速度通常很慢, 尤其是动态 MR 图像 (dynamic MRI, 简称 dMRI)。因此, 如何在保证图像质量的前提下, 提高成像速度一直是研究人员最关心的问题。

压缩感知 (compressed sensing) 是近十年来图像处理领域的研究热点之一。作为一种新兴的采样理论, 压缩感知可以突破传统 Nyquist 采样定理的瓶颈, 用远少于 Nyquist 采样所需要的数据精确地重建出图像。压缩感知理论有三个主要的部分, 即图像稀疏性、随机下采样和非线性重建算法。具体来说, 压缩感知是在假设图像在某个变换域稀疏的前提下, 对图像进行随机下采样, 利用非线性算法重建图像。因此, 压缩感知理论可以加快成像速度, 减少存储压力, 提高图像的时间分辨率和空间分辨率。自从压缩感知理论诞生以来, 就被广泛应用于医学成像领域, 尤其是核磁共振成像中。对于 MR 图像的压缩感知模型而言, 最重要一点是寻找适合于 MR 图像的稀疏表示。此外, 对于动态 MR 图像, 时间分辨率往往和空间分辨率同样重要, 因此如何选择合适的时间方向上的稀疏表示, 提高动态 MR 图像的时空分辨率一直是一个具有挑战的问题。

定量磁共振成像技术 (quantitative MRI, 简称 qMRI) 是利用某些特殊的成像序列, 测量组织参数的成像技术, 如横向弛豫时间 (T_1)、纵向弛豫时间 (T_2)、质子密度等。相比于传统的定性 (qualitative) MR 图像, 定量 MRI 从客观定量的角度研究人体组织, 起到了帮助诊断与评估治疗的作用。其中磁共振动态对比增强 (dynamic contrast enhanced MRI, 简称 DCE-MRI) 是近些年来定量成像的研究热点, 其通过获取注入对比剂前后的图像, 经过一系列的计算分析, 得到定量或半定量的参数, 用于指导临床诊断。目前, 压缩感知理论已经被应用在加速 DCE-MRI 成像中, 但对于胸部 DCE-MRI 图像而言, 如何选择和评估时间方向上的稀疏项一直是未知的。

磁共振指纹 (magnetic resonance fingerprinting, 简称 MRF) 是定量 MRI 的新方法, 可以在单次数据采集中同时获得多种组织参数。磁共振指纹主要分为三个部分, 即预定义字典的生成、信号采集和模式识别。具体来说, 给定某个 MR 序列, 首先用 MR 成像的数学模型模拟生成一个包含不同参数的组织随时间演化的字典, 然后选择某个模式识别算法将采集得到的信号 (通常被称为指纹) 与字典中的原子进行匹配, 从而重建出组织的参数图像。由于信号的长度一般在 1,000 以上, 字典中的原子个数也经常达到 10,000 以上, 字典生成与匹配所消耗的时间通常达到几十分钟甚至几小时。因此, 如何快速精确地生成字典并且进行字典匹配是一个亟待解决的问题。

针对以上问题, 本文研究的工作以及创新点主要有以下三个方面:

1. 针对动态 MR 图像, 利用压缩感知和图像分解的思想, 提出了基于二阶时空广义全变差 (total generalized variation, 简称 TGV) 和核范数的重建模型。模型将图像分解为低秩部分和稀疏部分, 其中低秩部分用核范数约束, 稀疏部分用二阶时空 TGV 约束。核范数用来建模动态图像中时间方向高度相关的背景部分, 可以很好地去除空间伪影; 而 TGV 泛函用来表示图像中的光滑部分, 可以在保证重建图像边界清晰的同时减少重建图像中的阶梯效应。我们利用 Primal-Dual 算法来求解模型, 并给出了保证算法收敛的范数估计。为了减少算法的计算时间, 我们也使用了图形处理单元 (graphics processing unit, 简称 GPU) 进行编程来加速程序运行。我们针对体模、心脏灌注与胸部 DCE-MRI 图像, 对比了四种最前沿的针对动态 MR 重建模型在不同采样模式和不同采样率下的表现。数值实验结果表明, 相比于其他四个模型, 我们提出的模型在不同采样模式和不同采样率下, 可以更好地消除空间伪影并且保证图像边缘的清晰, 对于胸部 DCE-MRI 图像的效果尤其显著。

2. 针对胸部 DCE-MRI 图像, 比较了五种不同的时间方向上的稀疏正则项在压缩感知重建模型中的表现, 并对重建结果进行了定量分析。这五种稀疏项分别为 Fourier 变换、Haar 小波变换、全变差 (total variation, 简称 TV)、二阶 TGV 和核范数。所有模型均使用 FISTA (fast iterative shrinkage threshold algorithm) 快速算法进行求解。数值实验结果表明, 基于核范数的模型可以得到最高的信噪比, 而基于 TV/二阶 TGV 的模型可以得到最精确地定量分析。因此, 对于胸部 DCE-MRI, 在重建模型中选择 TV/二阶 TGV 作为正则项可以更好地重建病灶部分, 而选择核范数则可以提高图像的整体信噪比。

3. 针对 MRF 重建参数图速度慢的问题, 我们利用 GPU 进行 MRF 字典生成与匹配, 并设计开发了一款开源程序 snapMRF。程序基于 Nvidia 公司的底层编程框架 CUDA, 使用 C 语言编写。snapMRF 将所有计算都放在 GPU 上进行, 最大限度地减少了拷贝数据的开销。snapMRF 可以快速并准确地重建参数图, 并且适用于不同的 MR 序列。数值试验表明, 相比于其他 MRF 重建的开源程序, 如 EPG-X (基于 CPU 的 MATLAB 语言编写) 和 PnP-MRF (基于 CPU 的 C 语言编写), snapMRF 的字典生成速度提高了 10–1000 倍, 字典匹配速度提高了 10–100 倍。对于较小的字典, 可以达到实时的效果。此外, 程序也给出了六种经典 MR 序列的单元测试, 保证对于不同输入的 MR 序列, 字典生成与匹配的正确性。

关键词: 核磁共振成像, 动态核磁共振, 压缩感知, 变分模型, 稀疏表示, 磁共振指纹成像, 图像分解, 核范数, 广义全变差, 定量分析, 图形处理单元

Abstract

Magnetic resonance imaging (MRI) is the most commonly used imaging methodology in medical clinical research, which can non-invasively acquire *in vivo* tissue information and help to guide lesion diagnosis and detection. However, as the speed of MRI is fundamentally limited by physical and physiological constraints, the imaging speed of MRI is usually slow, especially for dynamic MRI (dMRI). Thus, how to accelerate the imaging speed without degenerating the image quality has always been the most concerned issue for MR researchers.

Compressed sensing has been a hot topic for the last decade in image processing. As the state-of-the-art sampling theory, compressed sensing breaks through the bottleneck of traditional Nyquist sampling theorem and is able to accurately reconstruct the image using only a small number of sampling values – much smaller than the Nyquist sampling theorem. Compressed sensing has three main components, which are sparse representation, random undersampling and nonlinear reconstruction algorithm, respectively. Specifically, given that the image is sparse in a certain transform domain, compressed sensing demonstrates that it is possible to reconstruct the image through a small number of random linear measures using a nonlinear reconstruction algorithm. Therefore compressed sensing has the potential to accelerate imaging speed as well as decrease storage burden, and has been widely used in medical imaging, especially in MRI. Compressed sensing can accelerate the whole procedure of MRI and improve both spatial and temporal resolution. For compressed sensing based MR reconstruction models, the most essential procedure is to find a proper sparsifying transform. Furthermore, for dynamic MRI, temporal resolution is often as important as spatial resolution. Therefore, how to select a proper temporal sparse representation for dynamic MR images has always been a challenging problem.

Quantitative MRI (qMRI) is the technology of measuring tissue parameters, such as T_1 , T_2 and proton density, using certain MR sequences. In contrast to traditional qualitative MRI, qMRI studies human tissues from an objective and quantitative perspective, which can help to guide diagnosis and assess treatment. DCE-MRI (dynamic contrast enhanced MRI) has been a hot topic in qMRI these years, which can guide the clinical diagnosis and provide semiquantitative or quantitative parameters by acquiring and analyzing the images before and after the injection of the contrast agent. Nowadays compressed sensing theory has been adopted to accelerate DCE-MRI reconstruction, but for

breast DCE-MRI, how to select and assess temporal sparsifying transforms is unknown.

Magnetic resonance fingerprinting (MRF) is the state-of-the-art quantitative imaging technique, which is able to achieve multiple tissue parameters in one single acquisition. MRF is consisted of three main parts, which are generation of pre-defined dictionary, signal acquisition and pattern recognition, respectively. More specifically, given a certain MR sequence, a dictionary of simulated signal time courses is first generated for anticipated combinations of tissue properties using an appropriate signal model, and then pattern recognition is used to match each voxel in the acquisition (also called fingerprinting) to the dictionary to reconstruct the final parameter maps. As the signal length is often longer than 1,000, and the number of atoms in the dictionary is often over 10,000, it usually takes minutes or even hours to generate dictionaries and perform matching. Therefore, how to fast and accurately generate dictionaries and perform matching is an urgent problem.

Based on the above problems, the main work and innovation of the thesis can be listed as follows:

1. For dynamic MR images, we propose a novel decomposition based model exploiting the second-order spatio-temporal total generalized variation (TGV) and the nuclear norm for compressed sensing dynamic MR reconstruction. The nuclear norm can model time-coherent background and perform well in removing spatial artifacts, and the TGV functional can represent the smooth area, preserve the edges as well as reduce staircase effects. We first employ the Primal-Dual algorithm to solve the proposed model and then give the norm estimation for the convergence condition. We also implement the proposed model using GPU (graphics processing unit) in CUDA C to accelerate the MATLAB code. Numerical experiments on PINCAT, cardiac perfusion and breast DCE-MRI datasets indicate that the proposed model outperforms the state-of-the-art methods in both suppressing the spatial artifacts and preserving the edges under different acceleration factors and different sampling schemes, especially for breast DCE-MRI datasets.

2. For DCE-MRI of the breast, we compare five different sparse regularizers in the temporal direction and perform quantitative analysis. The five temporal regularizers are the Fourier transform, the Haar wavelet transform, total variation (TV), second-order TGV and the nuclear norm, respectively. All the models are solved using FISTA (fast iterative shrinkage thresholding algorithm). Numerical experiments demonstrate that the nuclear norm provides the highest SER while TV/TGV brings the most accurate parameter analysis. Therefore we recommend using TV/TGV as the temporal constraint in compressed sensing reconstructions of breast DCE-MRI.

3. Due to the long running time in MRF, we propose to use GPU to accelerate both

dictionary generation and matching, and release an open source online program called snapMRF. snapMRF is based on the framework CUDA (compute unified device architecture) of NVIDIA and written in C language. snapMRF manipulates all the computations in GPU and thus minimizes the cost of data copy between GPU and CPU. snapMRF can fast and accurately reconstruct parameter maps and apply to different MR sequences. Compared with other online MRF reconstruction programs, such as EPG-X (based on MATLAB, CPU) and PnP-MRF (based on C, CPU), snapMRF accelerates dictionary generation by $10 - 1000 \times$ and matching by $10 - 100 \times$. For dictionaries with small sizes, snapMRF can perform in a real-time manner. Besides, snapMRF provides the unit tests of 6 classic MR sequences, ensuring the correctness of dictionary generation and matching.

Keywords: MRI, Dynamic MRI, Compressed Sensing, Variational Models, Sparse Representation, MRF, Image Decomposition, Nuclear Norm, TGV, Quantitative Analysis, GPU

目录

摘要	I
Abstract	III
目录	IX
1 绪论	1
1.1 研究背景和意义	1
1.2 压缩感知的基本理论	5
1.2.1 稀疏性	5
1.2.2 采样矩阵	6
1.2.3 非线性重建算法	7
1.2.4 本节小结	10
1.3 基于压缩感知 MR 重建的研究现状	10
1.3.1 MR 图像的稀疏性	10
1.3.2 MR 采样模式	12
1.3.3 MR 图像的重建模型	15
1.3.4 动态 MR 图像的重建模型	22
1.3.5 本节小结	28
1.4 磁共振指纹的研究现状	28
1.4.1 信号获取	29
1.4.2 字典生成	30
1.4.3 参数图重建	31
1.4.4 本节小结	33
1.5 本文的主要研究内容和章节安排	33
2 背景知识	37
2.1 MRI 基本原理	37
2.1.1 MRI 的物理原理和数学模型	37
2.1.2 MR 成像序列	42
2.1.3 本节小结	44
2.2 扩展相图	44
2.2.1 配置状态矩阵	44

2.2.2 分区状态方法	45
2.2.3 其他因素的作用和最终模型	46
2.2.4 本节小结	47
2.3 本章总结	47
3 动态 MR 图像的低秩和稀疏分解的高阶变分模型和算法	49
3.1 模型的提出	51
3.1.1 低秩矩阵补全	51
3.1.2 二阶时空 TGV 泛函	52
3.1.3 基于二阶时空 TGV 和核范数的模型	53
3.2 模型的离散化和算法	53
3.2.1 离散化模型	53
3.2.2 Primal-Dual 算法及收敛条件	56
3.3 数值实验数据与评价方法	58
3.3.1 数值实验数据	58
3.3.2 采样模式与评价方法	58
3.4 数值实验结果	60
3.5 本章小结	69
4 胸部 DCE-MRI 压缩感知重建模型中时间稀疏正则项的量化评估	71
4.1 胸部 DCE-MRI 的重建模型	72
4.1.1 基于 Fourier 变换的重建模型	72
4.1.2 基于 Haar 小波变换的重建模型	72
4.1.3 基于 TV 的重建模型	73
4.1.4 基于二阶 TGV 的重建模型	73
4.1.5 基于核范数的重建模型	74
4.2 图像重建快速算法	74
4.2.1 FISTA 快速算法	74
4.2.2 模型参数的选择	75
4.3 数值实验数据与评价方法	76
4.3.1 数值实验数据	76
4.3.2 采样模式与评价方法	76
4.4 数值实验结果	77
4.4.1 重建图像质量	77
4.4.2 定量参数精度	78
4.5 数值实验分析	85
4.6 本章小结	85

5 基于 GPU 的实时 MRF 字典生成与参数图重建	87
5.1 MRF 重建参数图的主要问题	87
5.2 基于 GPU 的 MRF 重建	88
5.3 snapMRF 程序设计与单元测试	89
5.3.1 snapMRF 程序流程	89
5.3.2 snapMRF 核函数的功能	91
5.3.3 snapMRF 并行计算的难点与创新	92
5.3.4 单元测试	95
5.4 数值实验数据与评价方法	96
5.4.1 体模数据	96
5.4.2 活体人脑数据	97
5.5 数值实验结果	98
5.5.1 程序运行时间	98
5.5.2 体模数据的重建参数图	98
5.5.3 活体人脑数据的重建参数图	102
5.6 数值实验分析	105
5.7 本章小结	105
6 总结与展望	107
6.1 本文工作总结	107
6.2 下一步工作展望	108
致谢	125
附录	127

1 绪论

1.1 研究背景和意义

核磁共振成像^[1] (magnetic resonance imaging, 简称 MRI) 是医学临幊上常用的成像方法，其利用核磁共振现象来实现高对比度成像，可以非侵入式地获取人体内部组织的信息，被广泛地应用于指导病灶检测、诊断与治疗。MRI 在医学临幊和研究中有很多不同的分类方式，根据图像是否含有时间维度，可以将 MR 图像分为静态 MR 图像和动态 MR 图像 (dynamic MRI, 简称 dMRI)。常见的动态 MRI 有心脏电影成像 (cardiac cine)、心脏灌注成像 (cardiac perfusion)、磁共振动态对比增强 (dynamic contrast enhanced MRI, 简称 DCE-MRI)、功能核磁共振成像 (function MRI, 简称 fMRI) 等，如图1.1所示。根据研究方法的不同，可以将 MRI 分为定性 (qualitative) MRI

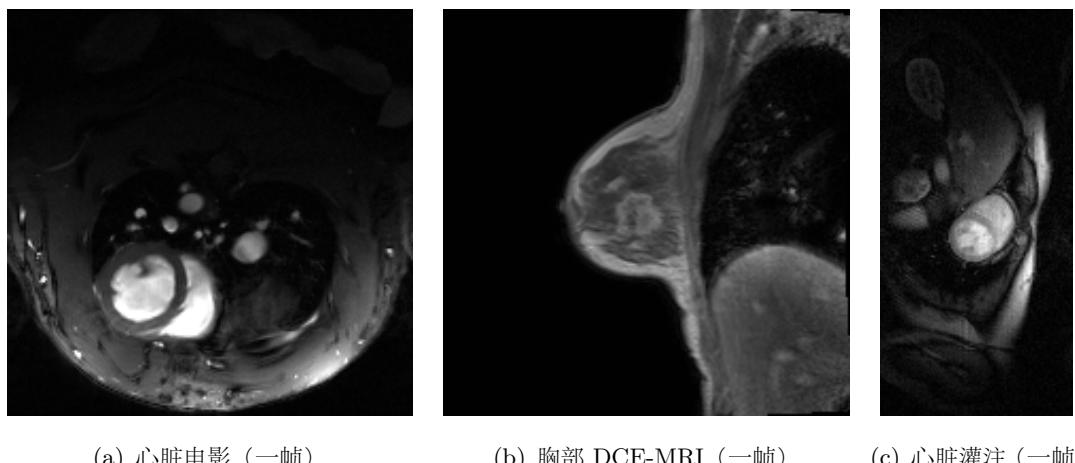


图 1.1 动态 MR 图像举例。

和定量 MRI (quantitative MRI, 简称 qMRI)，如图1.2所示。定性 MR 图像是医学临幊诊断中最常用到的类型，如 T_1 加权图像和 T_2 加权图像等，其对比度的来源分别为组织参数 T_1 和 T_2 的差异。但定性 MR 图像的像素值大小只是反映了 T_1 和 T_2 大小，并不是 T_1 和 T_2 本身。通过定性 MR 图像进行的诊断主要依靠医生的主观认识，缺乏定量指标。而定量 MRI 不仅能够提供组织的基本结构，也能获得组织的定量参数，从而减少了主观性，从客观定量的角度研究人体组织，达到帮助诊断与评估治疗的作用。DCE-MRI^[2] 是定量 MRI 的重要应用，被成功应用在胸部肿瘤定量分析以及脑灌注成像中。DCE-MRI 是通过注入对比剂引起的信号改变以评估组织灌注以及为血管通透性，常见的量化参数有体积转移常数 K^{trans} 和血管外细胞体积分数 v_e 等。我们将在第4章中介绍并讨论胸部 DCE-MRI 的重建模型和定量分析。磁共振指纹^[3] (magnetic resonance fingerprinting, 简称 MRF) 是近年来 qMRI 中的研究热点，如图1.3所示。MRF 可以在

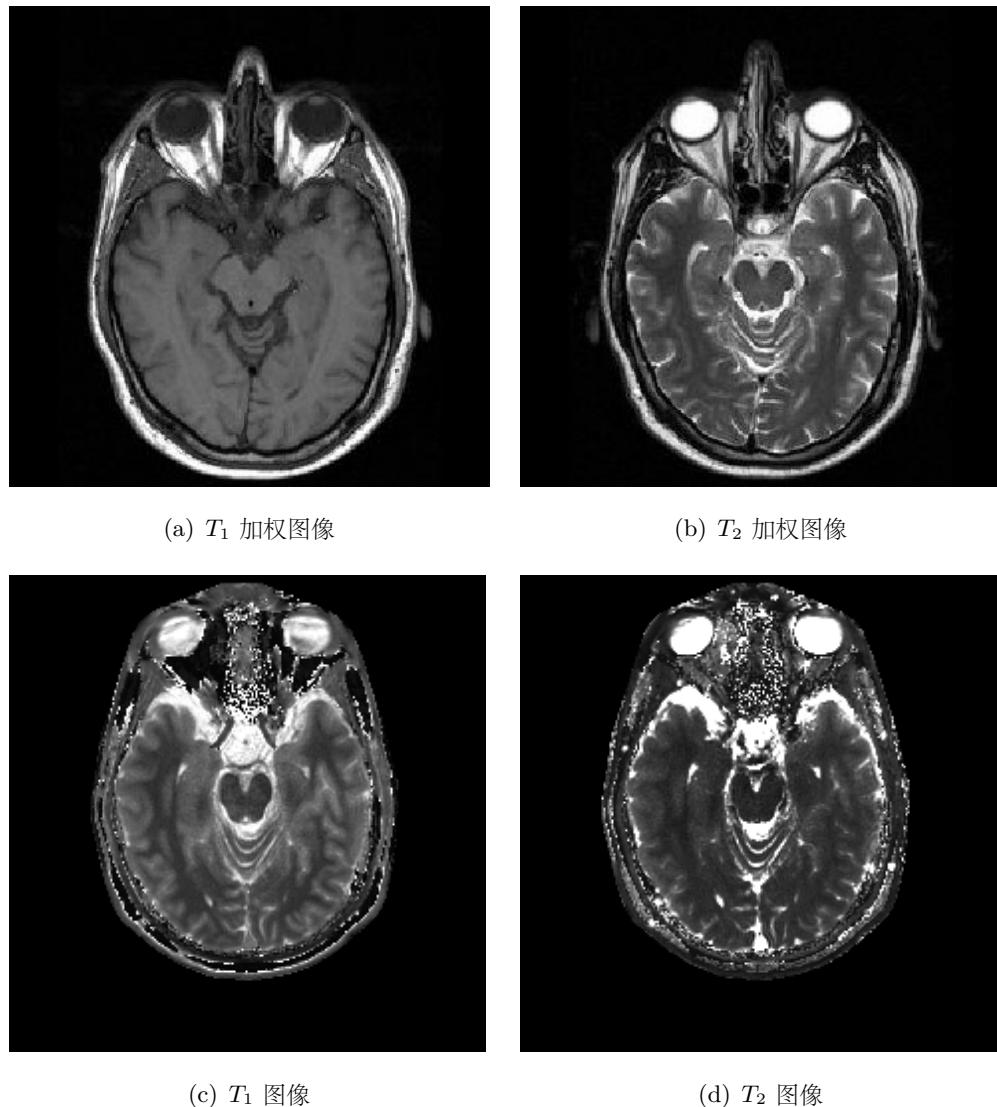


图 1.2 定性 MR 图像与定量 MR 图像。

单次采集中同时获得多个组织参数，并且重建的定量参数图有很高的信噪比。MRF 重建的过程主要包括预定定义字典的生成、信号采集和模式识别三个部分。具体来说，给定一个 MR 序列，首先使用数学模型生成一个包含不同参数的组织体素在该 MR 序列中随时间演化的字典，然后将采集到的信号与字典中的原子进行匹配，从而获得该体素的组织参数。我们将在第1.4小节介绍 MRF 的研究现状。

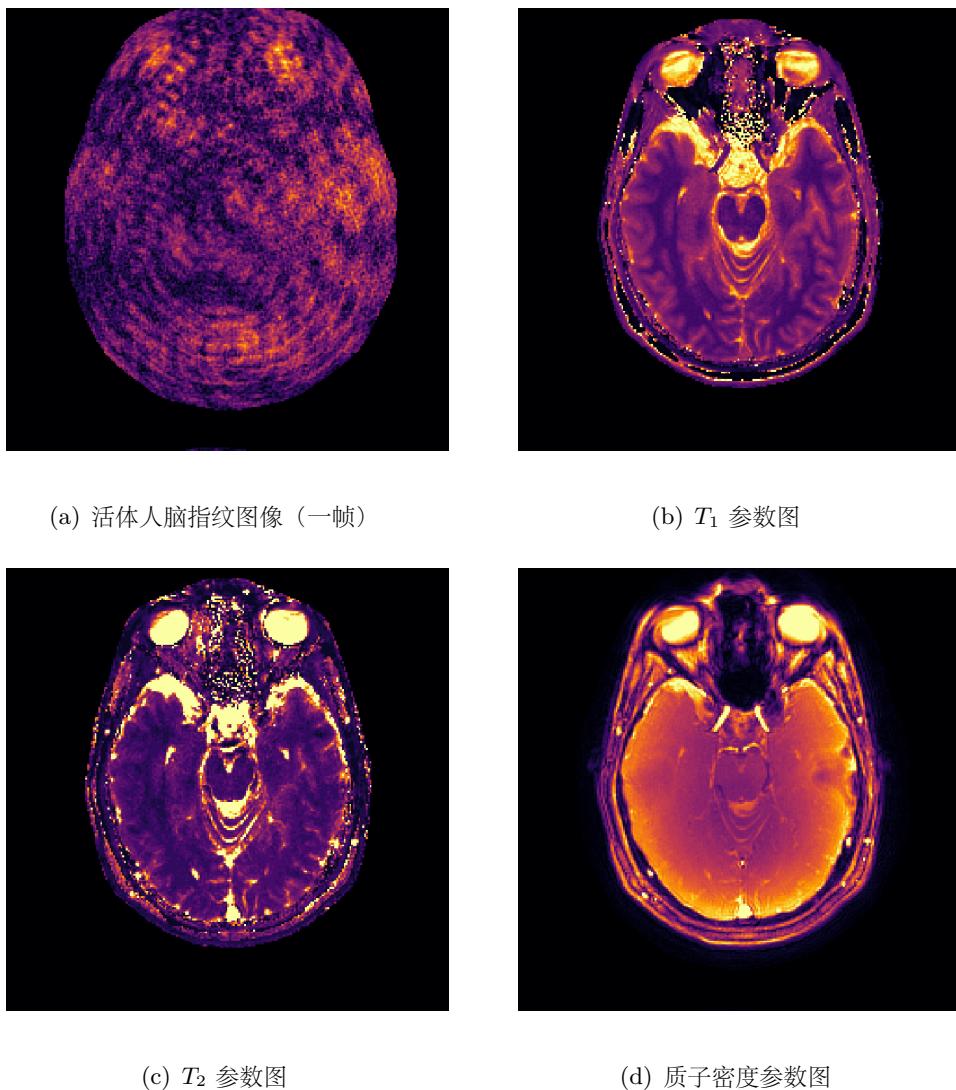


图 1.3 活体人脑 MRF 图像与重建参数图。

MR 成像过程通常被建模为二维或三维空间的 Fourier 编码，即首先从图像的 Fourier 域（通常称为 k-space）中采样，再通过算法将图像重建出来。详细的 MR 成像原理请参考第2章。为了精确地重建图像，采样率需要达到 Nyquist 采样定理^[4] 的要求，即当采样率达到信号最高频率两倍以上时，信号才可以被精确地重建出来。此外，由于受到物理和生理上的限制，MR 成像时需要在 k-space 中进行多次采样。所以，MR 成像速度通常很慢，尤其是 dMRI，这对成像设备和病人都是很大的负担。因此，如何在保证 MR 图像质量的前提下提高 MR 成像速度，一直是研究人员所关心的问题。目前，

加速 MR 成像通常的思路是减少采样数据，即在 k-space 空间中进行下采样，从而达到加速成像的效果。但由于在下采样时，采样率没有达到 Nyquist 采样定理的要求，通过简单的 Fourier 补零重建的图像（称为 Zero-filled 图像）会出现下采样伪影，如图1.4所示。例如，当对图像进行均匀下采样时，Zero-filled 图像中会出现相干伪影，使用传统的重建方法很难将其去除。压缩感知^[5-9] 是近些年来一种新兴的采样理论，可以加速信

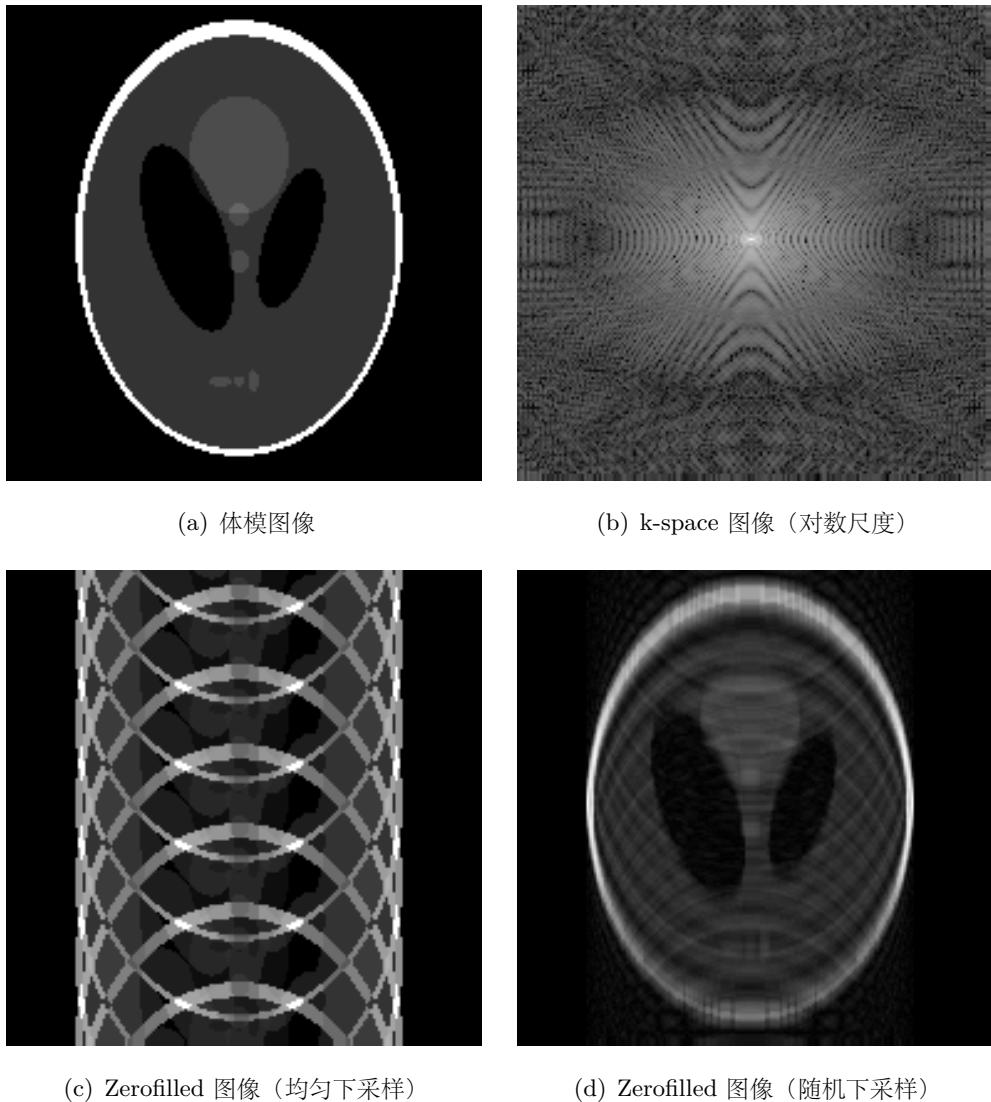


图 1.4 体模图像与 Zero-filled 图像。

号的采集与重建。压缩感知有三个主要部分，分别为稀疏性、非相干下采样和非线性重建算法。具体来说，首先假设信号在某个变换域内稀疏，对信号进行非相干下采样，然后通过非线性重建算法即可以高概率地将信号重建出来。压缩感知理论突破了 Nyquist 定理对高采样率要求的限制，可以用很少的数据来精确地重建信号。因此，压缩感知理论自诞生以来就被应用在 MR 成像中，可以显著地提高成像速度和质量，并逐渐成为快速 MR 成像的主流方法。本论文研究的重点之一即为压缩感知在 MR 成像中的应用，我们将在第1.2和1.3小节介绍基于压缩感知 MR 成像的研究现状。

1.2 压缩感知的基本理论

我们首先介绍压缩感知的基本理论和模型。压缩感知理论是 Candès, Romberg, Tao^[5-8] 和 Donoho^[9] 等于 2005 年提出的采样理论。数学上，压缩感知理论旨在求解以下模型：

$$b = Ax + \epsilon, \quad (1.1)$$

其中 $x \in \mathbb{C}^N$, $A \in \mathbb{C}^{M \times N}$, $b \in \mathbb{C}^M$, $\epsilon \in \mathbb{C}^M$, $M \ll N$ 。压缩感知的过程即是从 b 重建 x 的过程，在数学上被称为反问题。反问题在现实中有着广泛的应用，比如医学成像^[10-13]、计算机视觉^[14]、雷达成像^[15] 等。在 MR 成像中， A 通常被称作采样矩阵， x 为需要恢复的图像， b 是测量的到的 k-space 数据， ϵ 为测量噪声。由于 $M \ll N$ ，这个问题是病态的，所以我们需要对 x 有所约束。

1.2.1 稀疏性

压缩感知理论先验性地假设信号本身或者在某个变换域下是稀疏的。稀疏性是信号的固有属性，实际生活中的信号，如声音、图像、视频等一般在某个变化域下的系数都是稀疏的或者近似稀疏的。比如，分片常数的图像在梯度域（或称为有限差分域）中是稀疏的，而分片光滑的图像在小波变换域中是近似稀疏的。

在数学上，称向量 $x \in \mathbb{C}^N$ 为 s -稀疏的，如果

$$\|x\|_0 \leq s. \quad (1.2)$$

这里 $s < N$, $\|x\|_0$ 为 x 的 l_0 范数，其定义为

$$\|x\|_0 = |\text{supp}(x)| \quad \text{supp}(x) = \{j : x_j \neq 0\},$$

即 x 的 l_0 范数为 x 中非零元素的个数。对于 $s \in \{1, 2, \dots, N\}$ ，我们记

$$\Sigma_s = \{x \in \mathbb{C}^N : \|x\|_0 \leq s\}$$

为 s -稀疏向量的集合。进一步， $x \in \mathbb{C}^N$ 的最佳 s 项近似误差为：

$$\sigma_s(x)_p = \inf_{z \in \Sigma_s} \|x - z\|_p.$$

如果 $\sigma_s(x)$ 在 s 项内衰减很快，则称 x 为可压缩的。可以看出，可压缩向量可以由稀疏向量近似得到。这里向量 x 的 p 范数的定义为：

$$\|x\|_p = \left(\sum_{j=1}^N |x_n|^p \right)^{1/p}, \quad 0 < p < \infty. \quad (1.3)$$

在实际中，信号本身很少是真正稀疏的或者可压缩的，而是在某个变换域中才显示出稀疏性。常用的稀疏变换有 Fourier 变换、小波变换、紧框架、字典学习等。

1.2.2 采样矩阵

压缩感知重建模型旨在寻找模型 (1.1) 最稀疏的解。数学上, 这个过程可以描述为以下约束优化问题:

$$\begin{aligned}\hat{x} &= \arg \min_x \|x\|_0, \\ s.t. \quad &\|Ax - b\|_2^2 \leq \epsilon.\end{aligned}\tag{1.4}$$

我们通常称这个问题为 P0 问题。但是由于 (1.4) 非凸的, 而且是 NP 难的, 直接求解计算量很大。通常的求解方法有贪婪算法^[16,17]、重新加权范数算法^[18,19]、凸松弛法^[20]等。其中 l_1 范数凸松弛法是最常用的方法, 其旨在求解以下优化问题:

$$\begin{aligned}\hat{x} &= \arg \min_x \|x\|_1, \\ s.t. \quad &\|Ax - b\|_2^2 \leq \epsilon.\end{aligned}\tag{1.5}$$

l_1 优化问题也被称为基追踪 (basis pursuit^[21]) 或者 P1 问题。

为了保证压缩感知重建的精确性和鲁棒性, 采样矩阵 A 需要满足特定的性质。压缩感知理论目前给出的性质有三个, 它们分别为零空间性质 (null space property, 简称 NSP), 等距约束性 (restricted isometry property, 简称 RIP) 和非相干性 (incoherence)。其中零空间性质是 l_1 优化分析中的基本性质, 其定义为:

定义 1.2.2.1 称矩阵 $A \in \mathbb{C}^{M \times N}$ 满足 s -零空间性质, 如果

$$\|\eta_{\mathcal{O}}\|_1 \leq \gamma \|\eta_{\mathcal{O}^c}\|_1$$

对于所有集合 $\mathcal{O} \subset \{1, \dots, N\}$, $\#\mathcal{O} \leq s$ 和所有 $\eta \in \ker A = \{x | Ax = 0\}$ 均成立。其中, $\eta_{\mathcal{O}}$ 为下标与集合 \mathcal{O} 中元素一致的向量, \mathcal{O}^c 为集合 \mathcal{O} 的补集, $\#\mathcal{O}$ 为集合 \mathcal{O} 中元素的个数, $\gamma \in (0, 1)$ 为常数。以下定理保证了 P1 问题解的存在性。

定理 1.2.2.2^[22] 假设 $A \in \mathbb{C}^{M \times N}$ 满足 s -零空间性质。令 $x \in \mathbb{C}^N$, $b = Ax$ 并且 \hat{x} 是 P1 问题的解。那么如果 x 是 s -稀疏的, 则 $\hat{x} = x$ 。

可以看出, 零空间性质实际上等价于 l_1 稀疏重建。虽然 NSP 有着很好的性质, 但在实际中很难判断一个采样矩阵是否满足零空间性质。而等距约束性更加容易处理, 并且对噪声有鲁棒性。其定义如下:

定义 1.2.2.3 称采样矩阵 $A \in \mathbb{C}^{M \times N}$ 具有 s -等距约束性, 如果存在最小的常数 $0 \leq \delta_s < 1$ 使得

$$(1 - \delta_s)\|x\|_2^2 \leq \|Ax\|_2^2 \leq (1 + \delta_s)\|x\|_2^2\tag{1.6}$$

对所有 s -稀疏信号成立。

矩阵的等距约束性本质上要求矩阵 A 的所有 $M \times s$ 子矩阵接近等距, 即稀疏信号从高维空间向低维空间投影时的距离保持不变。从等距约束性可以推导出一些充分条件保证压缩感知可以求解 P1 问题。比如, 在没有噪声干扰的情况下, 如果 $\delta_{2s} < 1$, P0

问题有唯一的 s -稀疏解存在；如果 $\delta_{2s} < \sqrt{2} - 1$, P1 问题的解等价于 P0 问题的解。当有噪声存在时，以下定理保证了 P1 问题解的存在性。

定理 1.2.2.4^[22] 假设 $\delta_{2s} < \sqrt{2} - 1$, 则 P1 问题的解满足

$$\|\hat{x} - x\|_2 \leq c_0 s^{-1/2} \sigma_s(x)_1 + c_1 \epsilon,$$

其中 c_0 和 c_1 是与 δ_{2s} 有关的大于零的常数。

虽然等距约束性在理论上给出了采样矩阵 A 需要满足的条件，但是在实际应用中，等距约束性不容易计算，因此很难验证一个采样矩阵是否具有等距约束性。非相干性则给采样矩阵的验证提供了简单有效的方法，其定义为：

定义 1.2.2.5 矩阵 A 的列相干性为 A 中任意不同两列 a_i 与 a_j 内积的最大值：

$$\mu(A) = \max_{i \neq j} \frac{|\langle a_i, a_j \rangle|}{\|a_i\|_2 \|a_j\|_2}. \quad (1.7)$$

可以看出，矩阵列相干性的范围是 $\mu(A) \in [\sqrt{\frac{N-M}{M(N-1)}}, 1]$ 。

压缩感知理论表明，当采样矩阵 A 的列相干性接近于 1 时，压缩感知只需要 $O(s \log N)$ 次测量就可以很大概率恢复出原信号。

1.2.3 非线性重建算法

在实际计算中，我们一般将问题 (1.5) 转化成更容易处理的无约束问题求解，模型变为：

$$\hat{x} = \arg \min_x \|Ax - b\|_2^2 + \alpha \|x\|_1. \quad (1.8)$$

其中 α 是与噪声相关的权重参数。可以看出，问题 (1.8) 是凸的，但不可微。目前有很多快速算法来求解优化问题 (1.8)，有关算法的详细信息请参考文献^[23]。这些算法可以分为两类，一类是内点方法，另一类是一阶算法。一阶算法通常比内点算法计算更简便，也更适合处理大规模数据的情形。因此本文主要考虑一阶算法。

常用的一阶重建算法有向前向后分裂算法 (forward-backward splitting^[24], 简称 FBS), 分裂 Bregman 迭代算法 (split Bregman^[25]), 交替方向法 (alternating directional method of multiplier^[26-29], 简称 ADMM), Douglas-Rachford 分裂算法^[30], 快速阈值收缩算法 (fast iterative shrinkage threshold algorithm^[31], 简称 FISTA), 主对偶算法 (Primal-Dual^[32]) 等。其中 ADMM 算法最先被应用在压缩感知的重建中，也是最早被应用于 MR 压缩感知重建的算法之一。ADMM 算法首先将问题 (1.8) 转化为如下约束优化问题：

$$\begin{aligned} \hat{x} &= \arg \min_x \|Ax - b\|_2^2 + \alpha \|u\|_1, \\ s.t. \quad u &= x. \end{aligned} \quad (1.9)$$

问题 (1.9) 的增广 Lagrangian 形式为：

$$L_\mu(x, u, v) = \|Ax - b\|_2^2 + \alpha \|u\|_1 + \langle v, u - x \rangle + \frac{\mu}{2} \|u - x\|_2^2. \quad (1.10)$$

其中 v 为 Lagrangian 乘子, μ 为惩罚因子。通过交替方向法求解问题 (1.10), 其迭代步骤为:

$$\begin{cases} u^{n+1} = \arg \min_u L_\mu(x^n, u, v^n), \\ x^{n+1} = \arg \min_x L_\mu(x, u^{n+1}, v^n), \\ v^{n+1} = \arg \min_v L_\mu(x^{n+1}, u^{n+1}, v). \end{cases} \quad (1.11)$$

可以看出, l_1 重建的 ADMM 算法需要求解三个子问题。其中求解 u 子问题等价于求解

$$u^{n+1} = \arg \min_u \frac{\mu}{2} \|u - x^n + \frac{1}{\mu} v^n\|_2^2 + \alpha \|u\|_1,$$

可以使用软阈值收缩算法 (soft-shrinkage) 求解, 其定义为:

$$\mathcal{S}_\lambda(t) = \frac{t}{|t|} \max\{|t| - \lambda, 0\}.$$

取 $t = x^n - \frac{1}{\mu} y^n$, $\lambda = \frac{\alpha}{\mu}$, 则 u 子问题的解为:

$$u^{n+1} = \mathcal{S}_\lambda(t). \quad (1.12)$$

求解 x 子问题等价于求解

$$x^{n+1} = \arg \min_x \|Ax - b\|_2^2 + \frac{\mu}{2} \|v - x^n + \frac{1}{\mu} v^n\|_2^2, \quad (1.13)$$

可以通过共轭梯度算法^[33] (conjugate gradient) 进行求解。最后, 求解 v 子问题即为更新 Lagrangian 乘子:

$$v^{n+1} = v^n + \mu(u^{n+1} - x^{n+1}). \quad (1.14)$$

于是 l_1 稀疏重建的 ADMM 算法的详细流程如算法1所示。

算法 1 l_1 稀疏重建模型的 ADMM 算法

初始化: x^0, u^0, v^0 ;

迭代: 根据以下步骤更新参数:

1. 通过软阈值收缩算法求解 (1.12) 更新 u^{n+1} ;
2. 通过共轭梯度法求解 (1.13) 更新 x^{n+1} ;
3. $v^{n+1} = v^n + \mu(u^{n+1} - x^{n+1})$;

直到收敛, 返回 x^{n+1} 。

ADMM 算法是压缩感知的常用算法, 但一方面 ADMM 算法需要使用共轭梯度法求解子问题, 计算量大; 另一方面 ADMM 算法参数多, 调参的过程比较繁琐。FISTA^[31] 是也是压缩感知中常用算法, 收敛速度快, 而且需要调节的参数比 ADMM 少很多。FISTA 算法考虑极小化下面的问题:

$$\min_x f(x) + g(x), x \in \mathbb{C}^N, \quad (1.15)$$

其中 $f(x)$ 是凸的光滑函数，其 Lipschitz 常数为 L_f , $g(x)$ 为凸函数，可能是非光滑的。定义函数 $g(x)$ 的临近点算子为：

$$\text{prox}_\tau(g)(x) = \arg \min_u g(u) + \frac{1}{2\tau} \|u - x\|_2^2. \quad (1.16)$$

则 FISTA 的算法流程如算法2所示。

算法 2 FISTA 算法迭代流程

初始化： $\tau = 1/L_f, t^1 = 1, x^0 = r^1$;

迭代：根据以下步骤更新参数：

1. $\bar{x} = r^k - \tau \nabla f(x^n)$;
2. $x^n = \text{prox}_\tau(g)(\bar{x})$;
3. $t^{n+1} = (1 + \sqrt{1 + 4(t^n)^2})/2$;
4. $r^{n+1} = x^n + \frac{t^n - 1}{t^{n+1}}(x^n - x^{n-1})$;

直到收敛，返回 x^{n+1} 。

将 FISTA 算法应用到问题 (1.8) 中，令 $f(x) = \|Ax - b\|_2^2$, $g(x) = \alpha\|x\|_1$ ，则 l_1 稀疏重建的 FISTA 算法流程如算法3所示。FISTA 算法是压缩感知重建中常用的算子

算法 3 l_1 稀疏重建的 FISTA 算法

初始化： $\tau = 1/L_f, t^1 = 1, x^0 = r^1$;

迭代：根据以下步骤更新参数：

1. $\bar{x} = r^n - 2\tau A^*(Ax - b)$;
2. $x^n = \text{prox}_\tau(g)(\bar{x})$;
3. $t^{n+1} = (1 + \sqrt{1 + 4(t^n)^2})/2$;
4. $r^{n+1} = x^n + \frac{t^n - 1}{t^{n+1}}(x^n - x^{n-1})$;

直到收敛，返回 x^{n+1} 。

分裂算法之一，被广泛应用于信号处理^[34] 和多任务学习^[35] 中。但是 FISTA 只能解决比较简单的模型，不能应用在含有多个正则项的模型中。我们将在第4章中重点讨论和应用 FISTA 算法。

Primal-Dual^[32] 算法是近年来比较热门的一阶算法，收敛速度快，并且算法的收敛有着严格的定理保证。Primal-Dual 算法被广泛应用于凸-凹鞍点问题的极大极小问题中：

$$\min_{x \in \mathcal{X}} \max_{y \in \mathcal{Y}} \langle \mathcal{K}x, y \rangle + f(x) - g(y), \quad (1.17)$$

其中 \mathcal{X} 和 \mathcal{Y} 是 Hilbert 空间, $\mathcal{K} : \mathcal{X} \rightarrow \mathcal{Y}$ 是线性连续映射，并且泛函 $f : \mathcal{X} \rightarrow (-\infty, \infty]$ 和 $g : \mathcal{Y} \rightarrow (-\infty, \infty]$ 是适定的、凸的和下半连续的。这里我们仅给出 l_1 稀疏优化问题 (1.8) 的 Primal-Dual 算法的迭代步骤，如算法4所示。具体的推导过程和详细信息请参考第3章。

算法 4 l_1 稀疏重建的 Primal-Dual 算法

初始化： 选择 $\tau, \sigma > 0$, x^0, y^0, p^0 , 令 $\bar{x}^0 = x^0$;

迭代： 根据以下步骤更新参数:

1. $p^{n+1} = \mathcal{P}_\alpha(p^n + \sigma \bar{x}^n)$;
2. $r^{n+1} = [r^n + \sigma(A\bar{x}^n - b)]/(1 + \sigma)$;
3. $x^{n+1} = x^n - \tau A^* p^n$;
4. $\bar{x}^{n+1} = 2x^{n+1} - x^n$;

直到收敛，返回 x^{n+1} 。

1.2.4 本节小结

在本节中，我们回顾了压缩感知的数学模型和理论基础。压缩感知有三个主要部分，分别为稀疏性、采样矩阵和非线性重建算法。其中稀疏性是先验信息，采样矩阵保证了压缩感知重建模型解的存在性，而非线性重建算法决定了重建算法的速度和精度。

1.3 基于压缩感知 MR 重建的研究现状

MR 成像是从 Fourier 域（也称为 k-space）中采样并重建图像的过程。由于 MR 成像的固有时间长，如何加快成像速度一直是研究人员关心的话题。压缩感知理论为从下采样数据中精确重建出 MR 图像提供了理论基础和算法框架。压缩感知理论表明，假设 MR 图像在某个变换域稀疏的前提下，当采样矩阵满足等距约束性或者非相干性条件时^[9]，利用非线性重建算法可以大概率将图像精确地重建出来。压缩感知理论和 MR 成像有着天然的契合。为了使得压缩感知理论能够应用于 MR 成像，我们首先需要为 MR 图像找到合适的稀疏表示，并且采样矩阵也需要跟 MR 的成像域，即 Fourier 域非相干。在这一节中，我们重点回顾压缩感知在 MR 成像中的应用，MR 成像原理请见第2章。

1.3.1 MR 图像的稀疏性

一般来说，医学图像是分片光滑的，因此在合适的稀疏变化下，医学图像通常展现出稀疏性或可压缩性。对于 MR 图像而言，其稀疏性或者说冗余性主要体现在三个方面。

第一，空间上的冗余性。自然图像通常在小波变换域或者有限差分域是稀疏的，因此小波变换和 TV 泛函被广泛应用与图像去噪和重建中。受此启发，Lustig^[13] 通过数值实验验证了二维 MR 图像本身很少是稀疏的，但是其在小波变换域或者有限差分域上是稀疏的。图1.5展示了体模图像在小波变换和有限差分下的稀疏性。可以看出，体模图像本身并不是稀疏的，但在小波变换和有限差分域下均体现出稀疏性，因此它们是基于压缩感知的 MR 重建模型中最常用的稀疏变换。文献^[13] 是从压缩感知理论的角度研究 MR 图像稀疏性的第一篇文献，之后的文献均从这个想法出发，根据具体问题提出

了不同的稀疏变换，比如 TGV^[36]、Shearlet^[37]、字典学习^[38] 等。

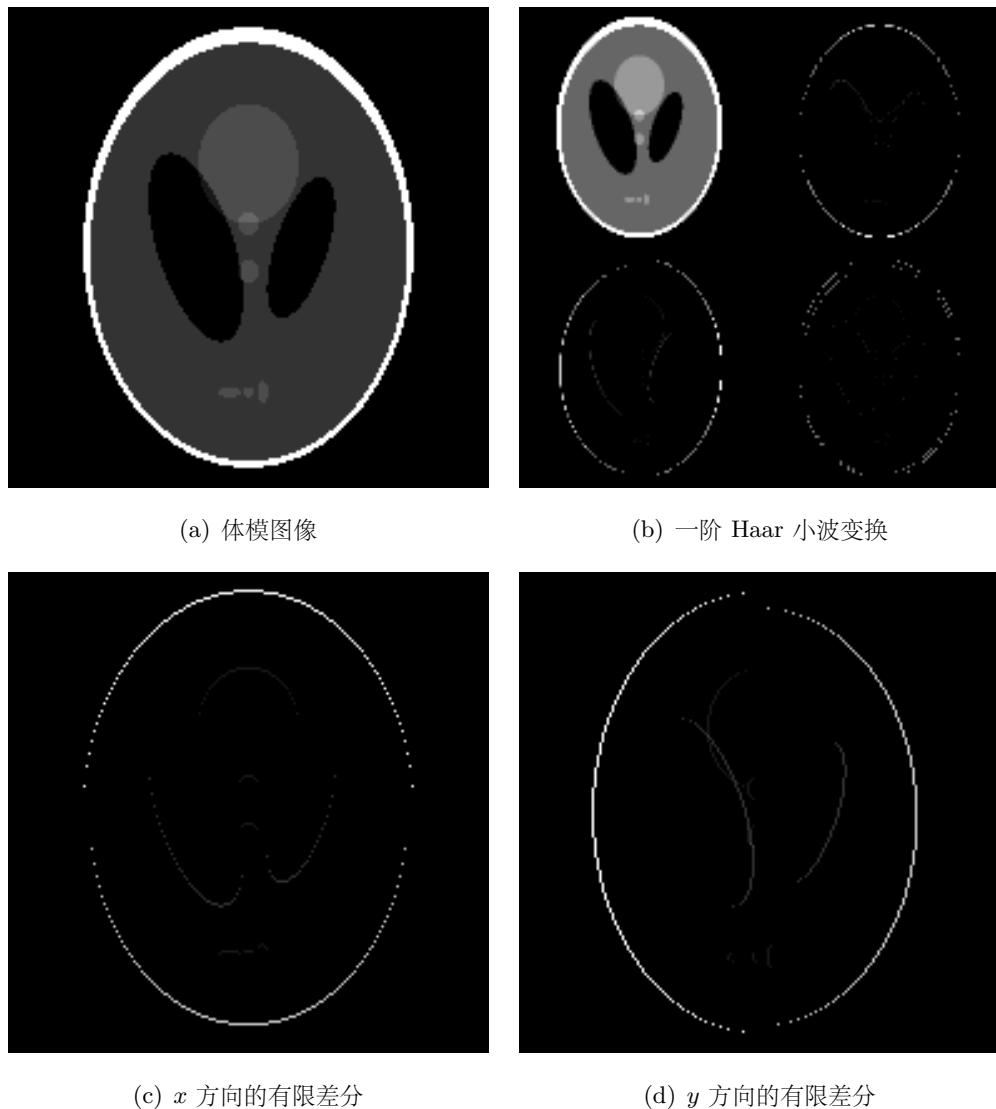


图 1.5 体模图像稀疏性举例。

第二，时间上的冗余性。对于动态 MR 图像而言，其稀疏性不仅体现在空间方向上，也体现在时间方向上。动态 MRI 是 MRI 的重要应用之一，主要的应用有心脏电影成像、磁共振动态对比增强、功能磁共振成像等。动态 MR 常用的稀疏表示有 Fourier 变换、核范数、有限差分、字典学习等。比如，Lustig^[13] 将一维 Fourier 变换作为时间方向上的稀疏变换，并将模型应用在心脏电影成像中。其想法是，心脏做着周期性的运动，而 Fourier 变换可以稀疏化具有周期性的信号。但是在实际应用中，心脏的运动并不具有完美的周期性，Fourier 变换不是最优的选择。因此，基于数据的方法，比如主成分分析、字典学习等可能是更好的选择。简而言之，对于动态 MR 图像而言，其稀疏变换的选择是十分重要而且困难的。我们将在第1.3.4小节详细介绍动态 MR 图像的压缩感知模型。

第三，线圈上的冗余性。平行成像是 MRI 中的重要应用，也是最早与压缩感知理

论结合的应用之一。平行成像使用多组线圈对同一目标从不同的空间位置进行数据采集，再通过算法将不同线圈的数据合成为最终的图像。由于每个线圈对目标的空间感受性（sensitivity）不同，线圈所采集到的数据之间有着很大的冗余性，我们称之为线圈上的冗余性。平行成像的常用方法有 SENSE^[39]，SMASH^[40]，GRAPPA^[41] 等。其中 SENSE 利用了图像空间域的冗余性，而 GRAPPA 利用了 Fourier 域的信息。压缩感知与平行成像相结合是目前 MR 成像领域研究的热点之一。在静态成像上，Block 等^[42] 将全变差的思想应用到平行成像上，而 Liang^[43] 等则给出了更一般形式的 SENSE。在动态成像上，文献^[44,45] 最早结合了平行成像与动态 MR 成像。

综上所述，MR 图像的稀疏性主要体现在空间、时间和线圈上。因此需要根据具体问题，在不同的维度上选择不同的稀疏变换。

1.3.2 MR 采样模式

在压缩感知理论中，采样模式起着关键的作用。一般来说，采样矩阵需要满足等距约束性或者非相干性。对于 MRI 而言，其采样模式也受到物理方面的限制。从第2.1.1节中我们知道，MRI 的采样模式或采样轨迹是通过调整梯度场的大小和方向来实现的，因此在实际中我们希望采样轨迹是光滑的。所以，虽然随机矩阵在理论上可以保证足够的非相干性，但是无法应用到 MR 成像中。在一般的二维 MR 成像中，采样的策略通常是只保证一个方向的自由度，即在数据读取方向（readout）方向上进行全采样，而在相位编码（phase encoding）方向上进行随机采样，以增加采样矩阵的非相干性。

压缩感知 MR 重建有以下四种常见的采样模式，如图1.6所示。其中图1.6（a）是 Cartesian 采样模式，被广泛地应用于医学临床中。由于其采样点均在 Cartesian 网格点上，重建过程十分简便，通过快速 Fourier 变换与逆变换即可实现图像域和频率域的转化。在图1.6（a）中，k-space 的中间部分进行了全采样，而外围部分则随机选取了几条采样线。这是因为 k-space 的中心区域是 MR 图像的低频部分，保留着图像的卡通信息；而外围是 MR 图像的高频部分，保留着图像的轮廓信息和噪声。虽然 Cartesian 采样计算简便，但其能达到的非相干性很低，这限制了它在压缩感知重建中的表现。

图1.6（b）和（c）是两种经典的 non-Cartesian 采样模式，分别为径向采样^[42,46] 和螺旋采样^[47]。与 Cartesian 采样模式相比，这两种采样模式的非相干性更高，因此可以达到更高的采样率。其中径向采样在动态 MR 成像中的应用更加广泛，因为其可以在多个维度上产生非相干伪影。此外，径向采样也对运动不敏感，可以捕捉到更多的动态信息^[48]。螺旋采样所需要的采样线的个数最少，当采样时间足够长时，甚至只用一条采样线即可覆盖整个 k-space。但是 non-Cartesian 采样的采样点不在 Cartesian 网格上，一般需要进行差值操作，重建相对繁琐耗时。

图1.6（d）是变化密度采样（variable density sampling），其在两个方向上均为随机采样，整体的采样模式服从二维高斯分布。这种采样模式可以达到最高的非相干性，在 MR 重建的数值试验中表现最好，但在实际应用中，变化密度采样几乎无法实现，一般只用于数值模拟。

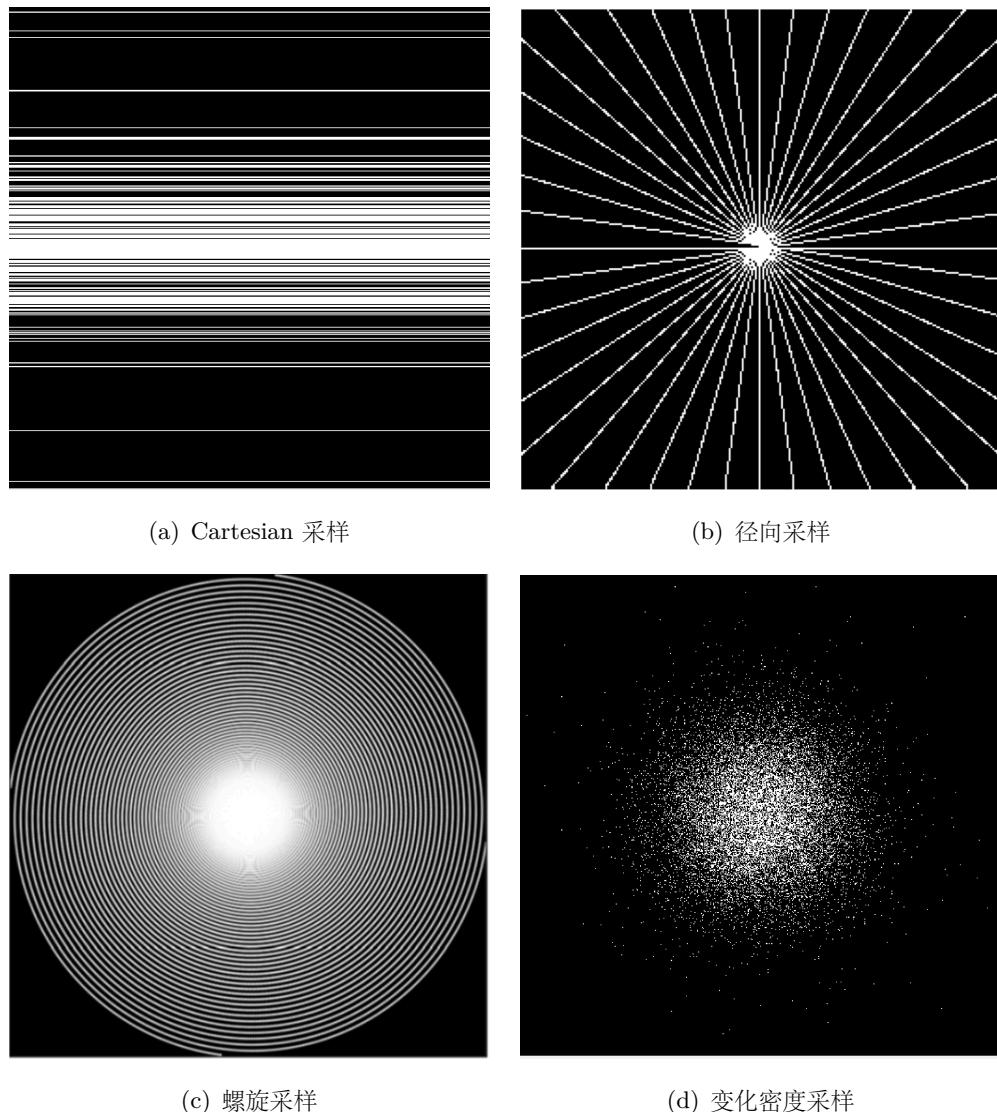


图 1.6 MRI 中常用的采样模式

如何设计压缩感知 MR 重建的采样模式一直是研究人员所关心的话题，也决定着重建图像质量的好坏和重建算法的精度。设计的思路之一是将以上四种采样模式进行组合。比如，对于二维 MR 成像，伪径向采样^[49]（pseudo radial sampling）也是常见的采样模式，如图1.7（a）所示。伪径向采样结合了 Cartesian 和径向采样的优点，其采样点均在 Cartesian 网格上，同时使得采样点尽量沿径向方向。这样既可以获得高的采样率，又简化了计算。对于动态 MR 成像或者三维 MR 成像，由于其冗余性体现在多个维度上，采样矩阵的设计有了更大的空间。文献^[50] 针对胸部 DCE-MRI 图像，结合了 Cartesian 采样和随机采样，如图1.7（b）所示。具体来说，在每一帧图像上使用 Cartesian 采样，并使得帧与帧之间采样有最大的正交性。文献^[49] 中针对动态 MR 图像，结合了伪径向采样和随机采样，如图1.7（a）所示。具体来说，在每一帧图像上均匀地选取一些伪径向采样线，并使得帧与帧之间的采样线随机转动一个角度以增加非相干性，即在空间上采用伪径向采样，在时间上采用随机采样，这也是动态 MR 采样模式的设计思路之一。文献^[51] 针对 fMRI 图像，采用了变化密度螺旋采样模式。具体来说，这种采样方式使得 k-space 中心区域可以达到 Nyquist 采样率，而 k-space 的外围则是下采样，从而增加了采样的非相干性。除了 fMRI，变化密度螺旋采样也被应用在心脏电影成像^[52]、扩散张量成像^[53]、心脏血管造影成像^[54] 中。

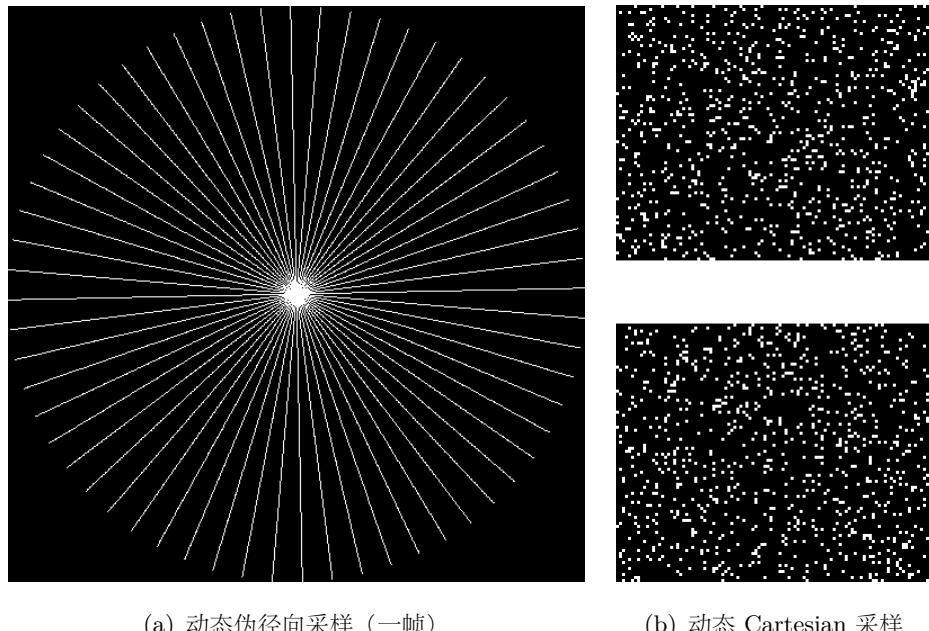


图 1.7 动态 MRI 中常用的采样模式。

综上所述，压缩感知 MR 重建的采样矩阵既需要满足非相干性，也需要考虑 MRI 的物理限制。对于二维 MR 成像，主要的采样模式有 Cartesian 采样、径向采样和螺旋采样；而对于动态 MR 成像和三维 MR 成像，冗余性体现在多个维度上，采样矩阵的设计有着更多的自由度。一般的思路是根据具体问题，在不同的维度上选择不同的采样模式。

1.3.3 MR 图像的重建模型

在这一小节，我们介绍压缩感知 MR 成像的经典模型，更详细的综述请参考文献^[55,56]。我们知道，MR 的数据获取发生在 k-space，即 Fourier 域中。于是 MR 成像的采样矩阵 $A \in \mathbb{C}^{M \times N}$, $M \ll N$ 可以写成

$$A = \mathcal{M} \cdot \mathcal{F} \quad (1.18)$$

其中 \mathcal{F} 为 Fourier 变换， \mathcal{M} 为采样模式。则 MR 成像的数据采集过程可以建模为：

$$b = Ax + \epsilon. \quad (1.19)$$

压缩感知理论旨在寻找最稀疏的解，于是 MR 成像的压缩感知模型为：

$$\begin{aligned} \hat{x} &= \arg \min_x \|\mathcal{S}x\|_1, \\ s.t. \quad &\|Ax - b\|_2^2 \leq \epsilon. \end{aligned} \quad (1.20)$$

这里 $x \in \mathbb{C}^N$ 为待重建图像， $b \in \mathbb{C}^M$ 为测量得到的 k-space 数据， \mathcal{S} 为某个稀疏变换。将问题 (1.20) 写成无约束优化问题为：

$$\hat{x} = \arg \min_x \|Ax - b\|_2^2 + \alpha \|\mathcal{S}x\|_1. \quad (1.21)$$

目前稀疏变换 \mathcal{S} 的选取方法有很多。根据稀疏变换选取方式的不同，可以把模型分为两大类，即基于模型 (model-based) 的方法和基于数据 (data-based) 的方法。

最早将压缩感知应用 MR 成像的是 Lustig^[13]。在文章中，作者研究了适用于 MR 的稀疏表示，讨论了不同采样模式非相干性的大小，并提出了基于小波变换和全变分的 MR 重建模型：

$$\min_x \frac{1}{2} \|Ax - b\|_2^2 + \alpha \|\mathcal{W}x\|_1 + \beta \|x\|_{\text{TV}}. \quad (1.22)$$

其中 \mathcal{W} 为小波变换， $\|\cdot\|_{\text{TV}}$ 定义为：

$$\|\cdot\|_{\text{TV}} = \|\nabla \cdot\|_1,$$

而

$$\|\nabla \cdot\|_1 = \sqrt{(\nabla_x \cdot)^2 + (\nabla_y \cdot)^2},$$

其中 ∇_x 和 ∇_y 分别为 x 和 y 方向上的梯度算子。

Lustig 使用了伪径向采样方式，并且利用共轭梯度法来求解模型 (1.22)，因此算法速度慢。针对模型 (1.22)，很多文章在重建算法上做了改进。Ma 等在 2008 年提出了一种算子分裂算法 TVCMRI^[57]，2009 年 Yang 等提出了基于交替方向方法的重建算法 RecRF^[58]。这两种方法都加快了算法的收敛速度，提高了重建图像的信噪比。2011 年 Huang^[59] 等提出了结合算子分裂和变量分裂的算法 FCSA (fast composite shrinkage algorithm)，解决了 FISTA 只能处理一个正则项的瓶颈。

由于模型 (1.22) 中含有 TV 正则项, 重建得到的图像往往会产生阶梯效应。为了解决这个问题, Bredies 等^[60] 提出了利用高阶梯度项来刻画图像的 TGV 泛函, 并将其应用在图像去噪^[61] 和压缩感知^[36] 中。在实际应用中, 一般使用二阶 TGV, 即 TGV_α^2 , 其定义为:

$$\text{TGV}_\alpha^2(x) = \arg \min_v \alpha_1 \|\nabla x - v\|_1 + \alpha_0 \|\mathcal{E}(v)\|_1.$$

这里 $\nabla \cdot = (\nabla_x \cdot, \nabla_y \cdot)$ 一阶梯度算子, v 属于离散复值向量场 $V \in \mathbb{C}^{2N}$, $\mathcal{E}(v) = (\nabla v + \nabla v^T)/2$ 是二阶对称导数算子。关于 TGV 的详细概念和性质请参考第3章。于是基于 TGV 的压缩感知重建模型为:

$$\min_x \frac{1}{2} \|Ax - b\|_2^2 + \text{TGV}_\alpha^2(x). \quad (1.23)$$

模型 (1.23) 使用 Primal-Dual 算法求解, 有关模型求解的详细推导过程和收敛性分析请参考第3章。基于二阶 TGV 的压缩感知模型可以在保持图像边缘的同时有效地抑制阶梯效应, 并且提高重建图像的信噪比。

模型 (1.22) 中使用的小波变换为 Haar 小波变换^[62], 其计算简便。在此之后, 很多文献将 Haar 小波推广到了小波家族中其他更加复杂的变换, 即多尺度几何分析。Curvelet^[63] 和 Contourlet^[64] 是多尺度几何分析中最早被提出的方法, 可以很好的稀疏化表示图像中的曲线状特征, 而传统的小波变换只能表示图像中的点状特征。Shearlet^[37] 变换具有对高维几何结构稀疏表示的能力, 对于卡通类图像的最优稀疏近似和 Curvelet 相同, 并且可以表示图像中边缘类纹理。在压缩感知的模型中加入多尺度几何分析, 可以更好地恢复出图像中的高维几何信息和纹理信息。比如文献^[65,66] 将模型 (1.22) 中的小波变换 \mathcal{W} 换成了 Curvelet 变换和 Contourlet 变换, 得到了以下两个模型:

$$\min_x \frac{1}{2} \|Ax - b\|_2^2 + \alpha \|\mathcal{CU}x\|_1 + \beta \|x\|_{\text{TV}} \quad (1.24)$$

和

$$\min_x \frac{1}{2} \|Ax - b\|_2^2 + \alpha \|\mathcal{CO}x\|_1 + \beta \|x\|_{\text{TV}}, \quad (1.25)$$

其中 \mathcal{CU} 和 \mathcal{CO} 分别为 Curvelet 变换和 Contourlet 变换。数值实验表明, 这两个模型可以很好地恢复出图像中曲线状的几何信息。Gao 等^[61] 基于图像分解的思想, 提出了利用了二阶 TGV 和 Shearlet 变换的变分模型:

$$\min_{x_1, x_2} \frac{1}{2} \|A(x_1 + x_2) - b\|_2^2 + \text{TGV}_{\alpha, \beta}(x_1) + \gamma \|\mathcal{SH}(x_2)\|_1, \quad (1.26)$$

其中 \mathcal{SH} 为 Shearlet 变换, x_1 和 x_2 分别代表图像的卡通部分和纹理部分。模型 (1.26) 可以应用在多种图像处理的任务中, 如图像去噪、图像去糊、压缩感知重建等。当用于压缩感知重建时, 模型 (1.26) 可以很好地恢复出图像的卡通部分和纹理部分。

除了利用小波变换和梯度信息, 近些年有一些方法利用了结构化稀疏 (structured sparsity) 的概念。结构化稀疏是一般稀疏正则化方法的推广和延伸, 其旨在选择特殊

结构的稀疏性，比如图（graph/group）或者网络（network）。对于 MR 图像而言，其稀疏表示的稀疏往往具有某种特殊结构，而利用这种结构可以增加图像的稀疏性。寻找小波变换系数的结构是结构化稀疏中常用的一类方法。Chen 等提出了一个基于小波树稀疏的压缩感知模型 WaTMRI^[67,68]：

$$\min_x \frac{1}{2} \|Ax - b\|_2^2 + \alpha \|x\|_{\text{TV}} + \beta (\|\mathcal{W}x\|_1 + \sum_{o \in \mathcal{O}} \|\mathcal{W}x_o\|_2). \quad (1.27)$$

这里 \mathcal{O} 表示所有母子系数组的集合， o 为 \mathcal{O} 的一个组， x_o 是属于母子系数组的所有数据。利用小波树进行压缩感知重建可以有效地消除边缘伪影。结构化稀疏的另一大类方法是基于图像非局部（nonlocal）^[69-71] 信息的方法，并且已经被应用于 MR 成像模型中，其表现比一般的正则项要好。非局部的方法的基本思想是将有相似结构特征的图像块（patch）聚合在一起进行处理，这样可以更好地利用图像的冗余性。Mehmet 等^[72] 基于非局部的思想提出了低维结构自学习和阈值算法（low-dimensional-structure self-learning and thresholding，简称 LOST），主要利用了下采样图像中的结构信息。算法首先使用 k-space 中心低分辨率图像学习得到具有相似解剖结构的图像块，然后将这些图像块通过某种相似性聚类算法整合在一起，这样做可以有效地去除重建图像中的空间伪影。Liang 等^[73] 研究了非局部 TV（NLTВ）正则化在 SENSE 重建中的应用。NLTВ 范数定义如下：

$$\|x\|_{\text{NLTВ}} = \|\nabla^w x\|_{2,1},$$

其中 $\|\cdot\|_{2,1}$ 是矩阵范数，即先对矩阵的每一列求 l_2 范数，再对所得结果求 l_1 范数。 $\nabla^w x$ 是一个 $N \times N$ 矩阵，表示加权非局部图像的梯度，其定义为：

$$\nabla^w x = [\nabla_{i,j}^w x],$$

其中 $\nabla_{i,j}^w x$ 是像素点 i 与其非局部相邻像素点 j 之间的加权图像梯度：

$$\nabla_{i,j}^w x = \sqrt{w(i,j)}(x(i) - x(j)),$$

而 $w(i,j)$ 是像素值 $x(i)$ 与 $x(j)$ 之间的图加权函数。将 NLTВ 范数应用到 MR 重建中，其模型为：

$$\hat{x} = \arg \min_x \frac{1}{2} \|Ax - b\|_2^2 + \alpha \|x\|_{\text{NLTВ}}. \quad (1.28)$$

数值实验表明，NLTВ 模型可以很好地保持图像信息和细小的结构，并抑制噪声。Liu^[74] 等结合 TV 和 NLTВ，提出了基于图像块的 JCTV 模型：

$$\begin{aligned} \{\hat{x}, \hat{z}_i\} = & \arg \min_{x, z_i} \frac{1}{2} \|Ax - b\|_2^2 + \sum_i \mu_{i,1} \|z_i\|_1 \\ & + \sum_i \mu_{i,2} \|z_i - \gamma_i\|_1, \\ s.t. \quad & z_i = \nabla \mathcal{P}_i x, \end{aligned} \quad (1.29)$$

其中 \mathcal{P}_i 算子表示从图像 x 中提取图像块 x_i , z_i 是图像块在有限差分域中稀疏表示的系数, γ_i 是真实图像的图像块在有限差分域中稀疏表示的系数, $\mu_{i,1}$ 和 $\mu_{i,2}$ 是正则化参数。Qu 等^[75] 提出了基于图像块的非局部算子 (patch-based nonlocal operator, 简称 PANO) 来对图像进行稀疏表示, 并将其应用到了压缩感知成像中。同样令 \mathcal{P}_i 为块分解算子, 图像的第 i 个块为 x_i , 则有 $x_i = \mathcal{P}_i x$ 。将图像块按相似度分为 J 组, 第 j 组的图像块记作 $\mathcal{O}_{v_j} x_i$, 其中 v_j 中为该组图像块的指标集。则非局部 PANO 算子定义为:

$$\mathcal{PANO}_j = \mathcal{SO}_{v_j} \mathcal{P}_i.$$

将 PANO 算子应用到压缩感知的模型中:

$$\hat{x} = \arg \min_x \frac{1}{2} \|Ax - b\|_2^2 + \alpha \sum_{j=1}^J \|\mathcal{SO}_{v_j} \mathcal{P}_i x\|_1. \quad (1.30)$$

PANO 算子可以稀疏化表示同一相似度的图像块, 在 MR 重建中也取得了很好地效果。BM3D (block matching 3D)^[76] 也是通过将图像块分组的方法来处理图像非局部结构信息的模型, 并被成功应用于图像去噪、去糊、超分辨率重建等任务中。Ender^[77] 首次将 BM3D 模型应用到了 MR 重建中, 并提出了模型 BM3D-MRI:

$$\hat{x} = \arg \min_x \frac{1}{2} \|Ax - b\|_2^2 + \alpha \|\Phi x\|_1, \quad (1.31)$$

其中 $\Phi \in \mathbb{C}^{K \times N}$ 为 BM3D 的分析框架, 且 $K >> N$, 即框架 Φ 是过完全的。文章使用了解耦算法来求解模型, 并给出了算法收敛的条件。非局部 BM3D-MRI 是 MRI 重建模型的很有前景的应用。注意, 利用结构化稀疏的模型中, 有些也使用了基于数据的方法, 比如 PANO。因此称它们为部分基于模型的方法更为准确。

到目前为止, 我们讨论的模型都属于基于模型的方法或者部分基于模型的方法。而压缩感知 MR 重建另一大类方法是学习图像的稀疏表示, 即基于数据的 (data-based) 方法, 或者称为盲压缩感知^[78] (blind CS)。2011 年, Ravishankar 等提出了基于字典学习的压缩感知模型 DL-MRI^[79], 这也是最早将学习的思想应用到 MR 重建中的模型。模型如下:

$$\begin{aligned} & \min_{x, \mathcal{D}, z} \frac{1}{2} \|Ax - b\|_2^2 + \alpha \sum_{j=1}^N \|\mathcal{P}_j x - \mathcal{D} z_j\|_2^2, \\ & s.t. \|z_j\|_0 \leq s, \|d_i\|_2 = 1, \forall i, j. \end{aligned} \quad (1.32)$$

这里 \mathcal{P}_j 算子同样为块分解算子, 即 $x_j = \mathcal{P}_j x$; $\mathcal{D} \in \mathbb{C}^{K \times N}$ 是待学习字典, K 为字典中原子的个数, z_j 是块 x_j 相对于字典 \mathcal{D} 的稀疏表示, 即 $z_j = \mathcal{D} x_j$ 。模型的求解分为两个子步骤, 即字典学习和图像重建。模型通过交替方向法求解, 在字典学习和图像重建之间迭代, 直到收敛。其中字典学习的子步骤是在已知图像 x 的情况下, 求解字典 \mathcal{D}

并计算图像 x 在该字典下的稀疏表示。模型为：

$$\begin{aligned} \min_{\mathcal{D}, Z} & \sum_{j=1}^N \|\mathcal{P}_j x - \mathcal{D} z_j\|_2^2, \\ \text{s.t. } & \|z_j\|_0 \leq s, \|d_i\|_2 = 1, \forall i, j. \end{aligned} \quad (1.33)$$

字典学习的方法有很多^[38,80,81]，其中 K-SVD^[38] 是字典学习中最经典的方法，DL-MRI 中即应用了这种算法，最后重建的结果比基于模型的方法有了显著的提高。作者在文献^[82] 中对 DL-MRI 模型进行了修改，提出了 SOUP-DIL (sum of outer products) 图像重建模型。模型利用了总 (aggregate) 稀疏度的概念，并提出了如下正则项：

$$\begin{aligned} \min_{\mathcal{D}, Z} & \sum_{j=1}^N \|\mathcal{P}_j x - \mathcal{D} z_j\|_2^2 + \alpha^2 \|z_j\|_0, \\ \text{s.t. } & \|d_i\|_2 = 1, \forall i. \end{aligned} \quad (1.34)$$

其中总稀疏度项为 $\sum_{j=1}^N \|z_j\|_0$ ，权重为 α^2 。数值实验表明，SOUP-DIL 算法可以提高重建图像的质量。虽然利用字典学习的方式进行压缩感知重建可以有效地学习到图像的特征，但字典学习过程中的超参数较多，比如图像块大小、字典大小等，并且没有理论上的方法来指导超参数的选择，只能凭借经验。这也是学习方法中普遍存在的问题。

由于字典学习的模型 (1.33) 中使用了 l_0 范数，求解过程中会涉及到一个非凸的且 NP 难的优化问题^[83]，计算代价很大。Ravishankar 针对这个问题提出了基于稀疏变换学习 (sparsity-transform leaning, 简称 TL) 的模型^[84]。稀疏变换学习的方法十分有效，在很多面都有应用，并且可以保证算法的收敛性。具体来说，稀疏变换学习的正则项的定义为：

$$\begin{aligned} \min_{\mathcal{S}, Z} & \sum_{j=1}^N \|\mathcal{S} \mathcal{P}_j x - z_j\|_2^2, \\ & + \alpha Q(\mathcal{S}), \text{s.t. } \|Z\|_0 \leq s, \end{aligned} \quad (1.35)$$

其中 $\mathcal{S} \in \mathbb{C}^{n \times N}$ 是一个未知的变换，函数 $Q(\mathcal{S})$ 是关于这个变换的函数：

$$Q(\mathcal{S}) = -\log|\det \mathcal{S}| + 0.5 \|\mathcal{S}\|_F^2. \quad (1.36)$$

于是基于稀疏变换学习的重建模型为：

$$\begin{aligned} \min_{\mathcal{S}, Z} & \frac{1}{2} \|Ax - b\|_2^2 + \sum_{j=1}^N \|\mathcal{S} \mathcal{P}_j x - z_j\|_2^2 \\ & + \alpha Q(\mathcal{S}), \text{s.t. } \|Z\|_0 \leq s. \end{aligned} \quad (1.37)$$

模型 (1.37) 的最重要的优势是更新 \mathcal{S} 的时候存在闭形式的解，因此也避免了字典学习更新中的繁琐计算。目前，基于稀疏变换学习的重建模型已经有了很多推广，比如

STL-MRI^[84]、UNITE-MRI^[85]、FRIST-MRI^[86]、STROLLR-MRI^[87] 等。详细的算法流程及综述请参考^[88]。

最近几年，由于深度学习^[89,90]（deep learning）的兴起和火热，出现了很多利用深度神经网络进行压缩感知重建的方法，如^[91–99]等。这大类方法的主要思路是通过神经网络来学习图像的特征，来建立下采样数据和真实数据之间的联系。根据处理图像域的不同，基于深度学习的 MR 重建可以分为四类，如图1.8所示。其中图1.8（a）是图像域的深度学习。此类方法首先从下采样的 k-space 数据中得到有高度空间伪影的 Zerofilled 图像，然后训练神经网络学习这些空间伪影，即训练一个从图像域到图像域的神经网络。此类网络的代表有深度残差学习^[94] 和域适应的神经网络^[95]，可以有效地去重建图像中的空间伪影。目前图像域的神经网络的发展趋势是寻找更复杂精妙的损失函数来克服平滑化的伪影。图1.8（b）是混合域的神经网络，这类网络是利用压缩感知模型中的数据项来进行训练和推断，将求解模型的优化算法展开成一个神经网络。这类方法是最常用的基于深度网络的重建方法。LISTA^[100] 是最早使用这个思想的网络，其将 ISTA 算法^[34] 展开成了神经网络，来学习权重和稀疏变换。VARnet^[98] 将如下优化问题展开成网络：

$$\hat{x} = \arg \min_x \frac{1}{2} \|Ax - b\|_2^2 + \alpha \sum_{i=1}^J \langle \sigma_i(\mathcal{C}_i x), \mathbf{1} \rangle. \quad (1.38)$$

这里 \mathcal{C}_i 为第 i 个通道的卷积算子， σ_i 为对应的激活函数。在 VARnet 中，需要通过网络学习卷积算子 \mathcal{C}_i 、激活函数 σ_i 和 α 。ADMM-NET^[92] 将 ADMM 算法的迭代步骤展开，并与深度神经网络的每一层一一对应。其他类似的网络有 Primal-Dual^[101] 网络和投影梯度^[102] 网络。图1.8（c）的方法直接通过深度神经网络学习频率域与图像域的映射，如 AUTOMAP^[96]。这种方法需要在卷基层后接上全连接层，因此对内存和存储的要求很高。最后一类是成像域（频率域）的神经网络，如图1.8（d）所示。这类方法直接通过神经网络学习频率域的差值并进行去噪操作，应用到 MR 重建中的网络有^[103,104]等。基于深度学习的重建算法本质上是通过神经网络学习图像的特征，如稀疏变换、空间伪影或者噪声等，其重建的效果也比一般基于模型的方法更好。但是，深度神经网络的超参数很多，比如卷积核的大小、激活函数的选择等，并且深度学习目前没有建立普遍适用的理论基础，这也是其最被诟病的一点。

综上所述，基于压缩感知的 MR 重建模型分为基于模型和基于数据两大类。其中基于模型的方法是根据图像的先验特征选择或者构造稀疏变换，比如 TV、TGV、小波变换、多尺度分析、非局部算子等。这类方法一般有着坚实的数学理论做支撑，算法的收敛性也有保证。基于数据的方法是从数据中学习图像的稀疏变换，主要有字典学习、稀疏变换学习、深度学习等。这类模型的重建结果一般比基于模型的方法要好，但是其超参数通常很多，而且计算量大。因此，在实际应用中，需要根据实际问题选择合适的模型。

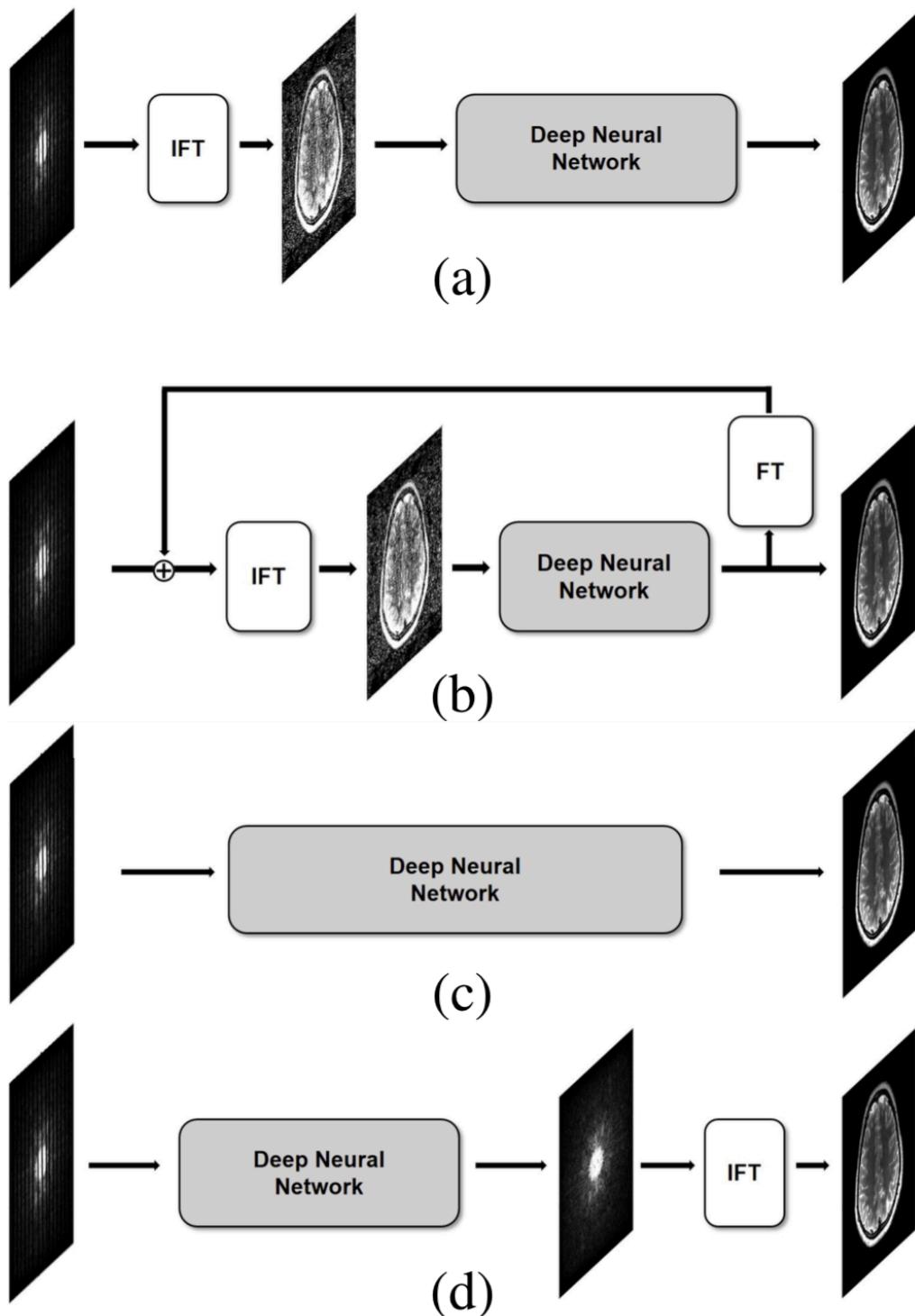


图 1.8 基于深度学习的 MR 重建。此图来自文献^[56]。

1.3.4 动态 MR 图像的重建模型

由于本论文主要针对动态 MR 图像，在这一小节我们介绍动态 MR 图像的压缩感知模型。动态 MRI 在医学临床中应用广泛，如心脏电影成像、心脏灌注、磁共振动态对比增强，功能磁共振成像等。正如第1.3.1小节所述，动态 MR 数据不仅包含了图像的空间信息，也包含了组织随时间变化的信息。因此，动态 MR 图像的冗余性不仅体现在空间方向上，也体现在时间方向上。设计动态 MR 图像的压缩感知模型的思路是利用好这两个方向上，尤其是时间方向上的冗余性。

为了与二维 MR 图像区分，我们记动态 MR 图像为 $X \in \mathbb{C}^{N_1 \times N_2 \times d}$ ，其中 d 为时间方向的维度。于是动态 MR 压缩感知的模型为：

$$\begin{aligned}\hat{X} = \arg \min_X & \| \mathcal{S}X \|_1, \\ s.t. & \| AX - B \|_F^2 \leq \epsilon.\end{aligned}\quad (1.39)$$

这里 F 为矩阵的 Fobienus 范数， B 为下采样数据矩阵，我们称之为 kt-space 数据。将问题 (1.39) 写成无约束优化问题：

$$\hat{X} = \arg \min_X \frac{1}{2} \| AX - B \|_F^2 + \alpha \| \mathcal{S}X \|_1. \quad (1.40)$$

最早将压缩感知应用到动态 MR 图像上的模型是 k-t SPARSE^[13]。模型使用了空间和时间方向的稀疏性，其中空间方向上的稀疏变换为小波变换，而时间方向上使用了一维 Fourier 变换。其模型如下：

$$\min_X \frac{1}{2} \| AX - B \|_F^2 + \alpha \| \mathcal{W}X \|_1 + \beta \| \mathcal{F}_t X \|_1. \quad (1.41)$$

其中 \mathcal{F}_t 为时间方向上的一维 Fourier 变换。模型采用了 Cartesian 随机采样，使用共轭梯度法求解。由于 Fourier 变换具有周期性，这个模型在心脏电影成像中的重建结果很好，但在其他动态 MR 图像中的应用有限。Jung 等提出了基于经典 k-t BLAST/SENSE^[105] 的改进算法 k-t FOCUSS^[45,106]，并且 k-t FOCUSS 最重要的贡献之一是揭示了动态 MR 压缩感知模型和传统的 k-t 方法的差别不大，而且可以通过简单修改 k-t BLAST/SENSE 来得到。k-t FOCUSS 的模型为：

$$\begin{aligned}\min_X & \| X \|_1, \\ s.t. & \| A\mathcal{F}_t X - B \|_F^2 \leq \epsilon.\end{aligned}\quad (1.42)$$

k-t FOCUSS 主要的想法是利用残差编码和预测。具体来说，我们将 X 分解为预测信号 X_0 和残差信号 ΔX ：

$$X = X_0 + \Delta X.$$

因此模型 (1.42) 变为：

$$\begin{aligned}\min_X & \| \Delta X \|_1, \\ s.t. & \| A\mathcal{F}_t X_0 + A\mathcal{F}_t \Delta X - B \|_F^2 \leq \epsilon.\end{aligned}\quad (1.43)$$

k-t FOCUSS 的目标是预测图像 X_0 并且使残差图像 ΔX 尽量稀疏。注意模型中只有时间方向的稀疏变换 \mathcal{F}_t , 而没有空间方向的稀疏变换。Li 等结合压缩感知、平行成像和黄金角度径向采样的想法, 提出了 iGRASP^[44], 模型如下:

$$\min_X \frac{1}{2} \|AX - B\|_F^2 + \alpha \|\nabla_t X\|_1. \quad (1.44)$$

这里黄金角度径向采样是指时间上相邻两条径向采样线的夹角是 111.25° , 称为黄金角度。这样选择的好处是采样线可以覆盖整个 kt-space 并且在时间上保证非相干性。注意同 k-t FOCUSS 一样, iGrasp 中只含有时间方向的稀疏变换。

以上的几个模型在时间方向上使用了一维 Fourier 变换或者梯度算子, 而另一类方法是将矩阵的低秩性应用到动态 MR 图像重建模型中。低秩矩阵补全^[107] (low-rank completion) 将压缩感知的概念扩展到了矩阵, 从而可以在低秩和非相干的条件下恢复出矩阵中丢失的信息。稀疏向量只有很少的非零元素, 而类似的, 低秩矩阵只有很少的非零特征值。低秩矩阵补全和压缩感知的结合可以进一步提高成像速度。这时, 我们一般将 X 转化为一个时空矩阵, 其中矩阵的每一列相当于动态图像的一帧, 每一行相当于一个体素点。矩阵的结构如下所示:

$$X = \begin{bmatrix} X(p_1, t_1) & \cdots & X(p_1, t_d) \\ \vdots & \ddots & \vdots \\ X(p_{N_1 \times N_2}, t_1) & \cdots & X(p_{N_1 \times N_2}, t_d) \end{bmatrix}, \quad (1.45)$$

其中 $p_i, i = 1, \dots, N_1 \times N_2$ 为空间坐标, $t_i, i = 1, \dots, d$ 。我们记为 $X \in \mathbb{C}^{N_1 N_2 \times d}$, 并称其为 Casorati 矩阵。可以看出, 由于时空上的相关性, X 通常是一个低秩矩阵。图1.9 (a) 展示了一个动态图像的 Casorati 矩阵。可以看出, 矩阵的每一列只有很少的一部分有变化, 大部分都是不变的, 即这个矩阵的列秩很低。

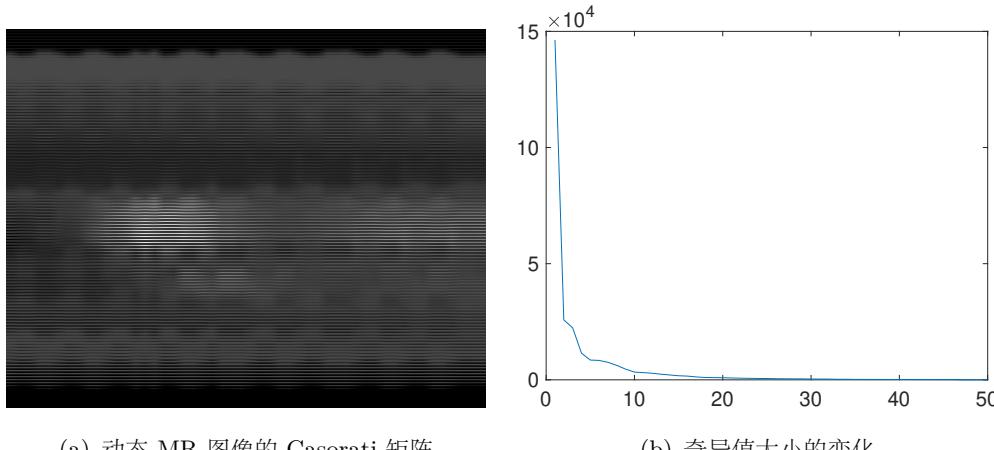


图 1.9 动态图像的低秩性。

最早将矩阵的低秩性应用到 MR 重建中的模型为^[49,108]。其中 k-t SLR^[49] 利用了三

维 TV 和核范数，其模型如下：

$$\min_X \frac{1}{2} \|AX - B\|_F^2 + \alpha \|X\|_{3DTV} + \beta \|X\|_*. \quad (1.46)$$

这里 $\|\cdot\|_*$ 是矩阵的核范数，其定义为矩阵奇异值的和，即

$$\|X\|_* = \sum_i \sigma_i,$$

其中 σ_i 为 X 的奇异值。 $\|X\|_{3DTV}$ 定义为：

$$\|\cdot\|_{3DTV} = \|\nabla_3 \cdot\|_1,$$

而

$$|\nabla_3 \cdot| = \sqrt{(\nabla_x \cdot)^2 + (\nabla_y \cdot)^2 + (\nabla_t \cdot)^2},$$

其中 ∇_x , ∇_y 和 ∇_t 分别为 x , y 和 t 方向上的梯度算子。核范数是矩阵的秩的凸松弛，被广泛应用于动态 MR 重建模型中。图1.9 (b) 展示了一个 Casorati 矩阵奇异值大小的变化。可以看出，Casorati 矩阵的奇异值只有很少一部分的值很大，大部分奇异值都接近于零，即其奇异值是 l_1 稀疏的。k-t SLR 是第一个使用核范数作为时间正则项的模型，可以很好地稀疏化非周期性运动。k-t SLR 采用伪径向采样，使用 ADMM 算法进行求解。模型在心脏体模数据和心脏电影成像的重建效果显著，但是重建算法的速度慢。

Otazo 等基于矩阵低秩与稀疏分解的思想，提出了适用于动态 MR 图像的压缩感知重建模型 L+S^[109]。矩阵分解的目标是将矩阵 X 分解为低秩部分 L 与稀疏部分 S :

$$\min_{L,S} \|L\|_* + \lambda \|S\|_1, \quad s.t. \quad X = L + S. \quad (1.47)$$

这个模型也被称为 RPCA。Candès 等^[110] 建立了其数学基础，并指出在特定的假设条件下，RPCA 可以完美地恢复出低秩和稀疏部分，其中 L 的秩需要很低并且 S 需要有很高的稀疏度。图1.10 展示了一个动态图像的低秩与稀疏分解的结果。可以看出，稀疏部分比原图像的稀疏度要高，而且图1.10 (d) 表明低秩部分的奇异值也比原图像要低。将矩阵分解的想法应用在动态 MR 重建中，模型变为：

$$\min_{L,S} \|L\|_* + \lambda \|\mathcal{S}S\|_1, \quad s.t. \quad A(L + S) = B, \quad (1.48)$$

其中低秩部分 L 建模时间上高度相关的背景，稀疏部分 S 表征背景之上的动态信息。将模型转换为无约束问题：

$$\min_{L,S} \frac{1}{2} \|A(L + S) - B\|_F^2 + \alpha \|L\|_* + \beta \|\mathcal{S}S\|_1. \quad (1.49)$$

文章中根据不同数据选择了不同的稀疏变换 \mathcal{S} 。比如在心脏电影成像中，选择 \mathcal{S} 为时间方向的一维 Fourier 变换 \mathcal{F}_t :

$$\min_{L,S} \frac{1}{2} \|A(L + S) - B\|_F^2 + \alpha \|L\|_* + \beta \|\mathcal{F}_t S\|_1. \quad (1.50)$$

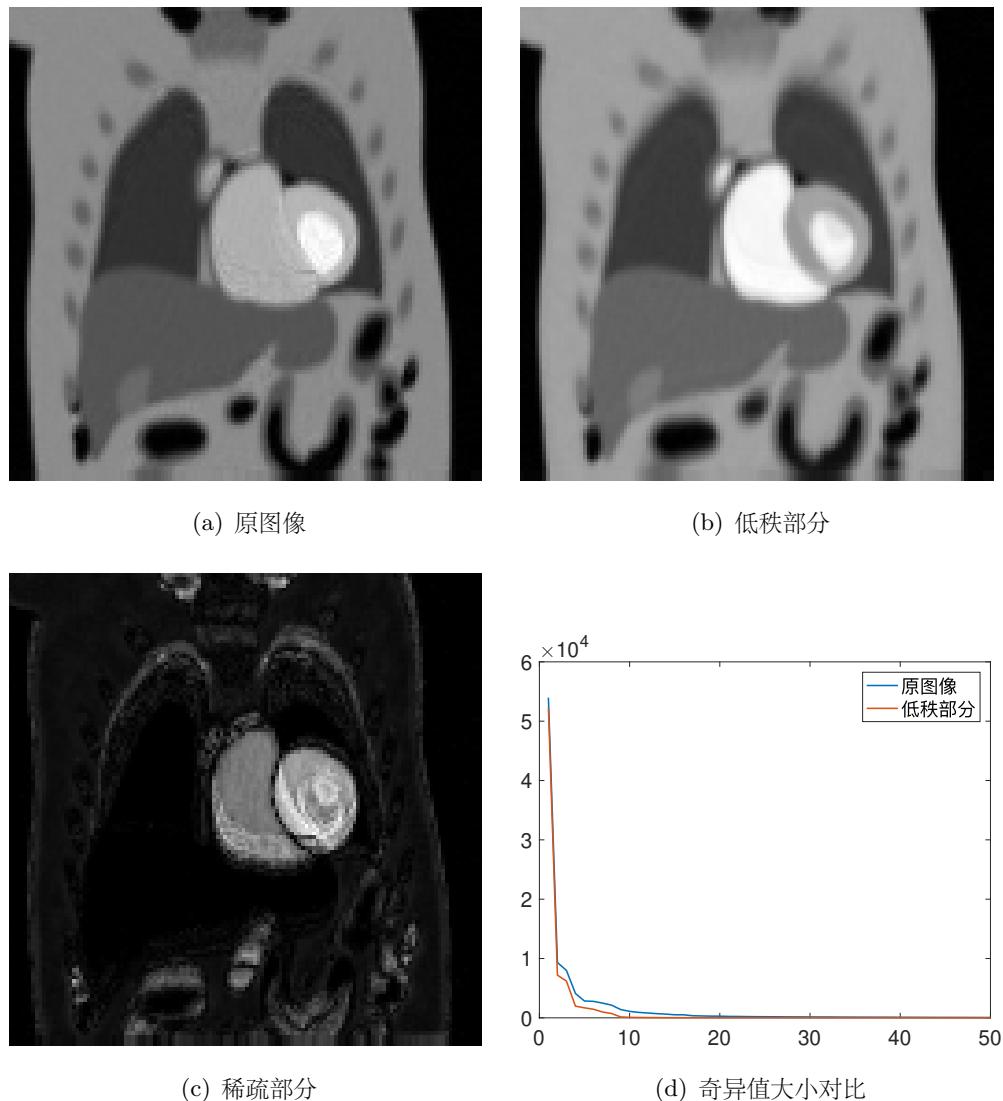


图 1.10 动态图像的低秩稀疏分解 (一帧)。

文献^[111] 中也使用了 \mathcal{F}_t 作为时间方向上的稀疏变换。在血管造影中，选择 \mathcal{S} 为单位矩阵（即没有稀疏变换）：

$$\min_{L,S} \frac{1}{2} \|A(L + S) - B\|_F^2 + \alpha \|L\|_* + \beta \|S\|_1. \quad (1.51)$$

在胸部 DCE-MRI 中，选择 \mathcal{S} 为时间方向上的差分算子 ∇_t ：

$$\min_{L,S} \frac{1}{2} \|A(L + S) - B\|_F^2 + \alpha \|L\|_* + \beta \|\nabla_t S\|_1. \quad (1.52)$$

同样利用图像分解思想，Schloegl 等^[61] 提出了卷积下确界 TGV 泛函（infimal convolution TGV，简称 ICTGV），其定义为：

$$\text{ICTGV}_{\alpha,\beta}^2(X) = \inf_{X=X_1+X_2} \text{TGV}_{\alpha_1}^2(X_1) + \beta \text{TGV}_{\alpha_2}^2(X_2), \quad (1.53)$$

相对应的重建模型为：

$$\min_X \frac{1}{2} \|AX - B\|_F^2 + \text{ICTGV}_{\alpha,\beta}^2(X). \quad (1.54)$$

有关 TGV 的详细概念请参考第3章。

与结构化稀疏的概念类似，最近结构化低秩 (structured low-rank) 的概念也获得了广泛的关注，并且在 MR 成像上有着显著的应用前景。这类方法通过巧妙地构造 Hankel 结构化矩阵，将卷积运算转化为矩阵运算，然后通过极小化 Hankel 的秩构造并求解模型。这类方法的基本模型为：

$$\begin{aligned} & \min_X \text{rank}(\mathcal{H}(X)), \\ & \text{s.t. } \mathcal{P}_{\mathcal{O}}(X) = \mathcal{P}_{\mathcal{O}}(B). \end{aligned} \quad (1.55)$$

其中 \mathcal{H} 为 Hankel 结构化矩阵算子， $\mathcal{P}_{\mathcal{O}}$ 为指标集 \mathcal{O} 上的投影算子。注意这里的 X 同 B 一样，为 kt-space 数据，且目前结构化低秩的方法大多在 kt-space 中构造 \mathcal{H} 矩阵。主要的模型有 SAKE^[112]、LORAKS^[113]、ALOHA^[114,115]、GIRAF^[116] 等，更详细的综述请参考文献^[56]。

除了以上基于模型的方法，基于学习的方法也被广泛地应用在动态 MR 成像中。Lingala^[117] 提出了针对动态 MR 图像重建的盲压缩感知框架。对于一个动态 MR 图像 $X \in \mathbb{C}^{N_1 N_2 \times d}$ ，盲压缩感知框架将其建模为两个矩阵的 $U \in \mathbb{C}^{N_1 N_2 \times K}$ 和 $V \in \mathbb{C}^{K \times d}$ 的乘积：

$$X = UV,$$

其中 K 为过完全字典中基函数的个数。这里 U 为空间系数矩阵，其通常是稀疏的；而 V 为包含时间基函数的字典。盲压缩感知模型即为从采样中估计 U 和 V 的过程：

$$\begin{aligned} \{\hat{U}, \hat{V}\} &= \arg \min_{U,V} \frac{1}{2} \|A(UV) - B\|_F^2 + \alpha \|U\|_1, \\ &\text{s.t. } \|V\|_F^2 \leq s. \end{aligned} \quad (1.56)$$

在自由呼吸心肌灌注 MR 图像上的数值实验表明，盲压缩感知模型对时空模糊更加鲁棒，并且更好地保持了精细的结构信息。Wang^[118] 等针对动态心脏电影 MR 图像，提出了基于三维时空字典的模型：

$$\begin{aligned} \min_{x, \mathcal{D}, z_j} & \frac{1}{2} \|AX - B\|_F^2 + \frac{\alpha}{2} \sum_j \|\mathcal{P}_j x - \mathcal{D}z_j\|_2^2 \\ & + \beta \sum_j \|z_j\|_0 + \gamma \|\mathcal{S}X\|_1. \end{aligned} \quad (1.57)$$

这里与二维的情形不同， \mathcal{P}_j 算子的作用是从图像 X 中提取时空图像块，且稀疏变换 \mathcal{S} 为 3DTV 算子。模型采用交替方向法求解，在字典学习和图像重建中进行迭代。类似的，Caballero 等^[119] 也基于心脏电影成像提出了时空字典学习模型。不同的是，Caballero 模型中的稀疏变换为时间方向的梯度算子 ∇_t 。Ravishankar 等^[84] 基于图像分解的框架，并结合低秩的思想和字典学习的方法，提出了 LASSI 模型：

$$\begin{aligned} \min_{L, S, \mathcal{D}, Z} & \frac{1}{2} \|A(L + S) - B\|_F^2 + \alpha \|\mathcal{S}_1(L)\|_* \\ & + \beta \left\{ \sum_j \|\mathcal{P}_j S - \mathcal{D}z_j\|_2^2 + \lambda^2 \|Z\|_0 \right\}, \\ s.t. \quad & \|Z\|_\infty \leq a, \text{rank}(\mathcal{S}_2(d_i)) \leq r, \|d_i\|_2 = 1, \forall i. \end{aligned} \quad (1.58)$$

数值试验表明，LASSI 模型适用于多种动态图像，并且比一般的盲压缩感知方法的表现好。

近几年来，随着深度学习的兴起，深度神经网络也被广泛应用在了动态 MR 成像中。我们已经在第1.3.3中介绍了深度学习在二维 MR 图像重建中模型，但是其在动态 MRI 上的应用还处于起步阶段，模型不多。Schlemper 等^[120] 首次将深度学习的思想应用在了动态 MR 图像重建中，提出了基于深度级联卷积神经网络的模型：

$$\min_X \|X - f_{\text{cnn}(X_z | \theta)}\|_F^2 + \alpha \|AX - B\|_F^2. \quad (1.59)$$

这里 f_{cnn} 为卷积神经网络的向前映射，参数为 θ ，其中包含了神经网络的权重，而 X_z 为 Zero-filled 图像。由于 Zero-filled 图像中包含着严重的空间伪影，CNN 重建图像的过程可以看成是去除空间伪影的过程。该方法可以在高采样率下保持图像的解剖结构，且重建速度快。Qin 等^[121] 提出了利用循环神经网络来重建动态 MR 图像。模型创新性地提出了双向卷积循环单元来学习图像在时空上的冗余性，在心脏图像重建中取得了很好地效果。Wang^[122] 等在模型 (1.59) 的基础上，提出了 DIMENSION 模型。模型同时利用了动态图像的 k-space 和空间先验信息，通过监督学习来训练网络。网络由两个主要部分构成，一部分为频率域网络，用于更新 k-space；另一部分为空间域网络，用来提取图像中的高级特征。两部分网络通过 Fourier 变换进行连接。数值实验表明，DIMENSION 可以提高重建图像质量和重建速度。Ke 等^[123] 提出了用于动态 MR 重建的级联残差密集网络 CRDN。具体来说，网络利用了卷积层的多层级局部和非局部信息，并且使用了

基于 TV 的增强边缘的损失函数。对比试验表明，CRDN 可以在提高重建质量的基础上减少重建时间。

综上所述，动态 MR 图像的稀疏性不仅体现在空间上，也体现在时间上。而且，时间上稀疏变换的选择是尤其重要的，可以很大程度上决定重建结果的好坏。目前，基于模型的方法既有传统的利用 Fourier 变换、TV 的模型，也有利用图像的低秩性或结构低秩性和基于分解思想的方法。基于数据的方法相对较少，主要有训练时空字典的模型和搭建深度神经网络的方法。

1.3.5 本节小结

本节主要介绍了压缩感知在 MR 成像中的应用。我们首先讨论了 MR 图像的稀疏性和常见的采样模式，然后回顾了压缩感知 MR 成像的经典模型。最后由于本论文主要针对动态 MR 图像，我们也给出了动态 MR 图像压缩感知模型的综述。

1.4 磁共振指纹的研究现状

磁共振指纹（magnetic resonance fingerprinting, 简称 MRF）是一种新的数据获取、后处理以及可视化的框架。与传统 qMRI 一次只能测量一个参数不同，MRF 可以在单次扫描中同时量化多个组织参数，比如 T_1 , T_2 , 质子密度等，并且提高参数图的信噪比和扫描效率。MRF 主要由三个部分组成，分别为信号获取、预定义字典的生成和模式识别重建参数图。具体来说，首先针对特定的问题，选择一个合适的 MR 序列进行信号采集。然后利用 MR 成像的数学模型，生成一个信号演化的字典，字典中的每一个原子模拟了不同参数的组织在该 MR 序列下的演化，即包含着组织的参数信息。最后，使用模式识别算法将采集的信号与字典进行匹配，选择字典中最配的原子，其包含的参数信息即为该体素的参数。MRF 的主要流程如图 1.11 所示。首先我们应该选取对所研

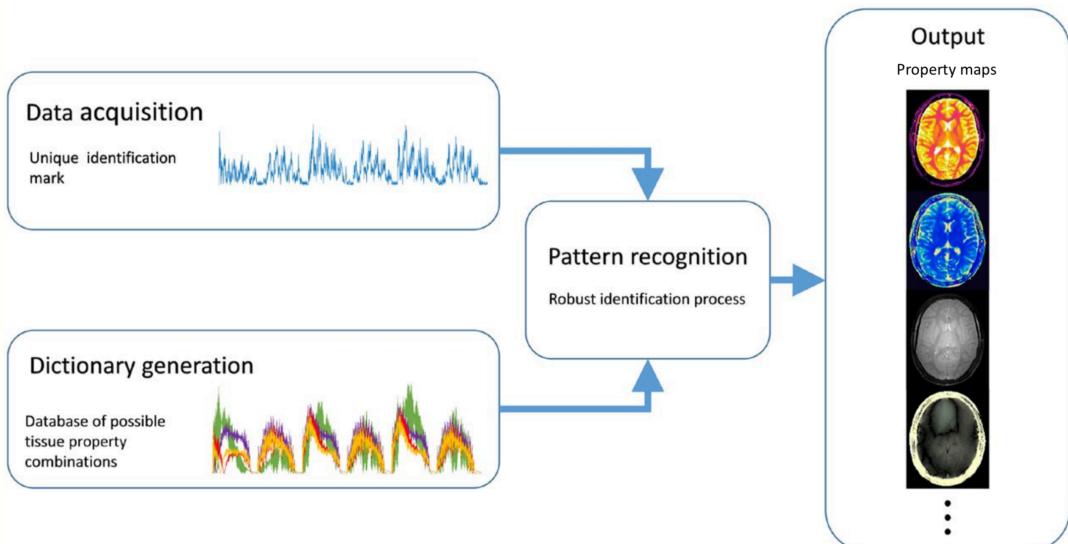


图 1.11 MRF 流程示意图。此图来自文献^[124]。

究参数敏感的 MR 序列对信号进行采样，并且序列的参数，如偏转角 (flip angle)，重复时间 (repetition time, T_R) 等需要随着时间随机变动，使得组织在 MR 序列中产生指纹状的信号演化。其次，字典中包含着不同参数的组织在该 MR 序列中的模拟演化。最后，模式识别算法用来比较每个像素指纹和字典中原子的匹配度，重建参数图。

1.4.1 信号获取

信号获取是 MRF 重建的第一步，我们首先需要根据具体问题，选择合适的 MR 序列进行信号采集。关于 MR 成像以及序列请参考第2章。MR 序列的选取需要满足以下三个方面。第一，选取的 MR 序列需要对我们所研究的参数敏感，如 T_1 , T_2 等。第二，在数据采集过程中，采集参数如偏转角，重复时间等需要连续随机变化。第三，每隔 T_R 时刻，使用伪随机下采样进行信号采集。图1.12展示了偏转角 (a) 和重复时间 (b) 随时间变化的例子。这样做的目的是使不同参数的组织可以产生独特的信号演化，我们称之为指纹，即每个指纹中包含着不同组织的参数信息，以便在后面进行参数图重建。图1.13是某一 T_R 时刻的体模和活体人脑的指纹图像。可以看出由于高度的下采样，图像中存在着感强烈的伪影。MRF 重建的目标即是从指纹图像中重建出参数图。

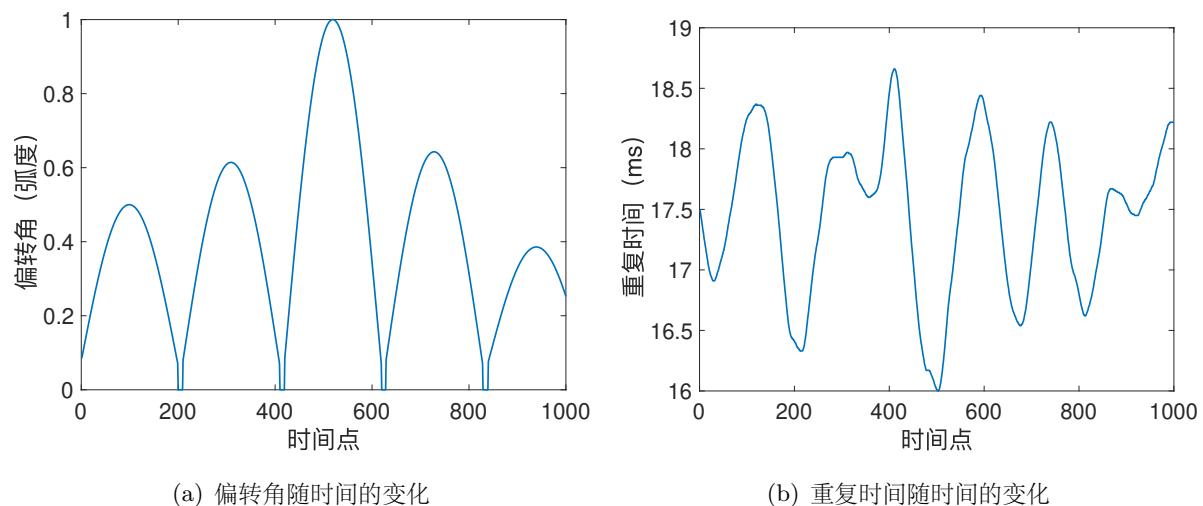


图 1.12 偏转角 (a) 和重复时间 (b) 随时间的变化。

目前常用的 MRF 序列主要有两种。Ma^[3] 在 MRF 的第一篇文章中使用了平衡稳态自由进动序列 (balanced steady state free precession, bSSFP^[125])，因为目前对 bSSFP 序列的研究最充分。bSSFP 对 T_1 , T_2 和 B_0 敏感，并且具有高信噪比、高扫描效率的性质。对于 MRF 来说，bSSFP 是一个适合的序列。另一种常用的序列为 FISP 序列，或者称为非平衡稳态自由进动序列 (unbalanced steady state free precession, uSSFP)。这个序列的特点是在每个 T_R 结束之前，加上了一个非平衡的梯度场。Jiang^[126] 等首先将 uSSFP 序列应用到了 MRF 上 uSSFP 序列有助于消除带状伪影，但是其只对 T_1 和 T_2 敏感，对 B_0 不敏感，因此不能获得 B_0 的参数图。其他种类序列的应用请参考文献^[127,128]。

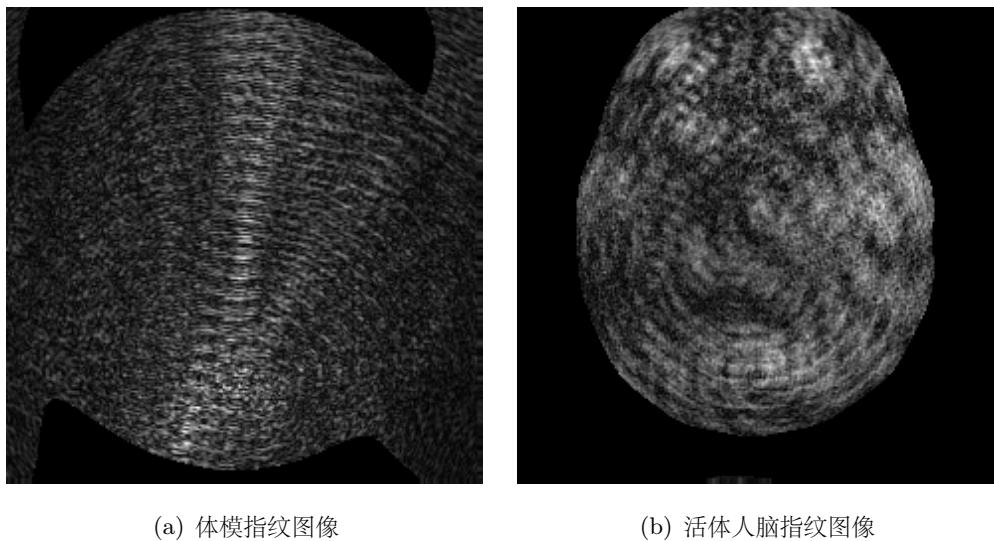


图 1.13 指纹图像举例（某一时间点）。

1.4.2 字典生成

MRF 字典中包含了不同参数的组织在 MR 序列中的模拟信号演化，字典中的每一个原子都包含着一组参数。因此，MRF 重建的精度取决于生成字典所用模型的精度和字典的大小。如果字典过小，则会导致重建参数图结果不精确；如果字典过大，生成字典和之后的匹配算法所需的时间就会增多。因此，需要在字典大小和重建速度之间寻找平衡。一般来说，对于一个给定的 MRF 序列来说，字典只需要生成一次，并且在之后的应用中保持不变。需要指出的是，MRF 的字典和字典学习中的字典是不同的。MRF 中字典的原子是组织的时间演化曲线，而字典学习中字典的原子是基于图像块的特征。

目前生成字典主要有两种数学模型。Ma 在^[3] 中使用了简单的 Bloch 方程来模拟 bSSFP 序列中射频场和时序对于单一等色子 (isochromat) 的影响。模型假设指纹中的每个体素都由单一等色子构成。Bloch 方程也可以应用在 uSSFP 中，但是会消耗更多的时间。这是因为在 uSSFP 中，由于非平衡梯度场的存在，指纹中的每个体素都是多个等色子的平均。因此 Bloch 方程在处理 uSSFP 时的效率不高。

扩展相图 (extended phase graph, 简称 EPG^[129]) 是模拟 uSSFP 的另一种常用的模型。图1.14展示了 EPG 模型生成字典中的原子。EPG 模型将体素内的自旋系统中描述为离散的状态矩阵，可以有效地表示自旋系统在非平衡梯度场的影响下随时间的演化。EPG 的具体细节请看第2.2小节。EPG 模型的另一个好处是在原有模型的基础上，加上其他参数的影响，如 B_0 , B_1^+ , 磁化转移等。因此 EPG 模型在 MRF 重建中有着广泛的应用。但是由于 EPG 模型需要计算和存储等色子的状态矩阵，其运行速度通常比较慢。

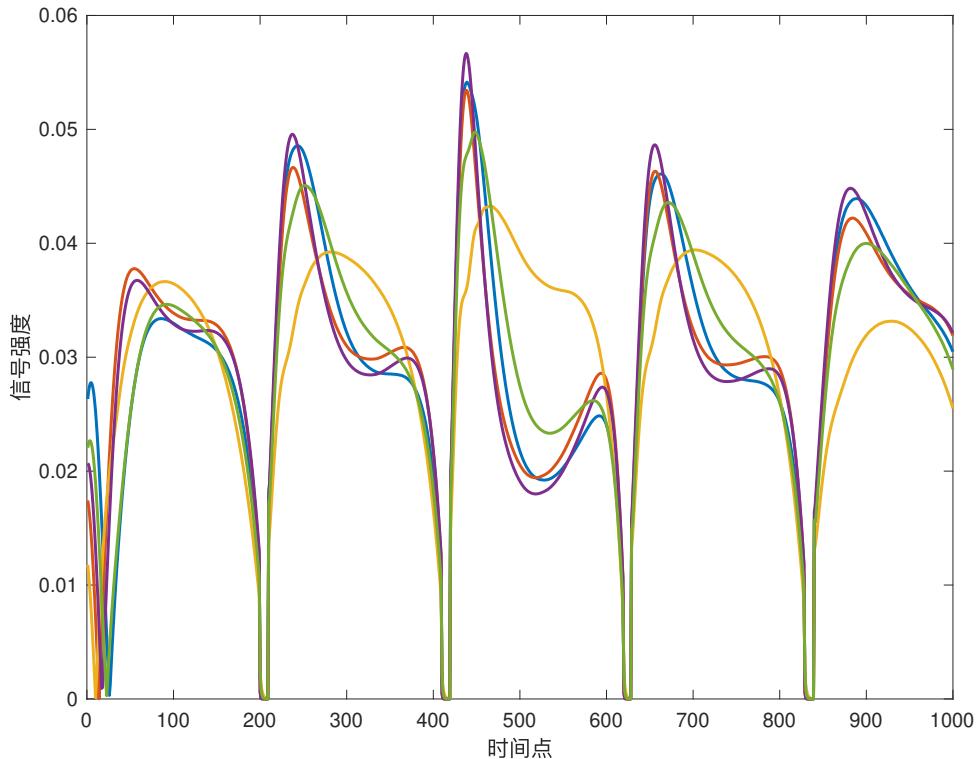


图 1.14 EPG 模型生成的字典原子。

1.4.3 参数图重建

MRF 重建的最后一步是选择合适的匹配算法，将采集到的指纹信号和生成的字典进行匹配，重建参数图。因此参数图的准确性取决于匹配算法的是否对噪声和下采样伪影具有鲁棒性。由于 MRF 数据维度一般很大，匹配算法所消耗的时间很多。因此目前 MRF 的研究重点是在不降低参数图质量的情况下加速匹配算法。参数图的重建算法可以分成以下四大类。

Ma 等^[3] 使用了模板匹配 (template matching) 的方法来重建参数图。对于每个指纹体素，模板匹配从字典中选择与该体素最配的原子，从而获得该体素的参数值。记 $X = \{x_n \in \mathbb{C}^L\}, n = 1, \dots, N$ 为采集到的指纹数据， $\mathcal{D} = \{d_k \in \mathbb{C}^L\}, k = 1, \dots, K$ 为生成的字典。那么模板匹配即为从字典 \mathcal{D} 中选取和 x_n 内积最大的原子：

$$\hat{k}_n = \arg \max_k |\langle \mathbf{d}_k, \mathbf{x}_n \rangle|, \quad (1.60)$$

且质子密度为：

$$\hat{\rho}_n = |\langle \mathbf{d}_{\hat{k}_n}, \mathbf{x}_n \rangle|. \quad (1.61)$$

模板匹配的方法可以精确地重建参数图，并且对噪声和伪影有鲁棒性。图1.15展示了使用模版匹配方法重建体模 MRF 数据的参数图。通过简单的计算可以看出，模板匹配的计算复杂度为 $O(NL^2K)$ 。一般来说，MRF 信号的时间点 L 通常在 1,000 以上，字典的个数通常在 10,000 以上，而体素的个数也会达到 10,000 以上。所以模板匹配所消耗的时间长，使用 MATLAB 软件在 CPU 上运行通常需要几十分钟甚至几小时。因此，虽

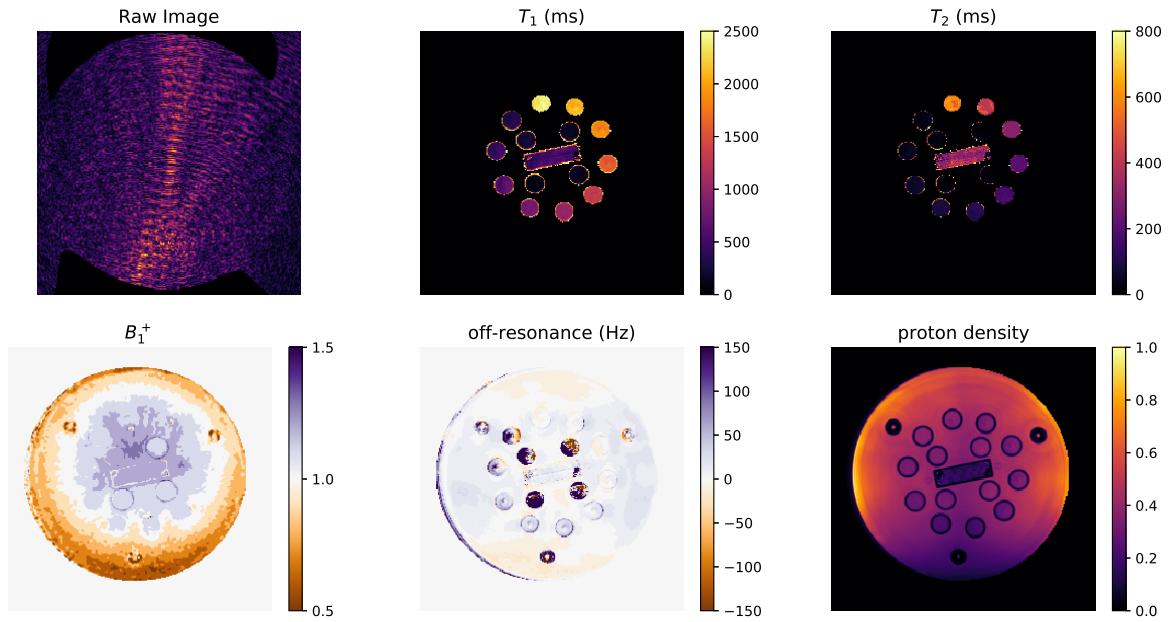


图 1.15 通过模板匹配重建的参数图。

然模板匹配的方法对噪声和空间伪影具有鲁棒性，在临床中需要更快速的方法来重建参数图。

第二类方法是利用降维的思想，先对字典或者指纹进行降维处理，然后在降维后的字典和指纹上进行匹配。这类方法本质上是对模板匹配的近似。Cauley^[130] 等利用了分组的思想进行 MRF 参数图重建。算法先将字典中的原子分成若干组，使得组内的原子高度相关，并且用组内所有原子的均值来表示这个组的指纹信号，然后再将指纹数据与分组之后的字典进行匹配。该算法大大加速了字典匹配的速度，但是由于依然需要生成字典，字典生成的时间并没有减少。McGivney^[131] 等人提出了利用奇异值分解 (singular value decomposition, 简称 SVD) 的方法将字典和指纹数据投影到一个低维子空间，并在子空间中进行匹配。这个方法本质上是压缩了时间方向的维度 L ，从而提高了运算速度。但是这种方法依然需要生成和存储字典，并且相比于模板匹配存在重建误差。

第三类方法是将压缩感知理论应用到 MRF 重建中，在重建模型中给参数图加上一些先验信息，以此提高 MRF 的重建效果。Davies^[132] 等给出了使用压缩感知进行 MRF 重建的一般框架，将信号的 Bloch 响应流型描述成连续信号，并在这个流型上进行下采样。Cline^[133] 等改进了 Davies 的模型，将 B_0 先验信息加入到了模型中，并用基于压缩感知的迭代算法求解模型。Pierre^[134] 等提出了迭代多尺度 MRF 重建算法。该方法利用了 k-space 的先验信息，在数据项和模式识别之间迭代计算直到收敛。Wang^[135] 等使用了基于小波变换的压缩感知框架对指纹数据的每一帧进行估计，然后再将重建后的图像和字典进行匹配。Zhao 等^[136] 利用统计学框架对组织参数进行最大似然估计。在 MRF 中，指纹图像 $X \in \mathbb{C}^{N \times L}$ 可以被参数化为：

$$X_l = \phi_l(T_1, T_2)\rho, \quad (1.62)$$

其中 $l = 1, \dots, L$, ρ 为质子密度, T_1 和 T_2 分别为纵向弛豫时间和横向弛豫时间矩阵, $\phi_l(\cdot)$ 为第 l 次采集时的对比度加权函数。于是, 图像采集可以被建模为:

$$B_l = AX_l + \epsilon_l = A\phi_l(T_1, T_2)\rho + \epsilon_l. \quad (1.63)$$

则 MRF 重建的过程即为从高度下采样的数据 B_l 中估计参数 $\{T_1, T_2, \rho\}$ 的过程。给定上述方程和噪声特征, 重建问题可以被建模为如下极大似然估计问题:

$$\{\hat{T}_1, \hat{T}_2, \hat{\rho}\} = \arg \min_{T_1, T_2, \rho} \sum_{l=1}^L \|B_l - A\phi_l(T_1, T_2)\rho\|_2^2. \quad (1.64)$$

该模型可以提高参数估计的精度并减少采集时间。综合以上方法可以看出, 基于压缩感知的 MRF 重建在一定程度上取得了良好的效果, 主要体现在减少采集时间和提高参数图的精度。

最后一大类方法是利用深度学习的框架来重建参数图。Cohen^[137] 等利用全连接网络来训练和重建参数图, Hoppe^[138] 和 Fabian^[139] 等构建了 MRF 的卷积神经网络。虽然这些深度神经网络取得了一定的效果, 但是由于网络均使用了监督学习方法, 训练所需的真实值依旧由模板匹配生成。因此, 网络学习的效果也取决于模板匹配的准确性。

综上所述, MRF 参数图的重建需要在精度与速度之间寻求平衡。模板匹配是最基本的重建方法, 重建的参数图的精度最高, 但是计算比较耗时。降维和深度学习的方法提高了重建的速度, 但重建的参数图不可避免地存在误差, 而且没有解决字典生成速度慢的问题。压缩感知的方法可以减少扫描时间, 但其重建的参数图同样误差偏高。

1.4.4 本节小结

本节主要回顾了磁共振指纹的研究现状。磁共振指纹有三个部分, 分别为数据采集、字典生成和模式识别。目前磁共振指纹的瓶颈在于字典生成和匹配的速度慢, 使用降维、压缩感知或者深度学习的方法会使得参数图的精度下降。因此, 如何在保证参数图精度的前提下提高 MRF 的重建速度是临幊上亟待解决的问题。

1.5 本文的主要研究内容和章节安排

加快 MR 成像速度一直是临幊中十分重要的问题, 而基于压缩感知理论的 MR 成像有着很好的重建效果, 并且具有完善的理论分析。MRF 是新的量化 MRI 的方法, 可以在单次扫描中同时得到多个组织参数。本文讨论的重点是基于压缩感知的动态 MR 图像的重建和基于 GPU 的 MRF 字典生成与参数图重建。全文章节安排的逻辑示意图如图1.16所示。

第二章, 我们首先简单回顾 MR 成像的基本概念, 包括射频场、梯度场、驰豫、k-space 和 MR 序列等。这些内容与 MR 重建有着密切的联系, 了解 MR 的成像过程可以更深入理解基于压缩感知的 MR 重建模型与 MRF 采样和匹配的过程。最后我们

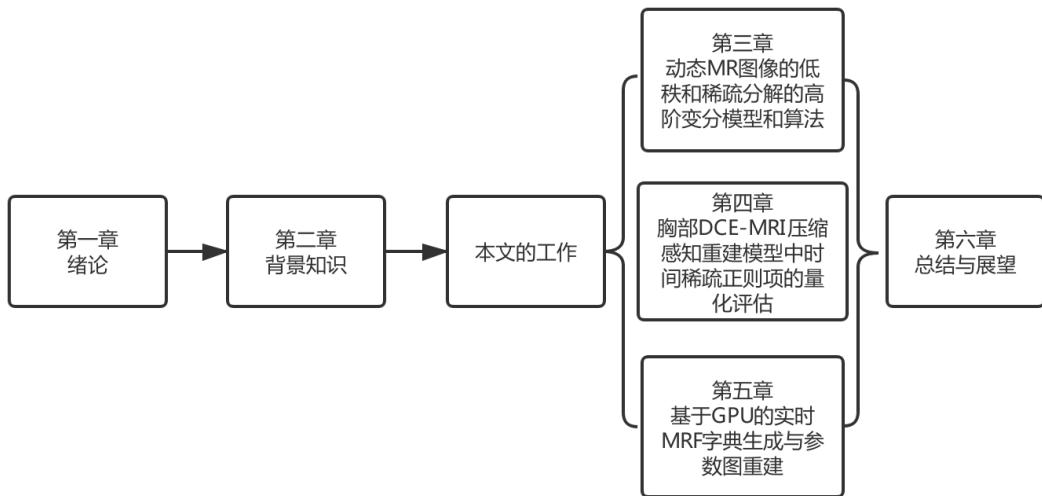


图 1.16 全文章节安排逻辑示意图。

给出了 EPG 模型的数学推导过程，这对有助于我们理解 MRF 字典的生成和第五章的程序设计。

第三章，针对动态 MR 图像，利用压缩感知和图像分解的思想，我们提出了基于二阶时空 TGV 和核范数的重建模型。模型将图像分解为低秩部分和稀疏部分，其中低秩部分用核范数约束，稀疏部分用二阶时空 TGV 约束。核范数用来建模动态图像中时间方向高度相关的背景部分，并且可以很好地去除空间伪影；而 TGV 泛函用来表示图像中的光滑部分，可以在保证重建图像边界清晰的同时减少重建图像中的阶梯效应。我们利用 Primal-Dual 算法来求解模型，并给出了算法收敛时的范数估计。为了减少算法的计算时间，我们也利用了图形处理单元来加速 MRTLAB 程序。我们针对体模、心脏灌注与胸部 DCE-MRI 图像，对比了四种最前沿的针对动态 MR 重建模型在不同采样模式和不同采样率下的表现。数值实验结果表明，相比于其他四个模型，我们提出的模型在不同采样模式和不同采样率下，可以更好地消除空间伪影并且保证图像边缘的清晰，对于胸部 DCE-MRI 图像的效果尤其显著。

第四章，针对胸部 DCE-MRI 图像，我们比较了 5 种不同的时间方向的稀疏变换，并对重建结果进行了定量分析。这 5 种稀疏变换分别为 Fourier 变换、Haar 小波变换、二阶 TGV 和核范数。所有模型均使用 FISTA 快速算法进行求解。数值实验结果表明，核范数可以得到最高的信噪比，而 TV/二阶 TGV 可以得到最精确地定量分析。因此，对于胸部 DCE-MRI，选择 TV/二阶 TGV 作为稀疏变换可以更好地重建病灶部分，而选择核范数则可以提高图像的整体信噪比。

第五章，针对 MRF 字典生成与匹配速度慢的问题，我们使用图形处理单元来进行字典生成和模板匹配，并开发了一款开源程序 snapMRF。snapMRF 将所有计算全部放在 GPU 上进行，可以快速并准确地重建参数图，并且适用于不同的 MR 序列。相比于其他 MRF 开源程序，snapMRF 的字典生成速度提高了 10–1000 倍，字典匹配速度提高

了 10–100 倍。我们也给出了六种不同 MR 序列的单元测试，保证了程序的准确性。

第六章，我们对本文的工作进行了总结并给出了下一步工作展望。

2 背景知识

为了保持文章的完整性，这一章我们介绍一些与 MR 成像和 MRF 重建相关的预备知识。

2.1 MRI 基本原理

在这一小节，我们简单回顾一下 MR 成像的物理原理和数学建模，这有利于我们加深对压缩感知 MRI 的理解。更详细的 MR 成像原理请参考文献^[1,140,141]。

2.1.1 MRI 的物理原理和数学模型

核磁共振成像可以非侵入性地获取人体内组织的图像，成像分辨率高，并且没有暴露在辐射中的危险。它主要利用了磁共振原理，反应了磁化粒子（主要是氢原子）和磁场的交互作用。MR 成像设备如图2.1所示，其内部系统如图2.2所示。其中有三个磁场



图 2.1 MR 成像设备。

的概念最为重要，第一个是主磁场 (the main magnetic field)，由主磁场线圈 (magnetic coil) 产生，如图2.3所示。主磁场是大小和方向都不变的静磁场，一般约定为 z 轴方向，记为 \mathbf{B}_0 。第二个磁场是射频场 (radio frequency, 简称 RF)，由射频线圈 (radio frequency coil) 产生，如图2.4所示。射频场是周期性变化的磁场，方向与主磁场垂直，记为 \mathbf{B}_1 。其作用是激活待成像粒子，使之发生共振，用于之后的成像。第三个磁场是

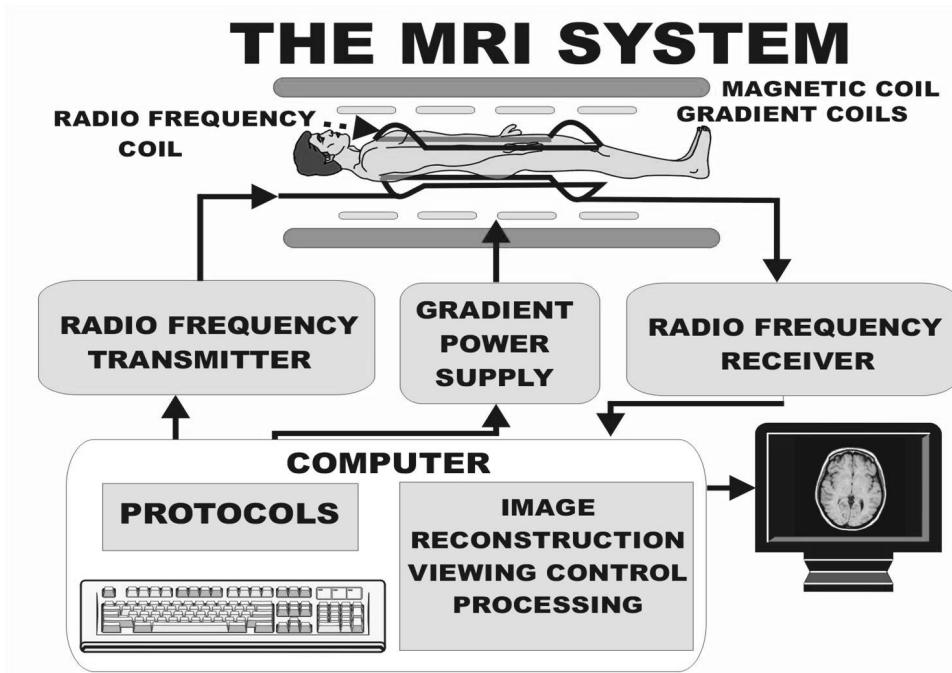


图 2.2 MR 成像系统。此图来自于文献^[142]。

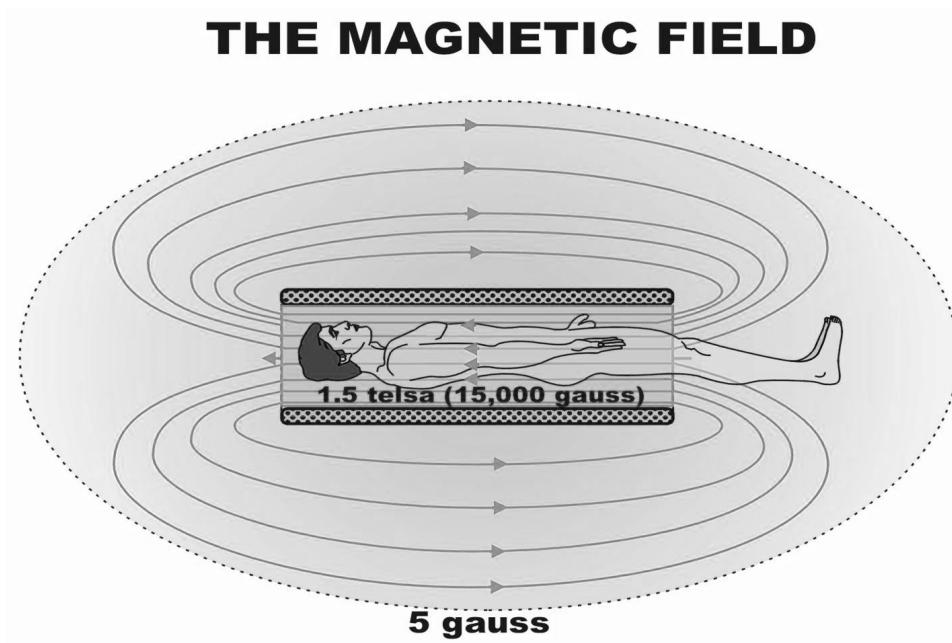


图 2.3 MR 主磁场。此图来自于文献^[142]。

THE RF COILS

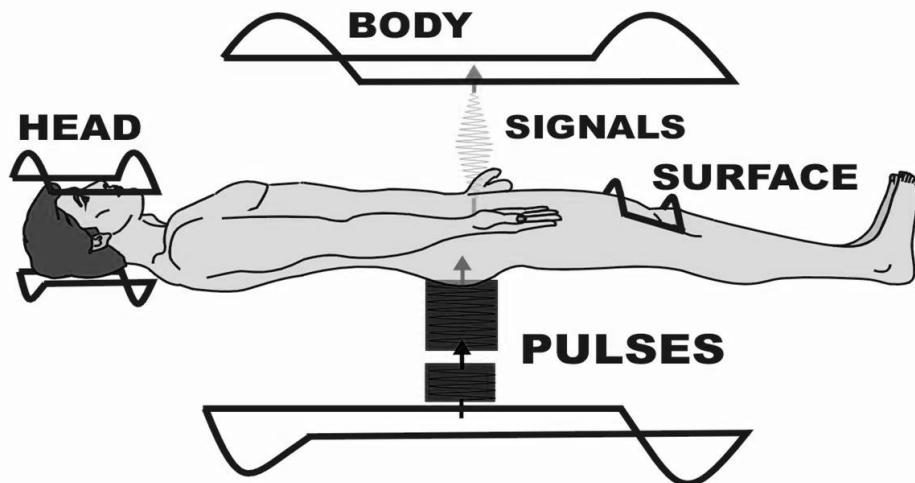


图 2.4 MR 射频场。此图来自于文献^[142]。

梯度场 (gradient fields)，由梯度线圈产生 (gradient coils)，如图2.5所示。梯度场的方向与主磁场平行，且大小随着空间位置的变化而变化。梯度场的作用是使得主磁场的大小发生改变，在 x , y 或者 z 方向上产生一个梯度，用于之后的成像过程。

成像前，所有待成像原子都做无规则的运动，其合磁场为 $\mathbf{0}$ 。开始成像时，打开主磁场 \mathbf{B}_0 。一段时间后，所有待成像的原子会被主磁场磁化，进入热平衡状态，产生一个平行于主磁场方向的磁动量 \mathbf{M} 。然后打开射频场，于此同时打开预先设定好的梯度场，这样会使得待成像原子产生共振，处于激发态。宏观上来看， \mathbf{M} 会绕着磁场 \mathbf{B}_0 以 Larmor 频率 ω 旋转：

$$\omega = -\gamma|\mathbf{B}|. \quad (2.1)$$

其中 γ 是旋磁率 (常数)。这个过程被成为进动 (procession)。可以看出，磁动量的 Larmor 频率的大小与其所处的磁场大小成正比。这里的磁场 \mathbf{B} 由四部分构成，分别为主磁场 \mathbf{B}_0 和三个方向的梯度场 $G_x(t)$, $G_y(t)$ 和 $G_z(t)$:

$$\mathbf{B}(x, y, z, t) = \mathbf{B}_0 + xG_x(t) + yG_y(t) + zG_z(t). \quad (2.2)$$

其中主磁场的大小和方向不变，而梯度场是大小随时间和空间变化的磁场，决定了磁场 \mathbf{B} 的大小和方向。记梯度场向量为 $\mathbf{G}(t) = (G_x(t), G_y(t), G_z(t))$ 。

\mathbf{M} 被射频场从热平衡态激发之后，逐渐偏离主磁场，并产生一个垂直与主磁场方向的磁化 M_{xy} 。我们将 \mathbf{M} 偏离主磁场的角度， \mathbf{M} 即与 z 轴的夹角称为偏转角 (flip angle)，记为 α ，将 \mathbf{M} 与 x 轴的夹角称为相位角 (phase angle)，记为 ϕ ，将磁动量在

A MAGNETIC FIELD GRADIENT

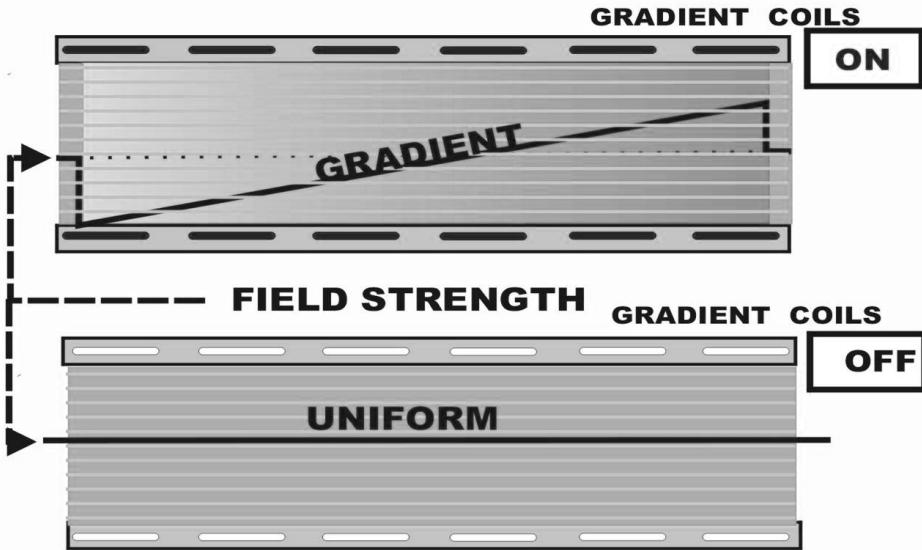


图 2.5 MR 梯度场。此图来自于文献^[142]。

x, y, z 处, t 时刻的密度 (density) 记为 $m(x, y, z, t)$, 则进动的角度或者相位累积为:

$$\begin{aligned}\phi(x, y, z, t) &= - \int_{t'=0}^t \gamma |\mathbf{B}(x, y, z, t')| dt' \\ &= \omega_0 t - 2\pi(xk_x(t) + yk_y(t) + zk_z(t)),\end{aligned}\tag{2.3}$$

其中

$$\omega_0 = -\gamma |\mathbf{B}_0|,\tag{2.4}$$

$$\begin{aligned}k_x(t) &= \frac{\gamma}{2\pi} \int_{t'=0}^t G_x(t') dt', \\ k_y(t) &= \frac{\gamma}{2\pi} \int_{t'=0}^t G_y(t') dt', \\ k_z(t) &= \frac{\gamma}{2\pi} \int_{t'=0}^t G_z(t') dt'.\end{aligned}\tag{2.5}$$

通过上面的方程可以看出, 位置 x, y, z 相对于初始位置的累积进动角度等于位置向量 $\mathbf{r} = (x, y, z)$ 与向量 $\mathbf{k}(t) = (k_x(t), k_y(t), k_z(t))$ 的内积的 -2π 倍。我们定义 $\mathbf{k}(t)$ 为 k-space。k-space 是 MR 成像中最基本的概念, 在之后的章节也会用到。有了相位, 我们就可以计算任意时刻在位置 \mathbf{r} 的磁场密度:

$$\begin{aligned}m_{xy}(x, y, z, t) &= m_{xy}(x, y, z, 0) e^{i\phi(x, y, z, t)} \\ &= m_{xy}(x, y, z, 0) e^{i\omega_0 t} e^{-2\pi i(xk_x(t) + yk_y(t) + zk_z(t))}.\end{aligned}\tag{2.6}$$

这里 $m_{xy} = m_x + im_y$ 。然后我们对视野内 (field of view, 简称 FOV) 所有原子的磁场

密度进行积分

$$\begin{aligned} M_{xy} &= \int_{\text{FOV}} m_{xy}(x, y, z, t) d\mathbf{r} \\ &= e^{i\omega_0 t} \int_{\text{FOV}} m_{xy}(x, y, z, 0) e^{-2\pi i(xk_x(t)+yk_y(t)+zk_z(t))} d\mathbf{r}. \end{aligned} \quad (2.7)$$

我们记 $x(\mathbf{r}) = m_{xy}(x, y, z, 0)$ 并且去掉检测阶段的 $e^{i\omega_0 t}$, 得到最终的 MR 信号:

$$b(\mathbf{k}(t)) = \int_{\text{FOV}} x(\mathbf{r}) e^{-2\pi i \mathbf{k}(t) \cdot \mathbf{r}} d\mathbf{r}. \quad (2.8)$$

从方程 (2.8) 中可以看出, MR 信号大小与图像的 Fourier 变换成正比。由于测量得到的信号是 Fourier 域的数据, 因此要最终的图像 $x(\mathbf{r})$, 需要从采集到的数据 $b(\mathbf{k}(t))$ 中恢复出 $x(\mathbf{r})$ 。

为了从采集到的数据 $b(\mathbf{k}(t))$ 中恢复出 $x(\mathbf{r})$, 我们假设以下两点。第一, 采集得到的 k-space 数据是有限的; 第二, 测量过程是带有噪声的。因此方程 (2.8) 可以写为:

$$b(\mathbf{k}_m) = \int_{\text{FOV}} x(\mathbf{r}) e^{-2\pi i \mathbf{k}_m \cdot \mathbf{r}} d\mathbf{r} + \epsilon(\mathbf{k}_m), m = 1, \dots, M. \quad (2.9)$$

由于 $b(\mathbf{r})$ 为连续形式, 而我们只有有限个 k-space 采样点, 因此这个重建过程是病态的。Nyquist 采样定理表明, 如果要精确地重建出信号, 采样率需要达到信号最高频率的两倍以上。这里我们假设采样率达到了 Nyquist 采样定理的要求, 称之为全采样。为了将图像重建出来, 我们可以对这个病态问题提出合理的假设。这里我们假设, $x(\mathbf{r})$ 属于有限维空间, 则 $b(\mathbf{r})$ 可以用这个空间的一组基来表示:

$$b(\mathbf{r}) = \sum_{n=1}^N b_n \psi_n(\mathbf{r}). \quad (2.10)$$

这里 $\psi_n(\mathbf{r}), n = 1, \dots, N$ 为基函数, N 为空间的维度。基函数的选择有很多种, 一般来说, 可以取 $\psi(\mathbf{r})$ 为 Dirac 函数或者矩形方程。将方程 (2.10) 和 $\psi(\mathbf{r})$ 带入方程 (2.9) 得到:

$$b(\mathbf{k}_m) = \sum_{n=1}^N A_{m,n} b_n + \epsilon(\mathbf{k}_m), m = 1, \dots, M. \quad (2.11)$$

这里 $A_{m,n} = e^{-2\pi i \mathbf{k}_m \cdot \mathbf{r}_n} \Psi(\mathbf{k}_m)$, $\Psi(\mathbf{k}_m)$ 为 $\psi(\mathbf{r})$ 在 \mathbf{k}_m 点的 Fourier 变换。将上式写成矩阵形式则为:

$$b = Ax + \epsilon, \quad (2.12)$$

其中 $b \in \mathbb{C}^M$ 为测量得到的数据, $A \in \mathbb{C}^{M \times N}$ 为采样矩阵, $x \in \mathbb{C}^N$ 为待重建图像, $\epsilon \in \mathbb{C}^M$ 为噪声。采样矩阵 A 取决于数据在 k-space 的采样轨迹 \mathbf{k} 和基函数 $\psi(\mathbf{r})$ 的选择。比如, 当 $\psi(\mathbf{r})$ 为 Dirac 函数, $M = N$ 并且所有采样点都在均匀 Cartetian 的网格上时, 采样矩阵 A 退化为 Fourier 变换矩阵。这时, 方程 (2.12) 可以通过快速 Fourier 变换 (FFT) 求解。

一般而言，如果采样矩阵 A 是列满秩的，即 $M \geq N$ ，方程 (2.12) 可以由最小二乘法 (least square) 求解

$$\hat{x} = \arg \min_x \|Ax - b\|_2^2. \quad (2.13)$$

方程 (2.13) 的解为：

$$\hat{x} = (A^* A)^{-1} A^* b. \quad (2.14)$$

其中， A^* 为 A 的共轭转置。当 $M \geq N$ 时，采样点的个数大于图像像素点的个数。由于 MR 成像速度慢，采集过多的数据会导致成像时间增长。因此在实际应用中，如平行成像、压缩感知等，通常是减少采样的个数来加快成像速度，即 $M < N$ ，这时方程 (2.13) 则变为病态问题。

在 MR 成像中，采样轨迹是通过调整梯度场 G_x, G_y 和 G_z 的大小和方向来实现的。在理想情况下，我们希望整个 k-space 可以被一条光滑的采样轨迹所覆盖。但在实际情况中，由于信号的衰减和主磁场 \mathbf{B}_0 的非均匀，我们一般需要多条采样轨迹来覆盖 k-space。这些采样轨迹合在一起称为采样模式。在第1.3小节中我们介绍过，MRI 中的采样模式可以分为 Cartesian 采样和非 Cartesian 两大类。其中 Cartesian 采样计算简便，在临床中应用最为广泛；而非 Cartesian 的采样方式，如径向采样和螺旋采样可以达到更高的下采样率，但是计算比 Cartesian 采样繁琐。

2.1.2 MR 成像序列

在这一小节中我们介绍 MR 中常用的成像序列。在上一小节我们提到，当射频场打开之后，磁化 \mathbf{M} 发生共振激发，沿着主磁场方向进动。当射频场关闭之后，磁化 \mathbf{M} 会从激发态慢慢恢复到热平衡态，这一过程成为驰豫 (relaxation)。我们将热平衡态的磁化大小记为 \mathbf{M}_0 ，磁化 \mathbf{M} 的垂直分量记为 M_z ，水平分量记为 M_{xy} 。驰豫的过程即为 M_z 慢慢恢复而 M_{xy} 逐渐消失的过程。于是根据方向的不同，我们将驰豫分为纵向驰豫和横向驰豫，他们满足以下 Bloch 方程^[143]：

$$\begin{cases} M_z(t) = \mathbf{M}_0(1 - e^{-t/T_1}), \\ M_{xy}(t) = \mathbf{M}_0 e^{-t/T_2}. \end{cases} \quad (2.15)$$

这里 T_1 为纵向驰豫时间，是 M_z 恢复到 $63\%\mathbf{M}_0$ 所需要的时间； T_2 是横向驰豫时间，是 M_{xy} 损失为原来的 37% 所需的时间。 T_1 和 T_2 是物质的固有属性，也是 MR 图像中对比度的主要来源。图2.6展示了横向驰豫和纵向驰豫的过程。 T_1 和 T_2 在定量 MR 分析中有着重要的作用，这一点我们后面再介绍。

前面我们介绍了射频场和梯度场在 MR 成像的作用，而 MR 序列即为不同射频脉冲序列和梯度脉冲序列的特定设置，每个不同的 MR 序列都会影响最终所成图像。MR 中常用的序列有两大类，一类是自旋回波序列 (spin echo, 简称 SE)，另一类是梯度回波序列 (gradient echo, 简称 GE)。

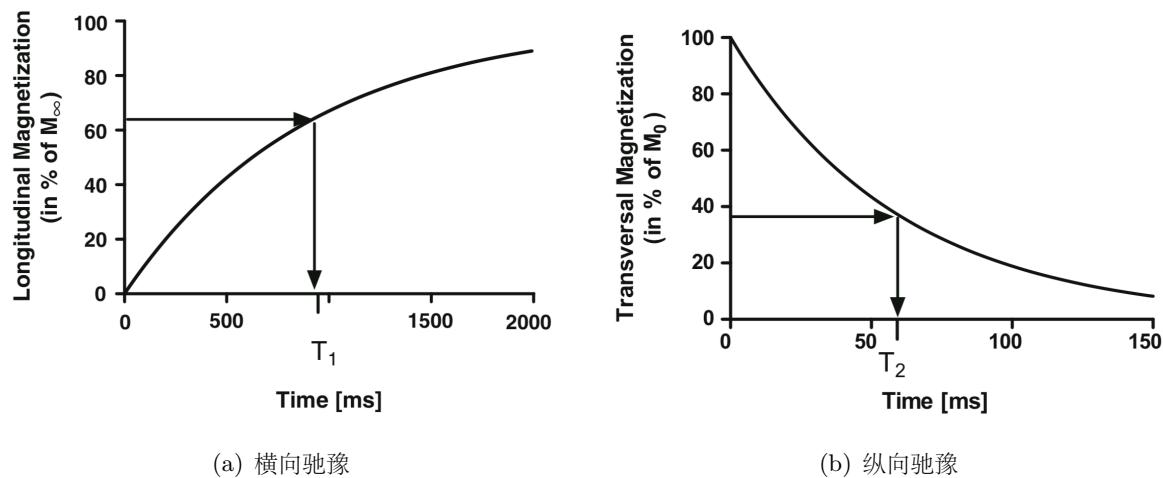


图 2.6 横向驰豫 (a) 与纵向驰豫 (b)。此图来自文献^[144]。

自旋回波是 MR 成像中的基本序列，也是最常用的序列。自旋回波的过程如图 2.7 所示。其中回波时间 (echo time, T_E) 是 90° 射频脉冲打开到收到回波的时间。自旋回波的特点是在 $T_E/2$ 时刻加入了 180° 的射频脉冲，使得磁化反向聚焦从而产生回波。而梯

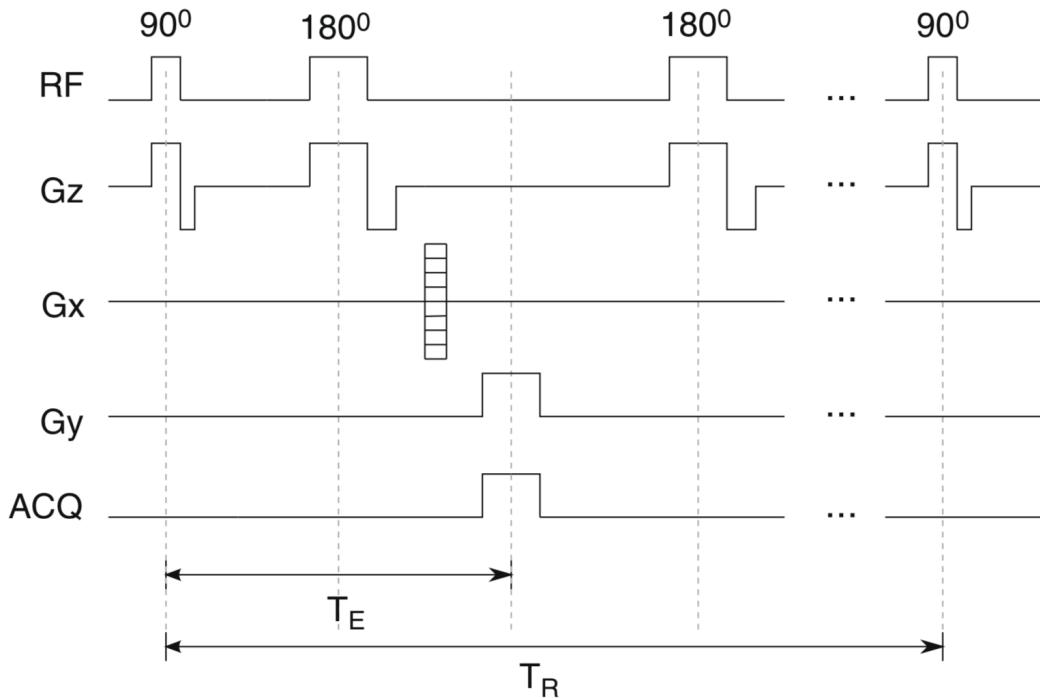


图 2.7 自旋回波序列。此图来自文献^[144]。

度回波序列与自旋回波不同，它是靠梯度场的反向来达成聚焦从而产生回波。梯度回波的过程如图 2.8 所示。其中，重复时间 (repetition time, T_R) 是连续两个射频脉冲的间隔时间，即序列的重复时间。临床中所得到的 MR 图像大部分都是通过这两种回波序列的改进或者组合而得到，因此它们在 MR 成像中有着重要的作用。

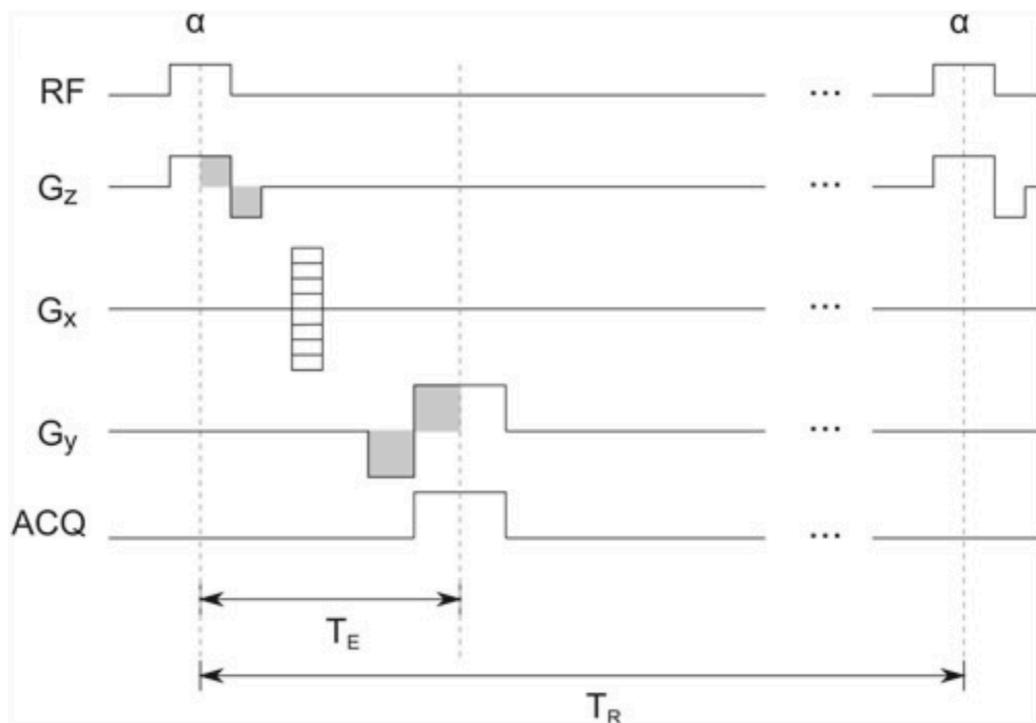


图 2.8 梯度回波序列。此图来自文献^[144]。

2.1.3 本节小结

在这一节中，我们主要介绍了 MR 成像的基本概念并推导了其数学模型。MR 成像可看成从 k-space 中采集数据并重建图像的过程，而不同的采样方式会对重建图像有着不同的影响，重建算法也会有所不同。MR 序列是 MRI 中的重要概念，序列的参数，如偏转角、相位角、重复时间和回波时间等，会对重建图像有着不同的影响。这些概念我们在后面的章节也要用到。

2.2 扩展相图

在这一小节中，我们简单介绍扩展相图的数学模型，详细的信息请参考文献^[129]。这有助于我们理解 MRI 中回波强度的计算和第5章中 snapMRF 的程序设计。

EPG 是用来描述和理解磁化在不同 MR 序列中响应的有力工具。在 EPG 中，梯度场、射频场、驰豫等对磁化的作用都可以用矩阵的简单计算来实现，因此在传统 MRI 和 MRF 中有着广泛的应用。相比与使用传统的 Bloch 方程来理解磁化响应的方法，EPG 可以快速并准确地量化回波强度，尤其当序列中存在多个梯度场时。

2.2.1 配置状态矩阵

EPG 模型中的第一个重要概念是用配置状态（configuration states）来表征梯度场使磁化失去相位和获得相位的过程。对于一个处于非中心位置 \mathbf{r} 的等色子，其受到的磁

场 \mathbf{M} 可分解为：

$$M_x(\mathbf{r}) = |\mathbf{M}| \cos \left(\gamma \mathbf{r} \cdot \int_0^t \mathbf{G}(t') dt' \right) = |\mathbf{M}| \cos(\mathbf{k} \cdot \mathbf{r}), \quad (2.16)$$

$$M_y(\mathbf{r}) = |\mathbf{M}| \sin \left(\gamma \mathbf{r} \cdot \int_0^t \mathbf{G}(t') dt' \right) = |\mathbf{M}| \sin(\mathbf{k} \cdot \mathbf{r}). \quad (2.17)$$

这里相位角 $\phi(\mathbf{r}) = \int_0^t \gamma \mathbf{G}(t') \cdot \mathbf{r} dt'$ 定义在给定时间 t 上，并且

$$\mathbf{k}(t) = \gamma \int_0^t \mathbf{G}(t') dt' \quad (2.18)$$

为失相的阶数 (order)。注意这里的阶数也就是上节所讨论的采样轨迹。将坐标从实数改变到复数，则上式变为：

$$\begin{aligned} M_+(\mathbf{r}) &= M_x(\mathbf{r}) + i M_y(\mathbf{r}) \\ &= |\mathbf{M}| e^{i\phi(\mathbf{r})} = |\mathbf{M}| e^{i\mathbf{k} \cdot \mathbf{r}} = (M_-)^*, \end{aligned} \quad (2.19)$$

$$\begin{aligned} M_-(\mathbf{r}) &= M_x(\mathbf{r}) - i M_y(\mathbf{r}) \\ &= |\mathbf{M}| e^{-i\phi(\mathbf{r})} = |\mathbf{M}| e^{-i\mathbf{k} \cdot \mathbf{r}} = (M_+)^*. \end{aligned} \quad (2.20)$$

通过这样的坐标转换可以更加有效的表征水平磁化。

由于在一个体素中可能会包含多个等色子，而不是单一的，在实际计算中，我们需要考虑所有等色子的影响。将 FOV 内的水平磁化进行积分得到：

$$\begin{aligned} \tilde{F}_+(\mathbf{k}) &= \int_{\text{FOV}} M_+(\mathbf{r}) e^{-i\mathbf{k} \cdot \mathbf{r}} d\mathbf{r} \\ \iff M_+(\mathbf{r}) &= \int_{\text{FOV}} \tilde{F}_+(\mathbf{r}) e^{i\mathbf{k} \cdot \mathbf{r}} d\mathbf{r}, \end{aligned} \quad (2.21)$$

$$\begin{aligned} \tilde{F}_-(\mathbf{k}) &= \int_{\text{FOV}} M_-(\mathbf{r}) e^{i\mathbf{k} \cdot \mathbf{r}} d\mathbf{r} \\ \iff M_-(\mathbf{r}) &= \int_{\text{FOV}} \tilde{F}_-(\mathbf{r}) e^{i\mathbf{k} \cdot \mathbf{r}} d\mathbf{r}. \end{aligned} \quad (2.22)$$

这个方程表明由梯度场导致的失相可以由配置状态 F_+ 和 F_- 来解释，因此所有等色子的状态也可以很容易地计算。

2.2.2 分区状态方法

EPG 模型中第二个重要的概念是分区状态 (partition state) 方法，用于描述射频场对于磁化的影响。根据 Bloch 方程，射频场对磁化 $[M_+, M_-, M_z]$ 的作用可以被定义为：

$$T_x(\alpha) = \begin{bmatrix} \cos^2 \frac{\alpha}{2} & \sin^2 \frac{\alpha}{2} & -i \sin \alpha \\ \sin^2 \frac{\alpha}{2} & \cos^2 \frac{\alpha}{2} & i \sin \alpha \\ -\frac{i}{2} \sin \alpha & \frac{i}{2} \sin \alpha & \cos \alpha \end{bmatrix} \quad (2.23)$$

和

$$T_z(\phi) = \begin{bmatrix} e^{i\phi} & 0 & 0 \\ 0 & e^{-i\phi} & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (2.24)$$

其中 α 是偏转角, ϕ 是相位角。将方程 (2.23) 和方程 (2.24) 经过以下变换

$$T_\Phi(\alpha) = T_z(\Phi)T_x(\alpha)T_z(-\Phi),$$

可以得到初始相位为 Φ 的射频场对磁化的作用矩阵, 我们用 T 来表示:

$$\begin{bmatrix} M_+ \\ M_- \\ M_z \end{bmatrix}^+ = \begin{bmatrix} \cos^2 \frac{\alpha}{2} & e^{2i\Phi} \sin^2 \frac{\alpha}{2} & -ie^{i\Phi} \sin \alpha \\ e^{-2i\Phi} \sin^2 \frac{\alpha}{2} & \cos^2 \frac{\alpha}{2} & ie^{-i\Phi} \sin \alpha \\ -\frac{i}{2}e^{-i\Phi} \sin \alpha & \frac{i}{2}e^{i\Phi} \sin \alpha & \cos \alpha \end{bmatrix} \times \begin{bmatrix} M_+ \\ M_- \\ M_z \end{bmatrix}^- . \quad (2.25)$$

这里“-”和“+”的意思是在射频场作用之前和之后。方程 (2.25) 给出了射频场对于单一等色子的作用。在射频场的作用之后, 磁化可以分为三个部分, 即失相部分 (M_+), 得相部分 (M_-) and 垂直部分 (M_z)。

将配置状态矩阵和分区状态方法结合起来, 就可以得到射频场对状态矩阵的作用:

$$\begin{bmatrix} \tilde{F}_+(k) \\ \tilde{F}_-(-k) \\ \tilde{Z}(k) \end{bmatrix}^+ = \begin{bmatrix} \cos^2 \frac{\alpha}{2} & e^{2i\Phi} \sin^2 \frac{\alpha}{2} & -ie^{i\Phi} \sin \alpha \\ e^{-2i\Phi} \sin^2 \frac{\alpha}{2} & \cos^2 \frac{\alpha}{2} & ie^{-i\Phi} \sin \alpha \\ -\frac{i}{2}e^{-i\Phi} \sin \alpha & \frac{i}{2}e^{i\Phi} \sin \alpha & \cos \alpha \end{bmatrix} \times \begin{bmatrix} \tilde{F}_+(k) \\ \tilde{F}_-(-k) \\ \tilde{Z}(k) \end{bmatrix}^- . \quad (2.26)$$

这里

$$\begin{aligned} \tilde{Z}(k) &= \int_{\text{FOV}} M_z(\mathbf{r}) e^{-i\mathbf{k}\cdot\mathbf{r}} d\mathbf{r} \\ \iff M_z(\mathbf{r}) &= \int_{\text{FOV}} \tilde{Z}(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{r}} d\mathbf{k}. \end{aligned} \quad (2.27)$$

2.2.3 其他因素的作用和最终模型

在上面两个小节, 我们推导了射频场对磁化的作用矩阵 T 。除了射频场的作用, EPG 模型也可以很容易地表示其他因素对磁化的作用, 比如梯度场、驰豫等。这也是 EPG 模型的优点之一。由梯度场造成的水平失相可以用平移矩阵 H 来表征:

$$H(\Delta k) : \quad \tilde{F}_k \rightarrow \tilde{F}_{k+\Delta k} \quad \text{and} \quad \tilde{Z}_k \rightarrow \tilde{Z}_k, \quad (2.28)$$

这里 $\Delta k = \gamma \int_{t'=0}^t \mathbf{G}(t') dt'$ 。驰豫作用可以由下面的驰豫算子 E 实现。对于 $k \neq 0$ 的状态：

$$\begin{aligned} E &= \begin{bmatrix} E_2 & 0 & 0 \\ 0 & E_2 & 0 \\ 0 & 0 & E_1 \end{bmatrix} \\ &= \begin{bmatrix} e^{-\tau/T_2} & 0 & 0 \\ 0 & e^{-\tau/T_2} & 0 \\ 0 & 0 & e^{-\tau/T_1} \end{bmatrix}. \end{aligned} \quad (2.29)$$

这里 τ 是驰豫进行的时间长度。当 $k = 0$ 时，需要在 $\tilde{Z}(0)$ 上进行额外的 T_1 驰豫，方程变为：

$$\begin{bmatrix} \tilde{F}_+(k) \\ \tilde{F}_+(-k) \\ \tilde{Z}(k) \end{bmatrix}^+ = \begin{bmatrix} E_2 & 0 & 0 \\ 0 & E_2 & 0 \\ 0 & 0 & E_1 \end{bmatrix} \begin{bmatrix} \tilde{F}_+(k) \\ \tilde{F}_+(-k) \\ \tilde{Z}(k) \end{bmatrix}^- + \begin{bmatrix} 0 \\ 0 \\ M_0(1 - E_1) \end{bmatrix}. \quad (2.30)$$

到现在为止，我们已经定义了射频场的作用矩阵 T ，梯度场的失相矩阵 H 以及驰豫矩阵 E 。那么对于一个给定的 MR 序列，磁化在其中的相应可以表示为：

$$\Gamma(t) = \cdots E \cdot H \cdot T \cdot E \cdot H \cdot T \cdot \Gamma (t=0), \quad (2.31)$$

其中状态矩阵按照以下形式存储：

$$\Gamma = \begin{bmatrix} \tilde{F}_0 & \tilde{F}_1 & \tilde{F}_2 & \tilde{F}_3 & \tilde{F}_4 & \cdots \\ \tilde{F}_0^* & \tilde{F}_{-1}^* & \tilde{F}_{-2}^* & \tilde{F}_{-3}^* & \tilde{F}_{-4}^* & \cdots \\ \tilde{Z}_0 & \tilde{Z}_1 & \tilde{Z}_2 & \tilde{Z}_3 & \tilde{Z}_4 & \cdots \end{bmatrix}. \quad (2.32)$$

2.2.4 本节小结

在本小节中，我们简单回顾了 EPG 模型，并推导了射频场、梯度场以及驰豫对于磁化的作用矩阵。注意这里我们只给出了这三种因素的作用，但 EPG 模型也可以加入其他因素的影响，如扩散和运动。

2.3 本章总结

在本章中，我们首先回顾了 MR 成像的基本原理和数学模型，并介绍了一些基本概念，如射频场、梯度场、驰豫、序列等。这有助于我们理解基于压缩感知的 MR 成像模型。之后我们简单回顾了扩展相图模型，并推导了射频场、梯度场和驰豫对于磁化的作用矩阵。这有助于我们理解第5章的程序设计。

3 动态 MR 图像的低秩和稀疏分解的高阶变分模型和算法

动态 MRI 是 MR 成像中的重要的技术，在医学临床中有着广泛的应用，比如心脏电影成像、磁共振动态对比增强 (DCE-MRI)^[2] 等，其目的是可视化形态学或者对比度随时间的变化。然而，传统 dMRI 方法成像速度慢，信噪比低，并且很难在空间分辨率和时间分辨率之间取得平衡^[105,145]。动态 MRI 重建中最重要的问题是如何平衡时间和空间分辨率，这是由于其数据采集速度慢所决定的。在对动态 MR 图像进行采样时，如果在每一帧图像上采集了更多的数据，那么在同样的时间内，我们就只能获得更少帧数的数据，造成图像时间分辨率的降低；反之，如果我们保证了时间分辨率，尽可能多的在每一帧图像上都采集数据，那么必然造成在每一帧图像上采集数据的减少，造成空间分辨率的降低。因此，动态 MRI 的时间分辨率和空间分辨率之间的平衡很难把握。

压缩感知理论^{[146][9]}在过去十年中一直是图像处理领域的一个热门话题，并被广泛应用于图像处理领域，尤其是动态 MRI^[13,49,109,147–150]领域。压缩感知理论证明，我们可以从下采样的 Fourier 数据中（通常被称为 k-space 数据）准确地重建出动态 MR 图像，这可以显著地减少扫描时间^[13,151]。因此，压缩感知理论可以在加速动态 MR 成像的同时保证重建图像的时间和空间分辨率。

目前压缩感知在动态 MRI^[49,109,147–149] 中已经有广泛的应用，我们在小节1.2中已经有所介绍。这里我们再来看看一下压缩感知重建动态 MR 的模型。记动态 MR 图像为 $X \in \mathbb{C}^{N_1 \times N_2 \times d}$ ，这里动态图像 X 有 d 帧图像，每帧图像的维度为 $N_1 \times N_2$ 。则动态 MR 成像相过程相当于在噪声的干扰下，在 kt-space 中进行动态采样

$$B = AX + \epsilon,$$

其中 $A = \mathcal{M} \cdot \mathcal{F}$ 是采样算子， \mathcal{F} 是作用在每一帧图像上的二维 Fourier 变换算子， \mathcal{M} 是作用在每一帧图像上的二维采样模式， ϵ 是加性高斯白噪声。于是基于压缩感知的动态 MR 重建模型可以表示为：

$$\min_X \frac{1}{2} \|AX - B\|_{\text{F}}^2 + \alpha \|\mathcal{S}X\|_1, \quad (3.1)$$

其中 \mathcal{S} 是某个稀疏变换， α 是平衡数据项和稀疏项的参数。

对于动态 MR 图像而言，稀疏变换 \mathcal{S} 的选择至关重要，而且很具技巧性，主要有以下三种策略。第一种也是最简单的一种是仅仅在模型中利用动态 MR 图像的稀疏性。kt-SPARSE^[13] 模型利用了时间和空间上的稀疏变换，其模型为：

$$\min_X \frac{1}{2} \|AX - B\|_{\text{F}}^2 + \alpha \|\mathcal{W}X\|_1 + \beta \|\mathcal{F}_t X\|_1, \quad (3.2)$$

其中 \mathcal{W} 是空间方向上的二维小波变换，而 \mathcal{F}_t 是时间上的一维 Fourier 变换。kt-SPARSE 在心脏电影成像中的表现很好，这是因为 Fourier 变换可以稀疏化周期性运动，比如心

脏的搏动。但是，对于其他动态 MR 图像而言，比如胸部 DCE-MRI，胸部组织和肿瘤区域通常不会表现出周期性，因此 kt-SPARSE 在这类数据中的表现有限。iGrasp^[44] 结合了压缩感知、平行成像和黄金角度径向采样的思想，提出了只含有一个稀疏约束的模型：

$$\min_X \frac{1}{2} \|AX - B\|_F^2 + \alpha \|\nabla_t X\|_1, \quad (3.3)$$

其中 ∇_t 是时间方向的梯度算子。iGrasp 在图像时间方向的稀疏性占主导地位时表现很好。但是对于绝大多数动态 MR 图像来说，空间分辨率和时间分辨率都是十分重要的——高空间分辨率可以帮助医生更好地从视觉上理解图像，而高时间分辨率在进行图像的量化分析时十分必要。因此仅仅利用时间方向的稀疏变换是不足的。

第二种策略是在模型中考虑动态图像的低秩性，目的是找到一个既稀疏又低秩的解。在这类方法中，动态 MR 图像通常需要写成二维矩阵形式，矩阵的每一行对应一个像素点，矩阵的每一列对应一帧图像。我们称这种形式的矩阵为 Casorati 矩阵，其维度为 $N_1 N_2 \times d$ 。由于动态图像帧与帧之间有着很强的相关性，Casorati 矩阵通常可以认为是低秩的。kt-SLR 模型^[49] 不仅利用时间和空间方向的梯度算子，也利用了非凸的 Schatten p -拟范数。其模型为：

$$\min_X \frac{1}{2} \|AX - B\|_F^2 + \alpha \|X\|_{3DTV} + \beta \|X\|_p. \quad (3.4)$$

这里 3DTV 算子的定义为：

$$\|\cdot\|_{3DTV} = \|\nabla_3 \cdot\|_1,$$

而

$$|\nabla_3 \cdot| = \sqrt{(\nabla_x \cdot)^2 + (\nabla_y \cdot)^2 + \mu(\nabla_t \cdot)^2},$$

其中 ∇_x , ∇_y 和 ∇_t 分别为 x , y 和 t 方向上的梯度算子， μ 是平衡空间稀疏性与时间稀疏性的参数， $\|\cdot\|_p$ 是 Schatten p -拟范数，并且 $0 < p < 1$ 。kt-SLR 模型可以提高重建图像的时间分辨率，在心脏灌注图像和体模图像中有良好的表现。但是由于 kt-SLR 模型使用了 TV 正则项，重建的图像经常会产生阶梯效应。另外，当 $0 < p < 1$ 时，Schatten p -拟范数是非凸的，计算一个稳定的数值解会更加困难。

最后一种策略是利用图像分解的思想，将动态 MR 图像分解为两个部分的加和，即稀疏部分和低秩部分，并且分别用不同的稀疏项来约束。对于动态图像而言，一方面其背景部分通常随时间变化缓慢，即在时间方向高度相关，这导致了背景部分的低秩性；另一方面其前景部分通常随着时间变化，而且前景部分相对于整幅图像通常比例很小，这导致了前景部分的稀疏性。这样处理的好处是，将背景部分从动态图像中减掉之后，剩下的前景部分会比原图像更加稀疏，因此也更加符合压缩感知理论的条件。Otazo 等^[109] 利用这个想法，提出了 L+S 模型：

$$\min_{L,S} \frac{1}{2} \|A(L + S) - B\|_F^2 + \alpha \|\nabla_t S\|_1 + \beta \|L\|_*, \quad (3.5)$$

其中 $\|\cdot\|_*$ 为核范数, L 和 S 分别为低秩部分和稀疏部分。类似的, Trémoulinéac 等^[11]提出了模型 kt-RPCA, 其利用时间方向的 Fourier 变换来约束稀疏部分:

$$\min_{L,S} \frac{1}{2} \|A(L+S) - B\|_F^2 + \alpha \|\mathcal{F}_t S\|_1 + \beta \|L\|_*. \quad (3.6)$$

但与模型 iGrasp 一样, $L+S$ 与 kt-RPCA 也仅仅利用了时间方向的稀疏变换。此外, 由于模型 kt-RPCA 中使用了时间方向的 Fourier 变换, 它只在心脏数据中有效果。Schloegl 等^[61] 提出了卷积下确界 TGV 泛函 ICTGV, 其定义为:

$$\text{ICTGV}_{\alpha,\beta}^2(X) = \inf_{X=X_1+X_2} \text{TGV}_{\alpha_1}^2(X_1) + \beta \text{TGV}_{\alpha_2}^2(X_2), \quad (3.7)$$

相对应的重建模型为:

$$\min_X \frac{1}{2} \|AX - B\|_F^2 + \text{ICTGV}_{\alpha,\beta}^2(X). \quad (3.8)$$

TGV 泛函的概念在文献^[60] 中首次提出, 是经典 TV 泛函的推广。TGV 泛函中含有高阶导数项, 可以更好地表示图像中的光滑区域, 因此可以有效地抑制阶梯效应。虽然模型 ICTGV 中没有明确地将图像分解为低秩部分和稀疏部分, 但是 ICTGV 可以自适应地将动态图像分解为两个部分, 一部分包含更多时间正则项, 而另一部分包含更多动态变化信息。因此我们也将 ICTGV 归入图像分解的方法。ICTGV 可以最优地平衡时间和空间正则项, 并且在心脏电影和心脏造影图像上表现良好, 但其在胸部 DCE-MRI 上的表现是未知的。

综上所述, 以上针对动态 MR 图像的压缩感知重建模型要么会导致阶梯效应, 要么只利用了时间稀疏变换。为了提高重建质量, 抑制重建图像中的阶梯效应, 并且利用图像分解模型的思想, 我们提出了利用二阶时空 TGV 泛函来约束稀疏部分, 利用核范数来约束低秩部分的新模型。具体来说, 我们利用核范数来建模时间高度相关的背景, 利用二阶时空 TGV 来刻画背景之上的动态信息。

3.1 模型的提出

3.1.1 低秩矩阵补全

最近几年, 低秩矩阵补全算法被广泛应用于动态 MR 图像重建中。其主要的想法是将动态 MR 图像看成一个时空矩阵, 矩阵的每一行对应动态图像的一个像素点, 矩阵的每一列对应动态图像的一帧。在文献^[110] 中, Candès 等建立了低秩和稀疏分解模型的数学基础, 即 RPCA。给定一个 Casorati 矩阵 $X \in \mathbb{C}^{N_1 N_2 \times d}$, RPCA 描述了以下优化问题:

$$\min_{L,S} \|L\|_* + \alpha \|S\|_1, \quad s.t. \quad X = L + S. \quad (3.9)$$

这里核范数作为矩阵的秩的凸松弛, 其定义为矩阵奇异值 (σ_i) 的和:

$$\|\cdot\|_* = \sum_i \sigma_i.$$

Candès 指出，在特定的假设下，RPCA 可以完美地恢复出低秩和稀疏部分，其中 L 的秩需要很低并且 S 需要有很高的稀疏度。

压缩感知和低秩矩阵填充算法的结合可以进一步加快成像速度，并且将基于图像分解的模型应用于动态 MR 图像重建是自然而然的：低秩部分可以用来建模时间高度相关的背景，稀疏部分可以建模背景之上的动态信息，如心脏电影图像中心脏的运动和胸部 DCE-MRI 中造影剂的变化等。目前，已有一些工作^[109,111] 将 RPCA 的想法应用与动态 MR 重建中，其中最常用的稀疏变化是时间方向的梯度算子和 Fourier 变换。

3.1.2 二阶时空 TGV 泛函

TGV 泛函的数学基础由 Bredies 等^[60] 所建立，是经典 TV 泛函的推广。TGV 泛函可以表示图像中的光滑部分，并且可以在保持图像边缘的同时抑制阶梯效应的产生。因此 TGV 泛函在很多图像处理问题中都有应用，比如图像去噪^[36]，压缩感知^[36] 等。在本文中，由于所处理的是动态 MR 图像，我们集中讨论 \mathbb{C}^3 空间中的二阶 TGV 泛函 (TGV_α^2)。

首先我们回顾 TGV_α^2 的定义。对于 $u \in L^1(\Omega), \Omega \subset \mathbb{C}^3$, TGV_α^2 定义为：

$$\begin{aligned} \text{TGV}_\alpha^2(u) = \sup \Big\{ & \int_{\Omega} u \operatorname{div}^2 q \, dx \mid q \in C_c^2(\Omega, S^{3 \times 3}), \\ & \|q\|_\infty \leq \alpha_0, \|\operatorname{div} q\|_\infty \leq \alpha_1 \Big\}, \end{aligned} \quad (3.10)$$

这里 $\alpha = (\alpha_0, \alpha_1)$ 是正参数， $S^{3 \times 3}$ 是对称矩阵的集合， $C_c^2(\Omega, S^{3 \times 3})$ 表示带有紧支撑集且映射到 $S^{3 \times 3}$ 的二次连续可微函数。散度算子 $\operatorname{div} q \in C_c^1(\Omega, \mathbb{C}^3)$ 和 $\operatorname{div}^2 q \in C_c(\Omega)$ 的定义如下：

$$\begin{aligned} (\operatorname{div} q)_i &= \sum_{j=1}^3 \frac{\partial q_{ij}}{\partial x_j}, \\ \operatorname{div}^2 q &= \sum_{i=1}^3 \frac{\partial^2 q_{ii}}{\partial x_i^2} + 2 \sum_{i < j} \frac{\partial q_{ij}}{\partial x_i \partial x_j}, \end{aligned}$$

其中 $q \in C_c(\Omega, S^{3 \times 3})$ 和 $p \in C_c(\Omega, \mathbb{C}^3)$ 的无穷范数为

$$\|q\|_\infty = \sup_{x \in \Omega} \left(\sum_{i=1}^3 |q_{ii}(x)|^2 + 2 \sum_{i < j} |q_{ij}(x)|^2 \right)^{1/2}$$

和

$$\|p\|_\infty = \sup_{x \in \Omega} \left(\sum_{i=1}^3 |p_i(x)|^2 \right)^{1/2}.$$

下面我们回顾一下 TGV_α^2 的基本性质。根据 TGV_α^2 的定义 (3.10)，我们可以得知其在空间 $L^p(\Omega)(1 \leq p < \infty)$ 中是适定的、凸的和下半连续的，并且具有旋转不变性和平移不变性。我们记有界二阶 TGV 的函数组成的空间为

$$\text{BGV}_\alpha^2(u) = \{u \in L^1(\Omega) \mid \text{TGV}_\alpha^2(u) < \infty\},$$

其范数为

$$\|u\|_{\text{BGV}_\alpha^2} = \|u\|_1 + \text{TGV}_\alpha^2(u).$$

我们已知 BVG_α^2 是一个 Banach 空间，并且 TGV_α^2 的核空间为：

$$\ker(\text{TGV}_\alpha^2) = \{u(x) = Kx + b \mid x \in \Omega, K \in \mathbb{C}^2, b \in \mathbb{C}\}.$$

可以看出 TGV_α^2 的核空间由仿射函数组成，所以基于 TGV_α^2 的模型可以很好地估计图像的光滑区域。而 TV 的核空间由常函数组成，只能估计分片常数的图像，因此基于 TV 的模型会产生阶梯效应，而基于 TGV_α^2 的模型则不会。

在本文中，我们使用 TGV_α^2 的等价定义：

$$\text{TGV}_\alpha^2(u) = \min_{\omega \in \text{BD}(\Omega)} \alpha_1 \|\nabla_3 u - \omega\|_1 + \alpha_0 \|\mathcal{E}(\omega)\|_1,$$

其中 $\mathcal{E}(\omega) = \frac{1}{2}(\nabla_3 \omega + \nabla_3 \omega^T)$ 是对称梯度算子， $\text{BD}(\Omega)$ 是有界形变函数空间。

3.1.3 基于二阶时空 TGV 和核范数的模型

基于图像分解的思想，利用二阶时空 TGV 和核范数，我们提出模型：

$$\min_{L,S} \frac{1}{2} \|A(L + S) - B\|_F^2 + \text{TGV}_\alpha^2(S) + \beta \|L\|_*. \quad (3.11)$$

其中核范数用来建模时间高度相关的背景，二阶时空 TGV 泛函用来建模背景之上稀疏的动态信息。可以看出，模型是凸的，适定的，并且对于动态 MR 图像而言是自然而然的。一方面，MR 图像通常在梯度域中是稀疏的，而 TGV_α^2 包含了一阶和二阶导数信息，可以很好地稀疏化 MR 图像，并且从原动态图像中减掉背景之后，剩下的动态信息会比原图像更加稀疏。另一方面，动态图像的背景在时间方向是高度相关的，因此核范数可以完美地建模背景部分。

3.2 模型的离散化和算法

在这一节，我们首先给出 TGV_α^2 的离散格式，然后给出 Primal-Dual 算法来求解模型的流程。

3.2.1 离散化模型

为了求解模型，我们首先给出 TGV_α^2 的离散格式。本节的离散化可参照文献^[32,36]。首先将区域 $\Omega \subset \mathbb{C}^3$ 离散为网格：

$$\Omega = \{(i, j, k) \mid 1 \leq i \leq N_1, 1 \leq j \leq N_2, 1 \leq k \leq d\},$$

其中 N_1 , N_2 和 d 是动态图像的维度。假设 $U = \mathbb{C}^{N_1 \times N_2 \times d}$, $V = U \times U \times U = U^3$ 和 $W = U \times U \times U \times U \times U \times U = U^6$ 是三个有限维空间。定义空间 U 中的标量内积

$\langle u, u' \rangle_U = \sum_{(i,j,k) \in \Omega} u_{i,j,k} u'_{i,j,k}$, $u, u' \in U$, 范数 $\|u\|_U = \sqrt{\langle u, u \rangle_U}$ 。同样的, V 和 W 中的标量内积分别为 $\langle v, v' \rangle_V = \sum_{i=1}^3 \langle v_i, v'_i \rangle_U$ 和 $\langle w, w' \rangle_W = \sum_{i=1}^3 \langle w_i, w'_i \rangle_U + 2 \sum_{i=4}^6 \langle w_i, w'_i \rangle_U$ 。

在这里, 用有限差分来离散 TGV_α^2 , 空间步长为 1, 时间步长由参数 μ 决定。向前向后差分算子记作:

$$\begin{aligned} (\partial_x^+ u)_{i,j,k} &= \begin{cases} u_{i+1,j,k} - u_{i,j,k}, & 1 \leq i < N_1, \\ 0, & i = N_1, \end{cases} \\ (\partial_y^+ u)_{i,j,k} &= \begin{cases} u_{i,j+1,k} - u_{i,j,k}, & 1 \leq j < N_2, \\ 0, & j = N_2, \end{cases} \\ (\partial_t^+ u)_{i,j,k} &= \begin{cases} (u_{i,j,k+1} - u_{i,j,k})/\mu, & 1 \leq k < d, \\ 0, & k = d, \end{cases} \end{aligned}$$

和

$$\begin{aligned} (\partial_x^- u)_{i,j,k} &= \begin{cases} u_{1,j,k}, & i = 1, \\ u_{i,j,k} - u_{i-1,j,k}, & 1 < i < N_1, \\ -u_{m-1,j,k}, & i = N_1, \end{cases} \\ (\partial_y^- u)_{i,j,k} &= \begin{cases} u_{i,1,k}, & j = 1, \\ u_{i,j,k} - u_{i,j-1,k}, & 1 < j < N_2, \\ -u_{i,n-1,k}, & j = N_2. \end{cases} \\ (\partial_t^- u)_{i,j,k} &= \begin{cases} u_{i,j,1}/\mu, & k = 1, \\ (u_{i,j,k-1} - u_{i,j,k})/\mu, & 1 < k < d, \\ -u_{i,j,k-1}/\mu, & k = d. \end{cases} \end{aligned}$$

梯度算子 ∇_3 , 对称梯度算子 \mathcal{E} 和它们相对应的散度算子为:

$$\nabla_3 : U \rightarrow V, \quad \nabla_3 u = \begin{pmatrix} \partial_x^+ u \\ \partial_y^+ u \\ \partial_t^+ u \end{pmatrix},$$

$$\mathcal{E} : V \rightarrow W, \quad \mathcal{E}(v) = \begin{pmatrix} \partial_x^+ v_1 \\ \partial_y^+ v_2 \\ \partial_t^+ v_3 \\ \frac{1}{2}(\partial_y^+ v_1 + \partial_x^+ v_2) \\ \frac{1}{2}(\partial_t^+ v_2 + \partial_y^+ v_3) \\ \frac{1}{2}(\partial_t^+ v_1 + \partial_x^+ v_3) \end{pmatrix},$$

和

$$\operatorname{div} : V \rightarrow U, \quad \operatorname{div} v = \partial_x^- v_1 + \partial_y^- v_2 + \partial_t^- v_3,$$

$$\operatorname{div}_2 : W \rightarrow V, \quad \operatorname{div}_2 w = \begin{pmatrix} \partial_x^- w_1 + \partial_y^- w_4 + \partial_t^- w_6 \\ \partial_x^- w_4 + \partial_y^- w_2 + \partial_t^- w_5 \\ \partial_x^- w_6 + \partial_y^- w_5 + \partial_t^- w_3 \end{pmatrix}.$$

由定义, 散度算子是梯度算子的复共轭, 即 $\nabla_3^* = -\operatorname{div}$ 和 $\mathcal{E}^* = -\operatorname{div}_2$ 。为了定义离散形式的 $\operatorname{TGV}_\alpha^2$, 还需如下的 L^1 和 L^∞ 范数:

$$u \in U : \quad \|u\|_1 = \sum_{(i,j,k) \in \Omega} |u_{i,j,k}|,$$

$$\|u\|_\infty = \max_{(i,j,k) \in \Omega} |u_{i,j,k}|,$$

$$v \in V : \quad \|v\|_1 = \sum_{(i,j,k) \in \Omega} (v_{1i,j,k}^2 + v_{2i,j,k}^2 + v_{3i,j,k}^2)^{1/2},$$

$$\|v\|_\infty = \max_{(i,j,k) \in \Omega} (v_{1i,j,k}^2 + v_{2i,j,k}^2 + v_{3i,j,k}^2)^{1/2},$$

$$w \in W : \quad \|w\|_1 = \sum_{(i,j,k) \in \Omega} (w_{1i,j,k}^2 + w_{2i,j,k}^2 + w_{3i,j,k}^2 + 2w_{4i,j,k}^2 + 2w_{5i,j,k}^2 + 2w_{6i,j,k}^2)^{1/2},$$

$$\|w\|_\infty = \max_{(i,j,k) \in \Omega} (w_{1i,j,k}^2 + w_{2i,j,k}^2 + w_{3i,j,k}^2 + 2w_{4i,j,k}^2 + 2w_{5i,j,k}^2 + 2w_{6i,j,k}^2)^{1/2}.$$

接下来, 我们给出模型 (3.11) 的离散格式。假设采样矩阵 $A : U \rightarrow Y$, Y 为有限维度的 Hilbert 空间, 并且为离散形式。则模型 (3.11) 可以写成:

$$\begin{aligned} \min_{(S,w,L) \in U \times V \times U} & \alpha_1 \|\nabla_3 S - w\|_1 + \alpha_0 \|\mathcal{E}(w)\|_1 + \beta \|L\|_* \\ & + \frac{1}{2} \|A(L + S) - B\|_F^2, \end{aligned} \tag{3.12}$$

其相应的鞍点问题为：

$$\begin{aligned} \min_{(S,w,L) \in U \times V \times U} & \max_{(p,q,\lambda) \in V \times W \times Y} \langle \nabla_3 S - w, p \rangle + \langle \mathcal{E}(w), q \rangle + \beta \|L\|_* \\ & + \langle A(L + S) - B, \lambda \rangle - \frac{1}{2} \|\lambda\|_F^2 \\ & - \mathcal{I}_{\|\cdot\|_\infty \leq \alpha_1}(p) - \mathcal{I}_{\|\cdot\|_\infty \leq \alpha_0}(q). \end{aligned} \quad (3.13)$$

3.2.2 Primal-Dual 算法及收敛条件

在本文中，我们使用 Primal-Dual 算法^[32] 来求解模型。Primal-Dual 算法已经被广泛地应用到寻找凸-凹鞍点问题的极大极小问题中：

$$\min_{x \in \mathcal{X}} \max_{y \in \mathcal{Y}} \langle \mathcal{K}x, y \rangle + f(x) - g(y), \quad (3.14)$$

其中 \mathcal{X} 和 \mathcal{Y} 是 Hilbert 空间, $\mathcal{K} : \mathcal{X} \rightarrow \mathcal{Y}$ 是线性连续映射，并且泛函 $f : \mathcal{X} \rightarrow (-\infty, \infty]$ 和 $g : \mathcal{Y} \rightarrow (-\infty, \infty]$ 是适定的、凸的和下半连续的。为了应用 Primal-Dual 算法，我们定义预解算子 $(I + \tau \partial f)^{-1}$ 和 $(I + \sigma \partial g)^{-1}$ ，其闭形式的表达分别为：

$$\hat{x} = (I + \tau \partial f)^{-1}(\bar{x}) = \arg \min_{x \in \mathcal{X}} \frac{\|x - \bar{x}\|_{\mathcal{X}}^2}{2} + \tau f(x)$$

和

$$\hat{y} = (I + \sigma \partial g)^{-1}(\bar{y}) = \arg \min_{y \in \mathcal{Y}} \frac{\|y - \bar{y}\|_{\mathcal{Y}}^2}{2} + \sigma g(y),$$

这里 $\tau, \sigma > 0$ 为迭代步长。于是，给定初始点 $(x^0, y^0) \in \mathcal{X} \times \mathcal{Y}$ 并且令 $\bar{x}^0 = x^0$ ，Primal-Dual 算法有如下迭代步骤：

$$\begin{cases} y^{n+1} = (I + \sigma \partial g)^{-1}(y^n + \sigma \mathcal{K} \bar{x}^n), \\ x^{n+1} = (I + \tau \partial f)^{-1}(x^n - \tau \mathcal{K}^* y^{n+1}), \\ \bar{x}^{n+1} = 2x^{n+1} - x^n. \end{cases} \quad (3.15)$$

若 $\tau \sigma \|\mathcal{K}\|^2 < 1$ ，则算法收敛。

接下来，我们将模型 (3.13) 转化成鞍点问题的结构 (3.14)。令

$$\mathcal{X} = U \times V \times U, \quad \mathcal{Y} = V \times W \times Y, \quad \mathcal{K} = \begin{pmatrix} \nabla_3 & -I & 0 \\ 0 & \mathcal{E} & 0 \\ A & 0 & A \end{pmatrix},$$

并且

$$\begin{aligned} f(x) &= \beta \|L\|_*, \\ g(y) &= \langle B, \lambda \rangle + \frac{\|\lambda\|_F^2}{2} + \mathcal{I}_{\|\cdot\|_\infty \leq \alpha_1}(p) + \mathcal{I}_{\|\cdot\|_\infty \leq \alpha_0}(q). \end{aligned}$$

于是模型 (3.13) 的求解过程如算法5所示。

算法 5 二阶时空 TGV 和低秩分解模型的 Primal-Dual 算法。

初始化: σ, τ, S_0, L_0 , 令 $L_0 = A^*B, S_0 = 0$;

迭代: 根据以下步骤更新参数:

1. $p^{n+1} = \mathcal{P}_{\alpha_1}(p^n + \sigma(\nabla \bar{S}^n - \bar{w}^n))$;
2. $q^{n+1} = \mathcal{P}_{\alpha_0}(q^n + \sigma \mathcal{E}(\bar{w}^n))$;
3. $\lambda^{n+1} = (\lambda^n + \sigma(A(\bar{L}^n + \bar{S}^n) - B))/(1 + \sigma)$;
4. $S^{n+1} = S^n - \tau(A^*\lambda^{n+1} - \text{div}_1 p^{n+1})$;
5. $w^{n+1} = w^n + \tau(\text{div}_2 q^{n+1} + p^{n+1})$;
6. $L^{n+1} = \mathcal{S}_\beta(L^n - \tau A^* r^{n+1})$;
7. $\bar{S}^{n+1} = 2S^{n+1} - S^n$;
8. $\bar{w}^{n+1} = 2w^{n+1} - w^n$;
9. $\bar{L}^{n+1} = 2L^{n+1} - L^n$;

直到收敛, 返回 x^{n+1} 。

算法步骤中的投影算子 \mathcal{P}_α 为:

$$\hat{t} = \mathcal{P}_\alpha(\bar{t}) = \arg \min_{\|t\|_\infty \leq \alpha} \frac{\|t - \bar{t}\|_F^2}{2} = \frac{\bar{t}}{\max(1, \frac{|\bar{t}|}{\alpha})}.$$

收缩算子 \mathcal{S}_β 为:

$$\hat{t} = \mathcal{S}_\beta(\bar{t}) = \arg \min_t \frac{\|t - \bar{t}\|_F^2}{2} + \beta \|t\|_* = U \mathcal{S}_\beta(\Sigma) V^T,$$

其中

$$\mathcal{S}_\beta(\Sigma) = \text{diag}\{\max(\sigma_i - \beta, 0)\}, i = 1, \dots, r,$$

U, Σ 和 V 是秩为 r 的矩阵 \bar{t} 的奇异值分解。

为了保证算法收敛性条件 $\tau\sigma\|\mathcal{K}\|^2 < 1$ 成立, 我们需要估计算子 \mathcal{K} 的模。易知 $\|\nabla_3\|^2 < 12$ 和 $\|\mathcal{E}\|^2 < 12$ 。记 $x = (S, w, L)$, 于是

$$\begin{aligned} \|\mathcal{K}x\|^2 &= \|\nabla_3 S - w\|^2 + \|\mathcal{E}w\|^2 + \|AL + AS\|^2 \\ &\leq (\|A\|^2 + 12)\|S\|^2 + 2\sqrt{12}\|S\|\|w\| + 13\|w\|^2 \\ &\quad + \|A\|^2\|L\|^2 + 2\|A\|^2\|S\|\|L\|. \end{aligned}$$

因为 $2\sqrt{12}\|S\|\|w\| \leq a^2\|S\|^2 + 12\|w\|^2/a^2, a > 0$, 并且 $2\|A\|^2\|S\|\|L\| \leq \|A\|^2(\|S\|^2 + \|L\|^2)$, 我们有

$$\begin{aligned} \|\mathcal{K}x\|^2 &\leq (2\|A\|^2 + 12 + a^2)\|S\|^2 + (12/a^2 + 13)\|w\|^2 + 2\|A\|^2\|L\|^2 \\ &\leq \max\{2\|A\|^2 + 12 + a^2, 12/a^2 + 13, 2\|A\|^2\}\|x\|^2 \\ &= \max\{2\|A\|^2 + 12 + a^2, 12/a^2 + 13\}\|x\|^2, \end{aligned}$$

易知, 当 $\max\{2\|A\|^2 + 12 + a^2, 12/a^2 + 13\}$ 取得最小值时, 满足 $2\|A\|^2 + 12 + a^2 = 12/a^2 + 13$, 经过计算可得:

$$\|\mathcal{K}\|^2 \leq \frac{2\|A\|^2 + 25 + \sqrt{(2\|A\|^2 - 1)^2 + 48}}{2}.$$

3.3 数值实验数据与评价方法

3.3.1 数值实验数据

我们回顾性地将提出的模型应用在三组不同的动态 MR 数据集中。第一组数据为生理学改进非均匀心脏躯干数值体模 (physiologically improved non uniform cardiac torso, PINCAT), 图像维度为 $128 \times 128 \times 50$ 。第二组数据为活体心脏灌注 MR 数据 (in vivo cardiac perfusion), 图像维度为 $190 \times 90 \times 70$ 。这两组数据也被应用在模型 kt-SLR 中, 选择它们是为了更公平地和其他模型进行比较。有关数据集的详细信息, 请参考文献^[152,153]。第三组数据是活体胸部 DCE-MRI 数据。这些数据都来自于不同的病人, 并经过了机构审查委员会的批准。这组数据使用了 Philips 公司的 Achieve 3T 扫描仪通过破坏性梯度回波序列 (spoiled gradient recalled echo, SPGRE) 获取, 重复时间 T_R 为 4.33 ms, 回波时间 T_E 为 2.12 ms, 偏转角为 12° 。采集所得的数据为线圈组合复值数据, 维度为 $192 \times 192 \times 10 \times 105$, 其中相位编码和频率编码均为 192, 切片数量为 10, 帧数为 105。空间分辨率为 $1.33 \text{ mm} \times 1.33 \text{ mm} \times 5 \text{ mm}$, 视野大小为 $256 \text{ mm} \times 256 \text{ mm} \times 50 \text{ mm}$ 。更详细的成像细节请参考文献^[154]。我们将这个数据记作 Breast1。另外两个胸部数据是在同样的协议不同的成像参数下采集得到的, 维度为 $192 \times 192 \times 20 \times 25$ 。我们将这两个数据分别记作 Breast2 和 Breast3。注意, 与 Breast1 不同, Breast2 和 Breast3 只有实值的幅值 (magnitude) 数据。为了简化计算, 我们只关注通过肿瘤中心的切片。对于 Breast1, 中心切片为第 6 个; 对于 Breast2 和 Breast3, 中心切片为第 10 个。最终数据的维度分别为 $192 \times 192 \times 105$ 和 $192 \times 192 \times 25$ 。

3.3.2 采样模式与评价方法

我们将提出的模型和最前沿的模型进行比较, 它们分别为 kt-SLR^{1[49]}, kt-RPCA^{2[111]}, L+S^{3[109]} and ICTGV^{4[61]}。以上模型的程序都使用 MATLAB 编写, 并且全部开源。所有数值实验程序均在双核 Xeon E5-2630 2.2 GHz CPU, 128GB 内存和 NVIDIA TITAN RTX GPU 的工作站上运行, MATLAB 的版本为 2018a。为了加速 MATLAB 程序, 我们也使用 CUDA C 语言在 GPU 上实现提出的模型, 并且比较了 MATLAB 和 CUDA 的运行时间。目前 MATLAB 版本和 CUDA 版本的程序均已开源, 详细程序请参考网址<https://github.com/chixindebaoyu/tgvnn>。

¹http://user.engineering.uiowa.edu/~jcb/Software/ktslr_matlab/Software.html

²<http://agsp.org/bt/ktrpca/>

³<http://cai2r.net/resources/software>

⁴<https://github.com/IMTTugraz/AVIONIC>

我们选取了两种不同的采样模式进行数值试验，并应用到所有模型中。图3.1展示了数值实验中使用的采样模式，其中图3.1 (a) 是伪径向采样模式。在每一帧图像上，我们均匀地选取一些径向采样线，并使得帧与帧之间的采样线随机转动一个角度以增加非相干性。伪径向采样是动态 MR 重建数值试验中常用的采样模式，其采样点均在 Cartesian 网格上，直接使用快速 Fourier 变换即可实现频率域到图像域的转化，计算快速简便。在所有数值实验中，如非特别说明，我们均使用 32 条采样线来保证数值实验的一致性。伪径向采样的总加速因子为 6.6，这里的总加速因子为 k-space 中所有点的个数与 k-space 采样点个数的比值。为了测试提出的模型在不同采样模式下的表现，我们也选取了 Cartesian 采样模式进行试验，如图3.1 (b) 所示。Cartesian 采样模式是临床 MR 成像最常用的采样模式。在每一帧图像上，我们对低频区域进行全采样，对高频区域进行随进采样。具体来说，我们在 k-space 中选取中心窗口大小为 20 的区域作为低频区域，而在外围的高频区域，我们随机选择一些相位编码线来增加非相干性，使得总体来说外围的每一条相位编码线被选取的次数相同，并且帧与帧之间达到最大的正交性。Cartesian 采样的总加速因子为 5.1。

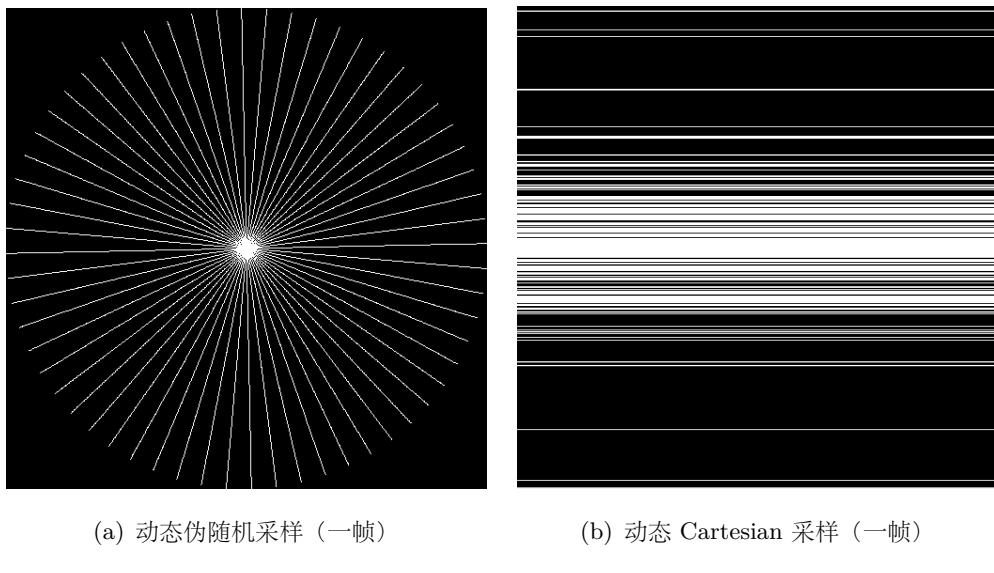


图 3.1 文章使用的动态采样模式（一帧）。

数值实验使用的评价方法为信噪比 (signal error ratio, SER) 和相似性测度^[155] (structural similarity index, SSIM)，它们是压缩感知重建中最常用到的评价方法。SER 的定义为：

$$\text{SER} = -10 \log_{10} \frac{\|X_{\text{rec}} - X_{\text{ori}}\|_{\text{F}}^2}{\|X_{\text{ori}}\|_{\text{F}}^2},$$

其中 X_{rec} 为重建后的图像， X_{ori} 为原始的、全采样图像。SSIM 的定义为

$$\text{SSIM} = \frac{(2\mu_{\text{rec}}\mu_{\text{ori}} + c_1)(2\sigma_{\text{rec},\text{ori}} + c_2)}{(\mu_{\text{rec}}^2 + \mu_{\text{ori}}^2 + c_1)(\sigma_{\text{rec}}^2 + \sigma_{\text{ori}}^2 + c_2)},$$

其中 μ_{ori} 和 μ_{rec} 分别为原始图像和重建图像的均值， σ_{ori} 和 σ_{rec} 分别为原始图像和重建图像的方差， $\sigma_{\text{rec},\text{ori}}$ 为原始图像和重建图像之间的协方差， c_1 和 c_2 为常数。由于我们

处理的是动态 MR 图像，这里的 SSIM 计算为每一帧的平均值。进行数值实验时为了保证一致性，所有数据都被缩放到 [0,1]。我们通过以下方法调试模型参数：对于 PINCAT 数据，通过对每个参数进行穷举搜索得到最高 SER 时的参数，并将此参数用于其他数据集中。表3.1展示了提出模型所使用的参数。为了保证算法收敛，伪径向采样的迭代次数设置为 500，Cartesian 采样的迭代次数设置为 1000。

表 3.1 文章提出模型使用的参数

采样模式 \ 参数	α_0	α_1	β	σ	τ	μ	iter
Pseudo radial	0.006	0.004	0.5	0.25	0.25	1	500
Cartesian	0.015	0.005	0.08	0.28	0.28	1	1500

3.4 数值实验结果

表3.2展示了在伪径向采样模式下，各个模型在不同数据上重建结果的比较。可以看出本文的模型在所有的数据集上都取得了很好地效果，尤其是在 cardiac 和胸部 DCE-MRI 数据上，SER 和 SSIM 均为最高。虽然 kt-SLR 在 PINCAT 数据上的 SER 最高，其在胸部数据上的表现很差。同样的，kt-RPCA 虽然在胸部数据上的表现仅次于本文提出的模型，但其在 PINCAT 和 cardiac 数据上的表现一般。

图3.2展示了各个模型在 PINCAT 数据上第 1 帧的重建结果。重点看红色箭头，虽然 kt-SLR 取得了最高的 SER (33.35)，其重建的图像中仍然残留有空间伪影，而我们提出的方法中则没有。此外，蓝色箭头表明我们的模型在保持边缘和抑制伪影方面的表现要优于 kt-RPCA, L+S 和 ICTGV。

图3.3展示了各个模型在心脏灌注数据上第 1 帧的重建结果。我们的模型取得了最高的 SER (19.58) 和 SSIM (0.9514)，并且重建图像中的伪影最少。红色箭头表明 kt-SLR, kt-RPCA 和 ICTGV 的表现很差，尤其在保持图像边缘上。虽然 L+S 在心脏灌注数据上的表现很好，蓝色箭头表明本文模型的重建图像比 L+S 更加光滑。

为了比较时间方向上的重建效果，我们针对心脏灌注数据展示了各个模型不同帧与时间序列图像的结果，如图3.4所示。可以看到，本文模型的重建图像在视觉上的质量最高。红色箭头显示出 kt-SLR 的重建图像很模糊，这导致了低 SER。蓝色箭头表明 kt-RPCA 的重建图像有严重的运动伪影，尤其在心脏的边缘部分。同时我们也可以从黄色箭头看出，kt-RPCA 和 ICTGV 在时间上有过平滑效应，尤其是 ICTGV。L+S 在 SER 和时间序列图像上的表现很好，但是由于 L+S 没有空间方向的约束，其重建图像的每一帧都有噪声（绿色箭头）。

图3.5, 3.6和3.7分别展示了各个模型在 Breast1 (第 105 帧), Breast2 (第 25 帧) 和 Breast3 (第 1 帧) 上的重建结果。本文的模型不仅取得了最高的 SER 和 SSIM，也可以很好地抑制空间伪影（蓝色箭头）并重建肿瘤区域（红色箭头）。kt-SLR, kt-RPCA

表 3.2 各个模型在不同数据上的重建结果

模型 \ 数据集		PINCAT	cardiac	Breast1	Breast2	Breast3
Zerofilled	SER	20.54	14.80	11.32	11.57	14.85
	SSIM	0.8321	0.8855	0.4820	0.5897	0.7086
kt-SLR	SER	33.35	17.58	17.42	13.61	18.38
	SSIM	0.9880	0.9412	0.7712	0.6956	0.8461
kt-RPCA	SER	29.27	18.33	19.31	14.82	19.81
	SSIM	0.9700	0.9447	0.8857	0.7920	0.9068
L+S	SER	27.99	19.12	17.60	14.74	19.91
	SSIM	0.9514	0.9490	0.7675	0.7319	0.8791
ICTGV	SER	26.88	17.87	16.31	12.18	16.84
	SSIM	0.9435	0.9405	0.6735	0.6080	0.7693
Proposed	SER	32.74	19.57	20.56	16.24	21.08
	SSIM	0.9917	0.9514	0.9402	0.9091	0.9356

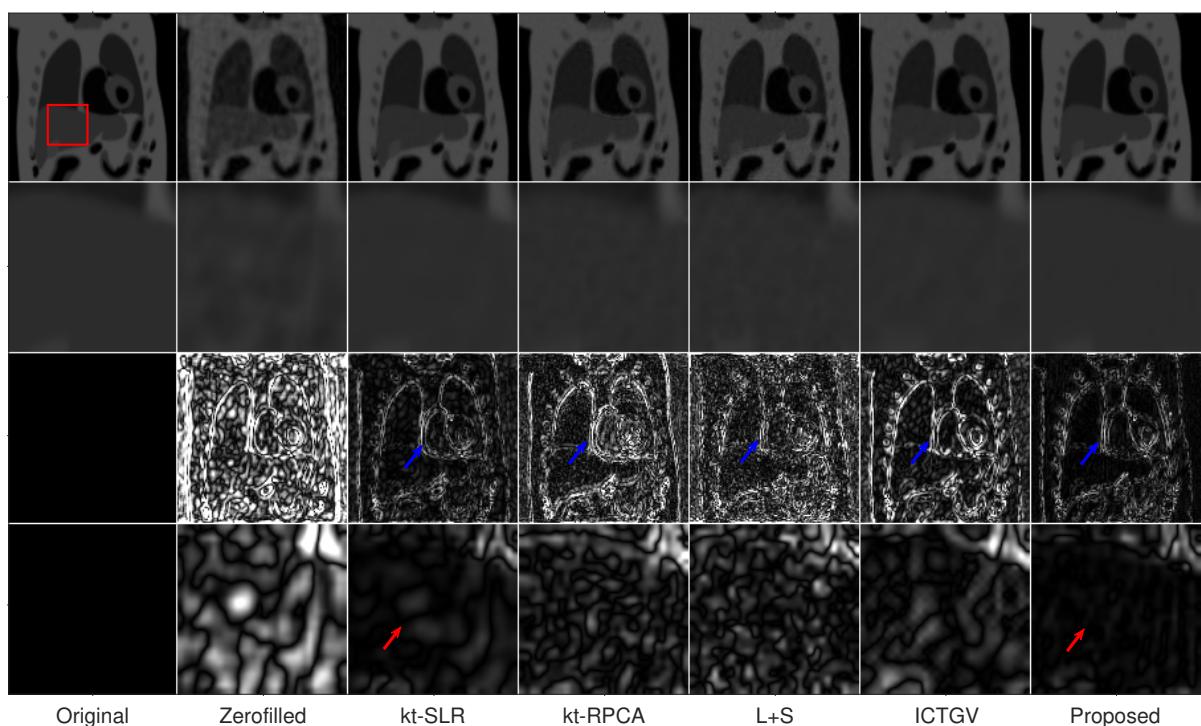


图 3.2 各个模型在 PINCAT 数据上的重建结果（第 1 帧）。第 1 行：重建图像；第 2 行：红色方框区域的放大图像；第 3 行：相对于原始图像的差值图像；第 4 行：红色方框区域放大的差值图像。为了增加可见度，差值图像放大了 30 倍。

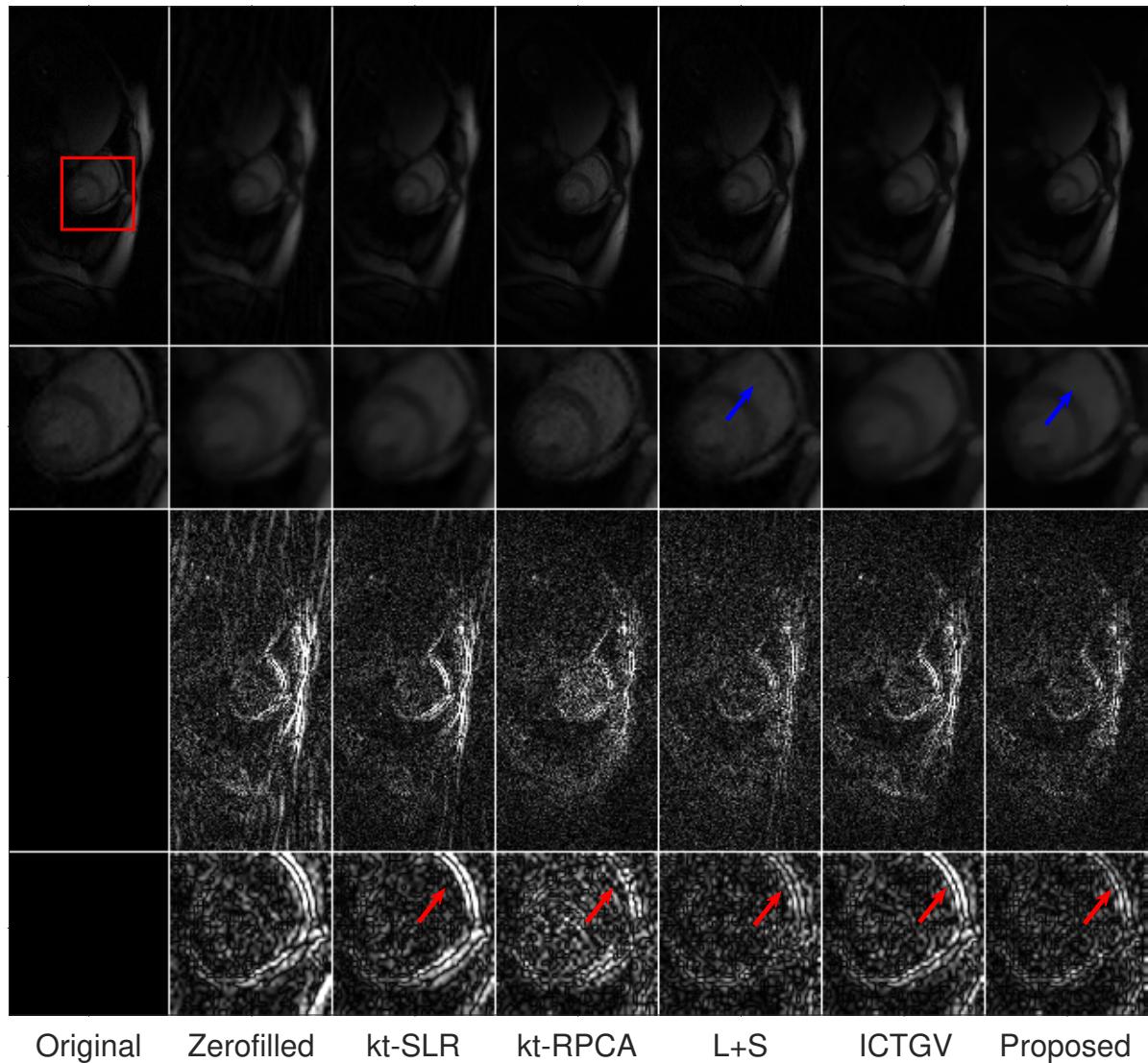


图 3.3 各个模型在 cardiac 数据上的重建结果（第 1 帧）。第 1 行：重建图像；第 2 行：红色方框区域的放大图像；第 3 行：相对于原始图像的差值图像；第 4 行：红色方框区域放大的差值图像。为了增加可见度，差值图像放大了 20 倍。

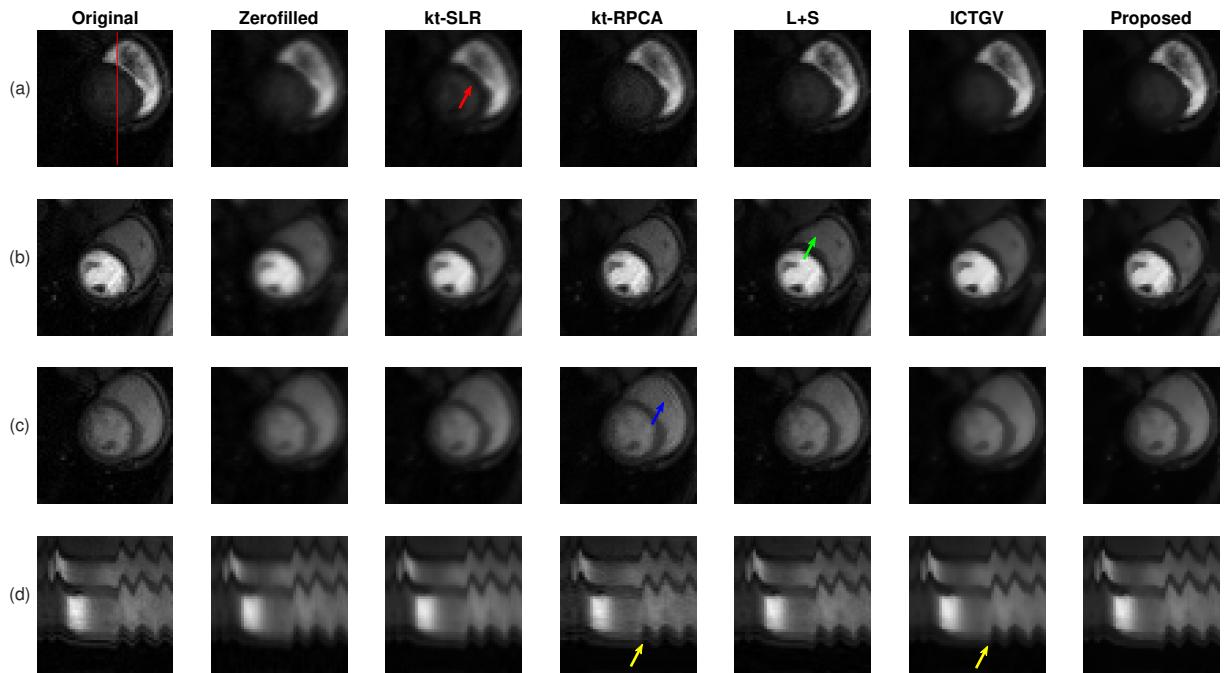


图 3.4 各个模型在 cardiac 数据上的重建结果（时间方向）。（a）、（b）和（c）行分别为重建图像的第 11，第 21 和第 54 帧；（d）行是通过心脏的一列（红线）在时间序列图像。

和 ICTGV 并不适用于胸部 DCE-MRI 数据集。有趣的是，L+S 在重建肿瘤方面的表现，在视觉上略优于我们的模型，但是 L+S 重建图像的 SER 和 SSIM 均很低，并且背景区域有残留的空间伪影（蓝色箭头）。

我们也比较了各个模型在 Breast1 数据上的运行时间，如表3.3所示。值得注意的是 ICTGV 也有 GPU 版本的代码，但在安装的时候出现了问题，导致程序无法在我们的工作站上运行，因此这里只展示了 ICTGV 在 CPU 上的运行时间。可以看出，kt-SLR 和 kt-RPCA 的运行速度相对较快，这是因为它们只包含时间方向上的稀疏变换。虽然本文模型的 MATLAB 代码运行速度较慢（2812.90 s），但 GPU 版本的运行时间只需要 28.35 s，GPU 的使用使得程序运行速度提高了将近 100 倍。

为了展示胸部肿瘤随时间的变化，我们也绘制了 Breast1 肿瘤区域的平均像素值时间曲线，如图3.8所示。平均像素值时间曲线是指肿瘤区域（红色方框）的平均像素值随时间变化的曲线，是胸部图像肿瘤区域重建效果的常用评价方法之一。可以看到，我们的模型在造影剂注入之前和之后都可以很好地拟合曲线，并且方差较低。相反的，

表 3.3 各个模型在 Breast1 数据上运行时间比较

模型	kt-SLR (CPU)	kt-RPCA (CPU)	L+S (CPU)	ICTGV (CPU)	Proposed (CPU)	Proposed (GPU)
时间 (s)	5794.88	641.71	493.80	2572.04	2812.90	28.35

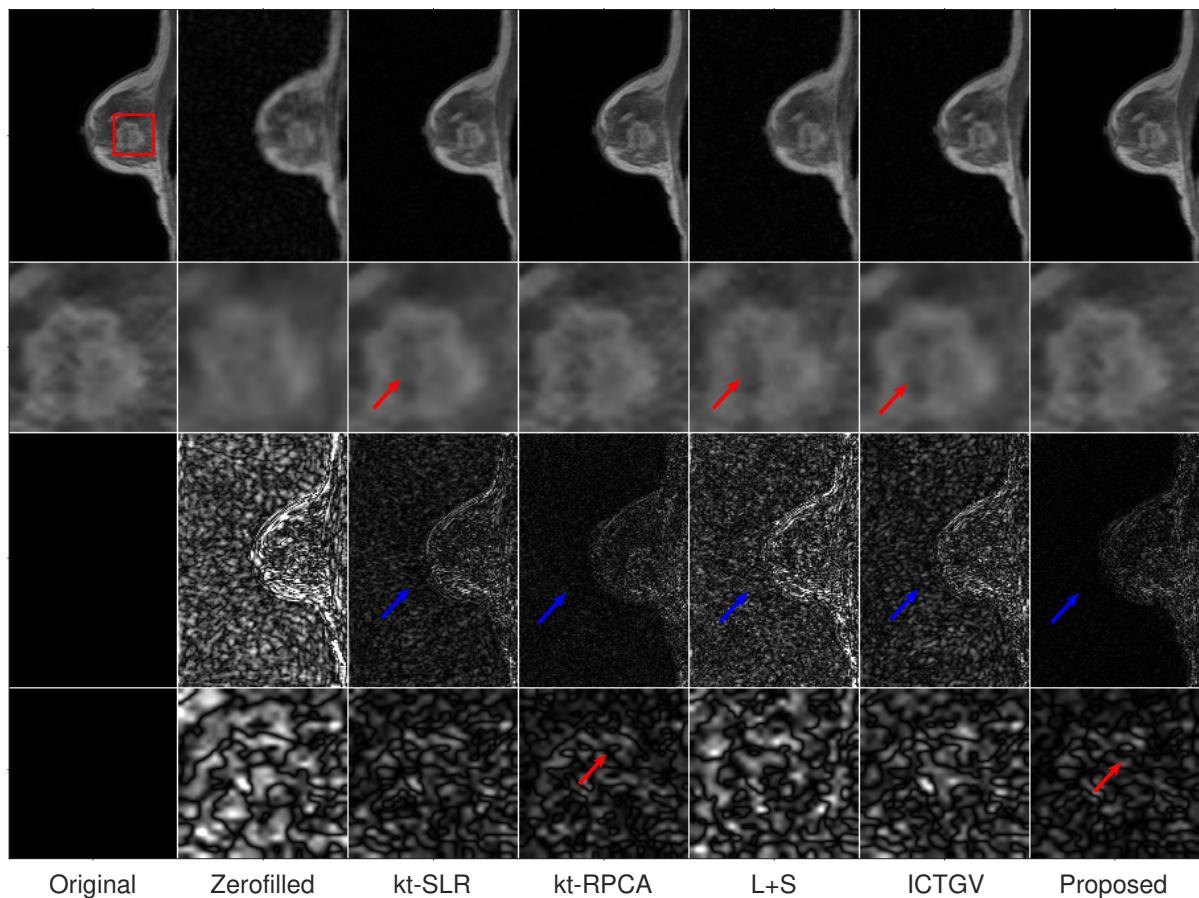


图 3.5 各个模型在 Breast1 上的重建结果 (第 105 帧)。为了在视觉上更好地展示胸部肿瘤, 我们切除了心脏部分。第 1 行: 重建图像; 第 2 行: 红色方框区域的放大图像; 第 3 行: 相对于原始图像的差值图像; 第 4 行: 红色方框区域放大的差值图像。为了增加可见度, 差值图像放大了 10 倍。

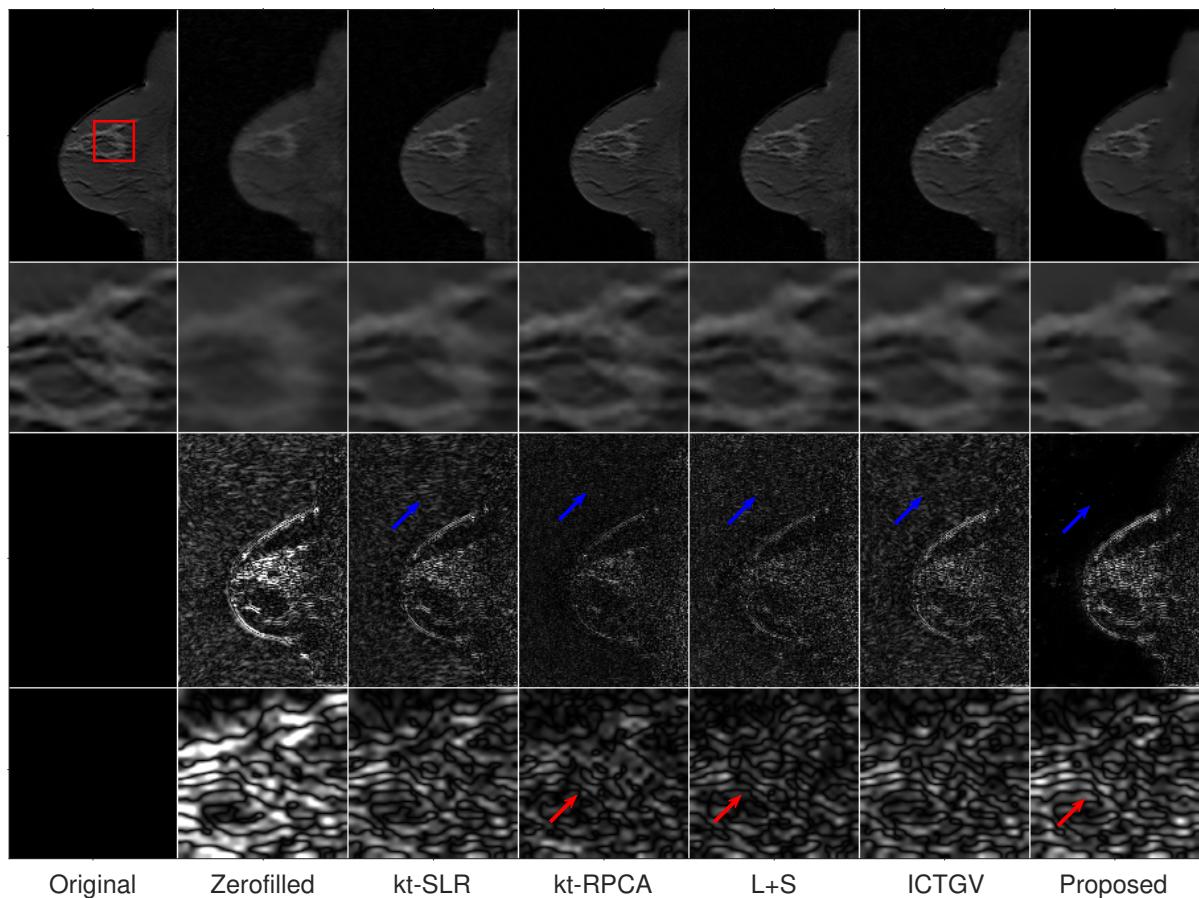


图 3.6 各个模型在 Breast2 上的重建结果（第 25 帧）。为了在视觉上更好地展示胸部肿瘤，我们切除了心脏部分。第 1 行：重建图像；第 2 行：红色方框区域的放大图像；第 3 行：相对于原始图像的差值图像；第 4 行：红色方框区域放大的差值图像。为了增加可见度，差值图像放大了 10 倍。

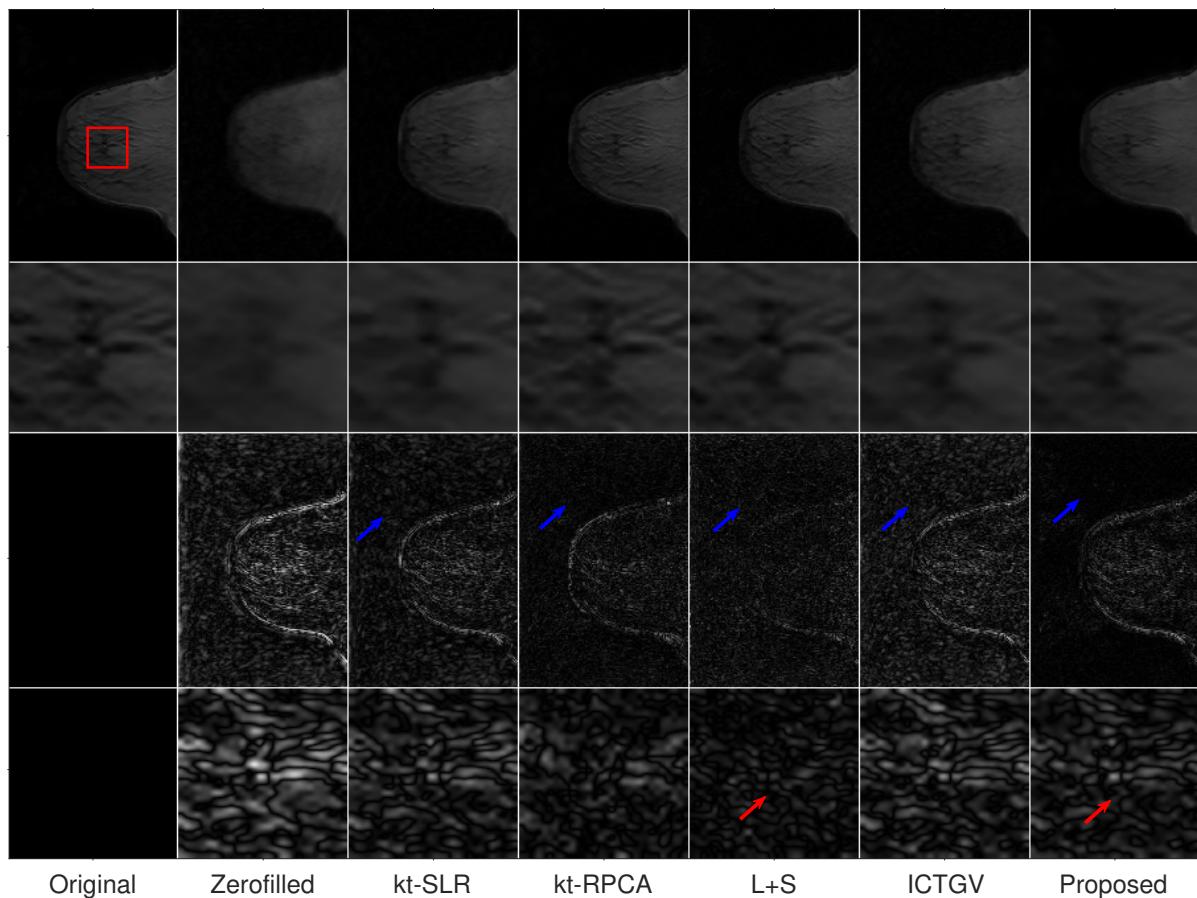


图 3.7 各个模型在 Breast3 上的重建结果（第 1 帧）。为了在视觉上更好地展示胸部肿瘤，我们切除了心脏部分。第 1 行：重建图像；第 2 行：红色方框区域的放大图像；第 3 行：相对于原始图像的差值图像；第 4 行：红色方框区域放大的差值图像。为了增加可见度，差值图像放大了 10 倍。

kt-SLR 和 kt-RPCA 拟合曲线的方差很高，尤其在注入造影剂之后（蓝色箭头）。黄色箭头表明 kt-RPCA 不能很好地拟合前 10–15 帧。L+S 最后 5 帧的拟合效果很差，这一点从绿色箭头中可以更容易看出。

为了测试各个模型在不同加速因子下的表现，我们选择了 5 个不同数量的采样线，即 12、22、32、42 和 52，相应的加速因子分别为 17.0、9.5、6.6、5.1 和 4.2。表3.4展示了不同模型在不同加速因子下在 Breast1 数据上的表现。可以看出，本文的模型在不同的加速因子下都可以得到最高的 SER 和 SSIM，这表明了本文模型的一致性。

为了评价我们的模型在不同采样模式下的表现，我们也比较了各个模型在 Cartesian 采样模式下（如图3.1 (b)）的重建结果。表3.5展示了各个模型在 Cartesian 采样模式下在 Breast1 数据上的重建结果。可以看出，我们的模型得到了最高的 SER (22.96) 和最高的 SSIM (0.9632)，而其他的模型在 Cartesian 采样模式下的表现不理想。

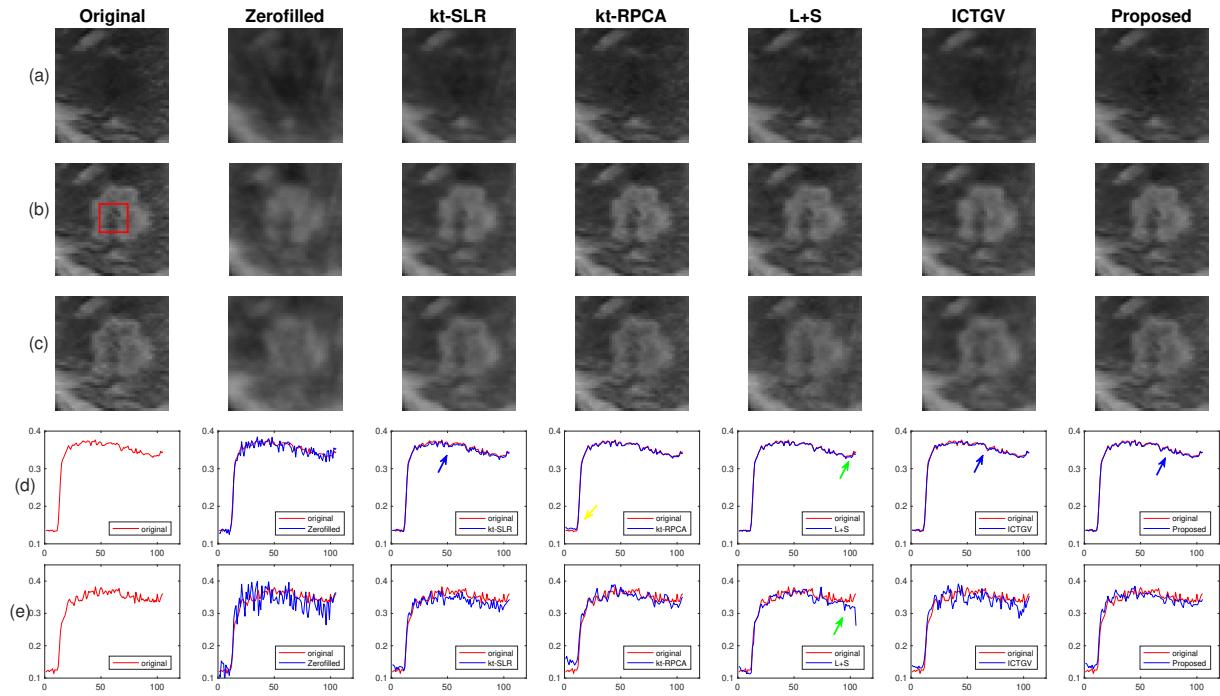


图 3.8 各个模型在 Breast1 上的重建结果（时间方向）。(a)、(b) 和 (c) 行分别为重建图像的第 1, 第 53 和第 105 帧; (d) 行是肿瘤区域（红色方框）的平均像素值时间曲线; (e) 行是肿瘤区域某一像素点的像素值时间曲线。

表 3.4 各个模型在不同加速因子下在 Breast1 数据上的重建结果

模型 \ 采样线		12	22	32	42	52
Zero-filled	SER	5.98	9.06	11.32	12.93	14.33
	SSIM	0.3218	0.4108	0.4820	0.5350	0.5818
kt-SLR	SER	15.16	16.10	17.42	19.07	19.11
	SSIM	0.7323	0.7312	0.7712	0.8087	0.7989
kt-RPCA	SER	13.10	17.74	19.31	20.26	21.13
	SSIM	0.6241	0.8356	0.8857	0.9061	0.9245
L+S	SER	12.75	16.00	17.60	18.73	19.64
	SSIM	0.5885	0.7119	0.7675	0.8018	0.8275
ICTGV	SER	13.56	15.30	16.31	17.03	17.56
	SSIM	0.6001	0.6485	0.6735	0.6895	0.7007
Proposed	SER	16.49	19.06	20.56	21.82	22.96
	SSIM	0.8620	0.9119	0.9402	0.9535	0.9632

表 3.5 各个模型在 Cartesian 采样模式下在 Breast1 数据上的重建结果

模型	Zerofilled	kt-SLR	kt-RPCA	L+S	ICTGV	Proposed
SER	11.29	16.60	16.86	14.55	14.33	17.79
SSIM	0.7370	0.8942	0.9126	0.8379	0.7681	0.9169

3.5 本章小结

在本章中，我们回顾了动态 MR 压缩感知的经典模型，并基于图像分解思想提出了利用二阶时空 TGV 泛函和核范数的重建模型。其中，核范数用来建模时间高度相关的背景部分，可以很好地消除空间伪影；二阶时空 TGV 泛函可以刻画图像中的光滑部分，在保持图像边缘的同时可以抑制阶梯效应的产生。我们使用 Primal-Dual 算法来求解模型，并给出了算法收敛时的范数估计。我们也利用了 GPU 来加速 MATLAB 程序。在 PINCAT、心脏灌注和胸部 DCE-MRI 数据上的数值实验表明，在不同的采样模式和加速因子下，本文的模型都比最前沿的模型在消除空间伪影、保持边缘上的表现更好，尤其是在胸部 DCE-MRI 数据上。

数值实验中，我们在 PINCAT 数据上手动调试了参数，然后将这组参数应用到其他数据中。这样的调参方法需要满足一个假设，即这组参数可以推广到其他数据中。但是在实际中，并不是所有的模型都满足这样的假设。这些模型的参数可能仅仅取决于噪声水平、加速因子等。因此本章中所展示的结果有一定的局限性，对于某个模型和某组数据而言，不同的参数可能会有更好的结果。

4 胸部 DCE-MRI 压缩感知重建模型中时间稀疏正则项的量化评估

磁共振动态对比增强涉及注入顺磁对比剂期间和之后连续采集 T_1 加权 MR 图像^[2]。对比剂通过改变组织固有的驰豫率来增加不同组织之间的对比度。在序列图像中，每个图像体素都会产生一个强度时间曲线，可以用于估计生理参数，例如体积转移常数 K^{trans} 和血管外细胞体积分数 v_e 等。高时间分辨率和高空间分辨率都有利于 DCE-MRI，高时间分辨率有利于精确地定量分析，而高空间分辨率有助于医生的临床阅读。但是，对重复高信噪比图像的要求限制了传统数据采集方法同时提高时间和空间分辨率。

为了加速动态 MR 成像，一种常用的方法来平衡空间和时间分辨率之间的权衡是在每一帧图像上进行下采样。许多成功的算法都用到了这个想法，如 keyhole^[145], k-t FOCUSS^[45], k-t BLAST 和 k-t SENSE^[105]。但是，这些方法所重建的图像通常信噪比低，而且在高下采样率时有严重的伪影。

压缩感知是一种加速 MR 成像的新方法^[9,146]。压缩感知可以用比传统加速方法少的 Fourier 数据精确地重建出 MR 图像，从而减轻数据获取的负担，加速成像。根据压缩感知的理论，MR 图像的重建可以建模为一个约束优化问题。这个优化模型一般分为两项，一项是数据项，另一项是稀疏项。数据项保证了重建后的图像与采集到的下采样数据的一致性，而稀疏项则假设图像在某个变换域下存在稀疏表示。对于动态图像而言，由于图像中的大部分在时间上是保持不变的，因此最显著的冗余性（即稀疏性）经常体现在时间方向上。

压缩感知在癌症 MR 成像中有着很大的应用潜力^[156]，因为很多癌症 MR 成像的协议都是动态的。目前，压缩感知在动态 MRI 中已经有很多成功的应用，如 k-t SPARSE^[13]、k-t SLR^[49]、iGRASP^[44] 等。比如，Han 等展示了基于压缩感知加速的 Cartesian 采样 FLASH (fast low angle shot)^[150] 序列在动物模型的 DCE-MRI 时空分辨率的增强。Chen^[157] 等比较了压缩感知重建中 4 种不同的时间方向 TV 正则项的效果，并且将其应用在胸部 DCE-MRI 中。Ji 和 Lang^[158] 使用时间方向的差分算子来增强空间信号的分辨率。Smith^[50,159] 等展示了空间方向的 TV 稀疏项在大量随机生成的采样方案中量化参数的预期方差。虽然压缩感知已经应用于胸部 DCE-MRI 中，但是目前还没有研究通过量化分析的方式来比较不同时间方向的稀疏项在 DCE-MRI 中的表现。因此，对于定量胸部 DCE-MRI，不同时间方向的稀疏项对重建误差的影响是未知的。

在压缩感知应用于临床肿瘤成像之前，其可靠性和准确性的研究是必要的。本章的主要目的是通过量化分析的方式来比较和评估 5 中不同的时间方向的稀疏项，它们分别是：

- (1) Fourier 变换的 l_1 范数 (Fourier transform, FT);

- (2) Haar 小波变换的 l_1 范数 (Haar wavelet transform, WT);
- (3) TV;
- (4) 二阶 TGV (TGV_α^2);
- (5) 核范数 (NN)。

我们假设这些稀疏项在基于压缩感知重建中的表现是不同的，即重建图像的信噪比和组织参数的准确性会有差异。

4.1 胸部 DCE-MRI 的重建模型

我们将重建的动态 MR 图像记为 $X \in \mathbb{C}^{N_1 \times N_2 \times d}$, 其中空间方向的维度为 $N_1 \times N_2$, 而时间方向的维度为 d 。则带噪声的动态 MR 采样相当于:

$$B = AX + \epsilon,$$

其中 $A = \mathcal{M} \cdot \mathcal{F}$ 是采样矩阵, \mathcal{F} 是对于每一帧图像的二维 Fourier 变换, \mathcal{M} 是每一帧图像的下采样矩阵, ϵ 为高斯白噪声。于是动态 MR 的压缩感知即是从下采样数据 B 中恢复出 X , 其重建模型为:

$$\hat{X} = \arg \min_X \frac{1}{2} \|AX - B\|_F^2 + \alpha \|\mathcal{S}_t X\|, \quad (4.1)$$

其中 $\alpha > 0$ 是平衡数据项和稀疏项的参数, F 是 Forbenius 范数, 而 \mathcal{S}_t 是某个时间方向的稀疏项。极小化稀疏项可以保证重建图像的稀疏性, 而极小化数据项可以保证数据的一致性。

4.1.1 基于 Fourier 变换的重建模型

动态 MRI 经常在时间方向上表现出冗余性, 有时这种冗余表现为周期性, 因此时间方向上的一维 Fourier 变换可以对信号进行稀疏表示。信号越是表现出周期性, 其在 Fourier 变换域中就越稀疏。并且由于下采样的去噪效应, 一个非周期性的带噪信号在 Fourier 域也是可压缩的。比如, 在心脏电影成像中, 由于心脏搏动的周期性, 图像可以很好地用时间方向的 Fourier 变换进行稀疏表示。在 k-t SPARSE^[160] 中, 作者使用了空间上的小波变换和时间上的 Fourier 变换作为压缩感知重建的稀疏项。记 \mathcal{F}_t 为一维时间上的 Fourier 变换, 我们可以得到模型:

$$\hat{X} = \arg \min_X \frac{1}{2} \|AX - B\|_F^2 + \alpha \|\mathcal{F}_t X\|_1. \quad (4.2)$$

4.1.2 基于 Haar 小波变换的重建模型

小波变换是基于压缩感知重建中常用的稀疏表示。对于 MR 图像而言, Haar 小波有两个感兴趣特征。第一, 分片常数的信号在 Haar 小波的表示下是稀疏的。第二, Haar 小波变换和 MR 图像的采集域, 即 Fourier 域有着高非相干性。根据压缩感知理

论, 这意味着对于给定的下采样率, Haar 小波变换可以获得更好的重建图像。比如, k-t SPARSE^[160] 将空间小波变换成功地应用在了心脏电影成像中。记 \mathcal{W}_t 为一维时间方向的 Haar 小波变换, 我们可以得到模型:

$$\hat{X} = \arg \min_X \frac{1}{2} \|AX - B\|_F^2 + \alpha \|\mathcal{W}_t X\|_1. \quad (4.3)$$

4.1.3 基于 TV 的重建模型

在压缩感知应用在 MR 成像的第一篇文章中, Lustig^[151] 在模型中使用了有限差分作为图像的稀疏表示。离散 TV 算子将时间和空间上的差分算子进行加权组合, 本质上是将图像用梯度进行稀疏表示。TV 泛函最初被应用于图像去噪中^[161], 最近也经常应用于压缩感知重建中。离散 TV 算子和 Fourier 域有着最大的非相干性, 因此, 对于给定的下采样率, 会比其他稀疏表示得到更好的重建图像。但实际上, 使用 TV 正则项经常会导致重建后的图像出现阶梯效应。这是因为真实的 MR 图像中并不是完美的分片常数, 而是存在高阶信息。当这些高阶信息用分片常数函数来刻画时, 在重建的过程中, 为了使得图像在梯度域中稀疏, 最小的梯度系数会被减小到零, 因此会损失很多高阶信息, 从而出现阶梯效应。在 k-t SLR^[49] 中, 作者使用了三维 TV 变换:

$$\|\cdot\|_{3DTV} = \|\nabla_3 \cdot\|_1,$$

而

$$|\nabla_3 \cdot| = \sqrt{(\nabla_x \cdot)^2 + (\nabla_y \cdot)^2 + (\nabla_t \cdot)^2},$$

其中 ∇_x, ∇_y 和 ∇_t 分别为 x, y 和 t 方向上的梯度算子。也有文献只使用了时间方向上的梯度算子就取得了很好的效果, 如 DLTG^[119] 和 iGRASP^[44]。在本文中, 我们仅考虑时间方向的梯度算子并评估以下模型的表现:

$$\hat{X} = \arg \min_X \frac{1}{2} \|AX - B\|_F^2 + \alpha \|\nabla_t X\|_1. \quad (4.4)$$

4.1.4 基于二阶 TGV 的重建模型

如上所言, 有限差分算子可以很好地对 MR 图像进行稀疏表示, 但是也会带来阶梯效应。因此, Bredies^[60] 等提出了基于高阶导数信息的 TGV 泛函, 可以更好地对图像中的高阶信息进行建模, 消除阶梯效应。对于二维图像 $u \in \mathbb{C}^{N_1 \times N_2}$, 其离散二阶 TGV 的定义如下:

$$\text{TGV}_{\alpha}^2(u) = \min_{\omega \in \text{BD}(\Omega)} \alpha_1 \|\nabla_3 u - \omega\|_1 + \alpha_0 \|\mathcal{E}(\omega)\|_1,$$

其中 $\nabla \cdot = (\nabla_x \cdot, \nabla_y \cdot)$ 一阶有限差分算子, $\mathcal{E}(\omega) = \frac{1}{2}(\nabla_3 \omega + \nabla_3 \omega^T)$ 是对称梯度算子, $\text{BD}(\Omega)$ 是有界形变函数空间。

从定义中可以看出, TGV_{α}^2 是 TV 的高阶推广, 可以很好地保持图像一阶和二阶导数之间的平衡。因此 TGV_{α}^2 比 TV 更适合建模图像中的光滑部分。在压缩感知模型中,

使用 TGV_{α}^2 可以在保持边缘的同时减少阶梯效应。在本文中，我们只考虑时间方向的 TGV_{α}^2 ：

$$\hat{X} = \arg \min_X \frac{1}{2} \|AX - B\|_{\text{F}}^2 + \text{TGV}_{\alpha}^2(X), \quad (4.5)$$

这里 $\nabla \omega = \nabla_t \omega$, $\omega \in \mathbb{C}^{N_1 N_2 d}$.

4.1.5 基于核范数的重建模型

低秩矩阵补全的想法也被应用在动态 MR 重建中。这时我们将动态 MR 图像视为一个时空矩阵，每一帧图像则为时空矩阵中的一列，每一个像素点为时空矩阵的一行。由于动态图像时间和空间上的相关性，这样的矩阵可以用低秩模型来刻画。因此，将压缩感知和低秩矩阵补全结合在一起可以进一步加快成像速度。在 k-t SLR 中，作者在模型中使用核范数，并在体膜数据和心脏灌注数据中取得了显著的提高。核范数记为 $\|X\|_*$ ，是矩阵 X 所有奇异值的和。本文我们考虑如下模型：

$$\hat{X} = \arg \min_X \frac{1}{2} \|AX - B\|_{\text{F}}^2 + \alpha \|X\|_*. \quad (4.6)$$

4.2 图像重建快速算法

4.2.1 FISTA 快速算法

在本文中，所有模型都使用 FISTA^[31] 算法进行求解。FISTA 是一种算子分裂算法，是压缩感知重建中常用的快速算法，可以有效地求解如下形式的模型：

$$\hat{x} = \arg \min_{x \in \mathbb{C}^n} f(x) + g(x), \quad (4.7)$$

其中， f 是光滑的凸函数，其 Lipschitz 常数为 L_f ； g 是凸函数，可能是非光滑的。FISTA 的迭代过程算法6所示。这里 $\text{prox}_{\tau}(g)(u)$ 是函数 $g(x)$ 在点 x 的临近算子，定义为：

$$\text{prox}_{\tau}(g)(u) = \arg \min_x g(u) + \frac{1}{2\tau} \|u - x\|_2^2. \quad (4.8)$$

而投影算子 \mathcal{P} 定义为：

$$x = \mathcal{P}(x, [l_1, l_2]) = \begin{cases} x, & \text{if } l_1 \leq x \leq l_2; \\ l, & \text{if } x < l_1; \\ u, & \text{if } x > l_2. \end{cases} \quad (4.9)$$

FISTA 中最重要的一步是找到高效的算法求解临近点子问题。由于在压缩感知模型中，我们通常使用 l_1 范数作为稀疏项，临近点问题可以由软阈值收缩求解。比如，令 $g(u) = \|u\|_1$ ，则方程 (4.8) 变为：

$$\text{prox}_{\tau}(g)(u) = \arg \min_u \|u\|_1 + \frac{1}{2\tau} \|u - x\|_2^2. \quad (4.10)$$

算法 6 稀疏重建的 FISTA 算法

初始化: $\tau, l_1, l_2, t^1 = 1, x^0 = r^1$;

迭代: 根据以下步骤更新参数:

1. $\bar{x} = r^n - \tau A^*(Ax - b)$;
2. $x^n = \text{prox}_\tau(g)(\bar{x})$;
3. $x^n = \mathcal{P}(x^n, [l_1, l_2])$
4. $t^{n+1} = (1 + \sqrt{1 + 4(t^n)^2})/2$;
5. $r^{n+1} = x^n + \frac{t^n - 1}{t^{n+1}}(x^n - x^{n-1})$;

直到收敛, 返回 x^{n+1} 。

方程 (4.10) 的解为

$$u = \mathcal{S}_\tau(x), \quad (4.11)$$

其中软阈值收缩算子 $\mathcal{S}_\tau : \mathbb{R}^n \rightarrow \mathbb{R}^n$ 的定义为:

$$\mathcal{S}_\tau(x) = \max(0, x - \tau). \quad (4.12)$$

当 $g(u)$ 为 TV 或者 TGV_α^2 时, 临近点问题不能用简单地软阈值收缩进行求解, 因此我们用 Primal-Dual 算法进行求解。Primal-Dual 算法的详细流程请参考第3章。

4.2.2 模型参数的选择

压缩感知模型中, 参数的选择是至关重要的, 因为只有在每个模型都达到最优解时的比较才是有意义的。在本文的数值实验中, 所有胸部 DCE-MRI 数据都被缩放到 $[0,1]$, 这样有助于简化调参。表4.1显示了以上所有模型所用的参数。其中 α 是稀疏项的权重, iter 是算法的迭代次数, σ , τ 和 λ 分别是 TV 或 TGV_α^2 模型中, 一阶迭代步长、二阶迭代步长和数据项的权重。参数的调节方法如下: 我们从很低的权重开始 ($\alpha = 10^{-5}$) 并且逐步增加权重的大小直到空间伪影在视觉上从重建的图像中消失。在每一步迭代中, 我们跟踪连续两次输出图像的相对差值, 直到其小于预设的阈值 (10^{-3})。

表 4.1 各个模型所使用的参数

稀疏项	α	iter	σ	τ	λ
FT	0.059	35	-	-	-
WT	0.008	60	-	-	-
TV	0.5	100	0.2	0.2	0.5
TGV_α^2	0.5	100	0.2	0.2	0.5
NN	0.3	40	-	-	-

4.3 数值实验数据与评价方法

4.3.1 数值实验数据

我们回顾性地将所有模型应用于活体胸部 DCE-MRI 数据上，这些数据经过了机构审查委员会的批准。数据使用 Philips 公司的 Achieve 3T 扫描仪通过破坏性梯度回波序列 (Spoiled Gradient Recalled Echo, SPGRE) 获取，重复时间 T_R 为 4.33 ms，回波时间 T_E 为 2.12 ms，偏转角为 12°。数据的维度为 $192 \times 192 \times 10 \times 105$ ，其中相位编码和频率编码均为 192，切片数量为 10，动态数量为 105。空间分辨率为 $1.33 \text{ mm} \times 1.33 \text{ mm} \times 5 \text{ mm}$ ，视野大小为 $256 \text{ mm} \times 256 \text{ mm} \times 50 \text{ mm}$ 。更详细的成像细节请参考文献^[162]。

数据通过 Cartesian 全采样模式获得，然后通过一些列的随机采样模式进行回顾性下采样。虽然前瞻性下采样更加符合现实，但是前瞻性地获取上百种不同的下采样模式是不切实际的。为了减少计算，我们只选取了肿瘤中心所在的切片。在我们的数据中，这是第 6 个切片。同时在所有的重建中，我们切除了胸腔的后边部分来增加图像的稀疏性，这是因为跳动的心脏会在相位编码方向上产生伪影，从而降低重建图像的信噪比。切掉胸腔的过程可以在下采样数据中进行，因为即使在充满伪影的图像中，胸腔也是很容易被划分出来的。最终图像的维度为 $192 \times 128 \times 105$ 。

我们生成了 200 个不同的 Cartesian 下采样模式，每个采样模式的维度都和数据相同，即 $192 \times 128 \times 105$ 。在每一帧图像上，我们对低频区域进行全采样，对高频区域进行随机采样。具体来说，我们在 k-space 中选取中心窗口大小为 20 的区域作为低频区域，而在外围的高频区域，我们随机选择一些相位编码线，使得总体来说外围的每一条相位编码被选取的次数相同。这 200 个下采样模式的平均下采样率为 4.5。这里我们选择 4.5 是因为 Cartesian 采样的固有限制，更高的采样率可以使用 non-Cartesian 采样模式获得，但这不在本文考虑范围内。图4.1 显示了一个 Cartesian 采样模式的例子。

4.3.2 采样模式与评价方法

我们使用两种方法对重建结果进行评估。首先，为了和之前的文献一致，我们使用基于图像像素值本身的信噪比 (SER) 对图像进行评估。SER 的定义如下：

$$\text{SER} = -10 \log_{10} \frac{\|X_{\text{rec}} - X_{\text{ori}}\|_F^2}{\|X_{\text{ori}}\|_F^2}, \quad (4.13)$$

其中 X_{ori} 为全采样图像，即原图像， X_{rec} 为算法重建后的图像。

我们使用的第二种评估方法是通过 DCE-MRI 的药代动力学 (pharmacokinetic) 模型来进行定量分析。本文我们使用最常用的标准 Tofts-Kety 模型^[163] 来计算 K^{trans} 和 v_e 的值。首先我们分别计算全采样和重建图像 K^{trans} 和 v_e ，然后用一致性相关系数 (concordance correlation coefficients, CCC) 来度量他们之间的相似性程度。这是 DCE-

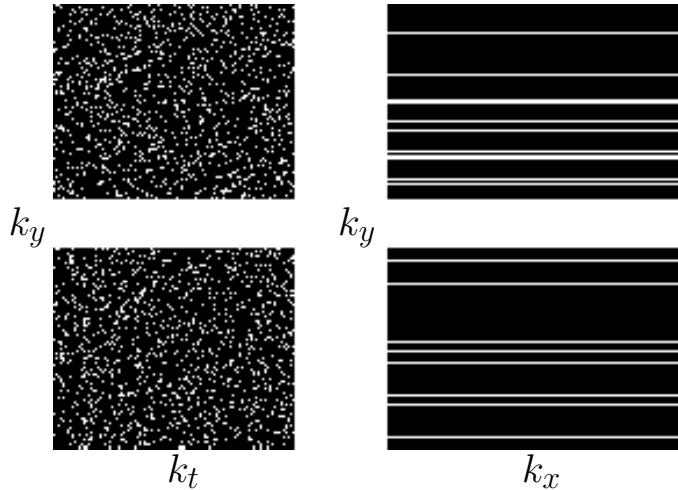


图 4.1 Cartesian 采样模式的例子。左图显示的是所有帧的采样模式，右图显示的是某一帧的采样模式。右图的二维采样模式通过左图的一列在 x 方向上重复得到。在每一帧图像上，中心的低维区域进行全采样，而外围的高维区域进行随机采样。整体的下采样在 4.5 左右。

MRI 中进行定量参数分析的常用方法。在统计学上，CCC 的定义如下：

$$\text{CCC} = \frac{2\sigma_{\text{ori}}\sigma_{\text{rec}}}{\sigma_{\text{ori}}^2 + \sigma_{\text{rec}}^2 + (\mu_{\text{ori}} - \mu_{\text{rec}})^2}, \quad (4.14)$$

其中 μ_{ori} 和 μ_{rec} 分别为全采样图像和压缩感知重建图像 K^{trans} 或者 v_e 的期望， σ_{ori} 和 σ_{rec} 是相对应的方差。

为了从视觉上直观地评估重建图像时间剖面的准确性，我们计算了重建图像肿瘤区域的时间曲线相对于全采样图像之间的差异。我们首先手动将肿瘤区域分割出来，形成一个掩膜，并应用于所有重建图像。我们也分别计算了全采样图像和重建图像肿瘤区域内 K^{trans} 和 v_e 的平均值，并用 Bland-Altman 图来显示其一致性。

重建程序全部使用 MATLAB 语言编写，DCE-MRI 的参数分析使用 Julia 语言的 DCEMRI.jl^[164] 工具包运行。所有数值实验均在双核 Xeon E5-2665 2.40 GHz, 20 GB 内存的工作站上运行，MATLAB 版本为 2015b (Mathworks, Natick, MA)，Julia 版本为 0.4.3。程序已经开源，详细程序请参考网址 <https://github.com/chixindebaoyu/qetsr>。

4.4 数值实验结果

4.4.1 重建图像质量

我们首先评估了这 5 种时间方向的稀疏项在图像质量上的表现。表 4.2 的第 2 列显示了 200 个不同 Cartesian 采样模式下重建图像的平均 SER。可以看出，在 5 种稀疏项中，核范数得到了最高的 SER (29.1)，小波变换的 SER 最低 (21.8)，而 Fourier 变换的效果介于小波变换和 TV (27.7) 之间。其中， TGV_α^2 (27.8) 与 TV 有着相似的表现。

图 4.2 (a) 显示了第 1 个 Cartesian 采样模式下 5 种时间稀疏项重建结果的第 105

表 4.2 200 个 Cartesian 采样模式重建的均值 SER 与 CCC

稀疏项	SER	CCC	
		K^{trans}	v_e
Zerofilled	15.1	0.694	0.636
FT	26.4	0.763	0.575
WT	21.8	0.878	0.733
TV	27.7	0.974	0.916
TGV	27.8	0.974	0.917
NN	29.1	0.842	0.799

帧图像。图中红色箭头显示，背景伪影在 TGV_α^2 和 TV 中依然存在，而 NN 基本消除了伪影。虽然 FT 和 WT 也很好地抑制了伪影，但是它们在肿瘤部分重建的表现是最差的。

图4.2 (b) 显示了同样 Cartesian 采样模式下重建结果的第 1 帧图像。图中红色箭头显示，在视觉上 TGV_α^2 和 TV 重建肿瘤的效果是最好的，这也是我们之后进行定量参数分析所用到的部分。FT 和 WT 在重建肿瘤方面的表现是最差的，而 NN 虽然 SER 最高，但从视觉上看，肿瘤部分的误差却高于 TGV_α^2 和 TV。

图4.3显示了重建图像 SER 的箱型图。我们可以观察到，对于每个稀疏项而言，SER 的方差都相对很小。因此除了 TV 和 TGV_α^2 之间，所有结果在统计学上都有显著差异。从箱型图中也可以看出，NN 的平均 SER 最高 (29.1)，而 WT 的 SER 最低 (21.8)。

4.4.2 定量参数精度

表4.2的第 3 列和第 4 列分别显示了 5 种稀疏项重建结果的 K^{trans} 和 v_e 的 CCC。TV 和 TGV_α^2 都得到了最高的参数精度，其中 TV 的参数精度分别为 0.974 和 0.916，而 TGV_α^2 的精度分别为 0.974 和 0.917。虽然 NN 得到了最高的 SER (29.1)，但其参数精度却不是最高的，分别为 0.842 和 0.799。同样的现象也发生在 WT 与 FT 上，WT 的 SER (21.8) 比 FT (26.4) 低，但是其参数精度则比 FT 高。

图4.4显示了 K^{trans} 和 v_e 参数图肿瘤部分放大后的图像。在视觉上，TV 和 TGV_α^2 都精确地重建出 K^{trans} 和 v_e ，与全采样的参数图最接近，没有任何去噪与模糊的效果。而 NN 的参数图像则变得模糊，有过平滑的效果。

图4.5显示了 5 种稀疏项 K^{trans} 和 v_e 的 Bland-Altman 图。从图中可以看出，TV 和 TGV_α^2 得到了几乎无偏差的 K^{trans} 和 v_e ，而 WT、FT 和 NN 都低估了 K^{trans} 和 v_e 的值。虽然从表4.2中可以看出，WT 得到了相对准确地 CCC，但图4.5表明，WT 也严重低估了肿瘤部分的均值。

图4.6显示了全采样图像与重建图像肿瘤区域像素时间曲线的比较。图4.6 (a) 是肿瘤区域所有像素平均值的时间曲线，而图4.6 (b) 是肿瘤中心像素点 (96,96) 的时间曲

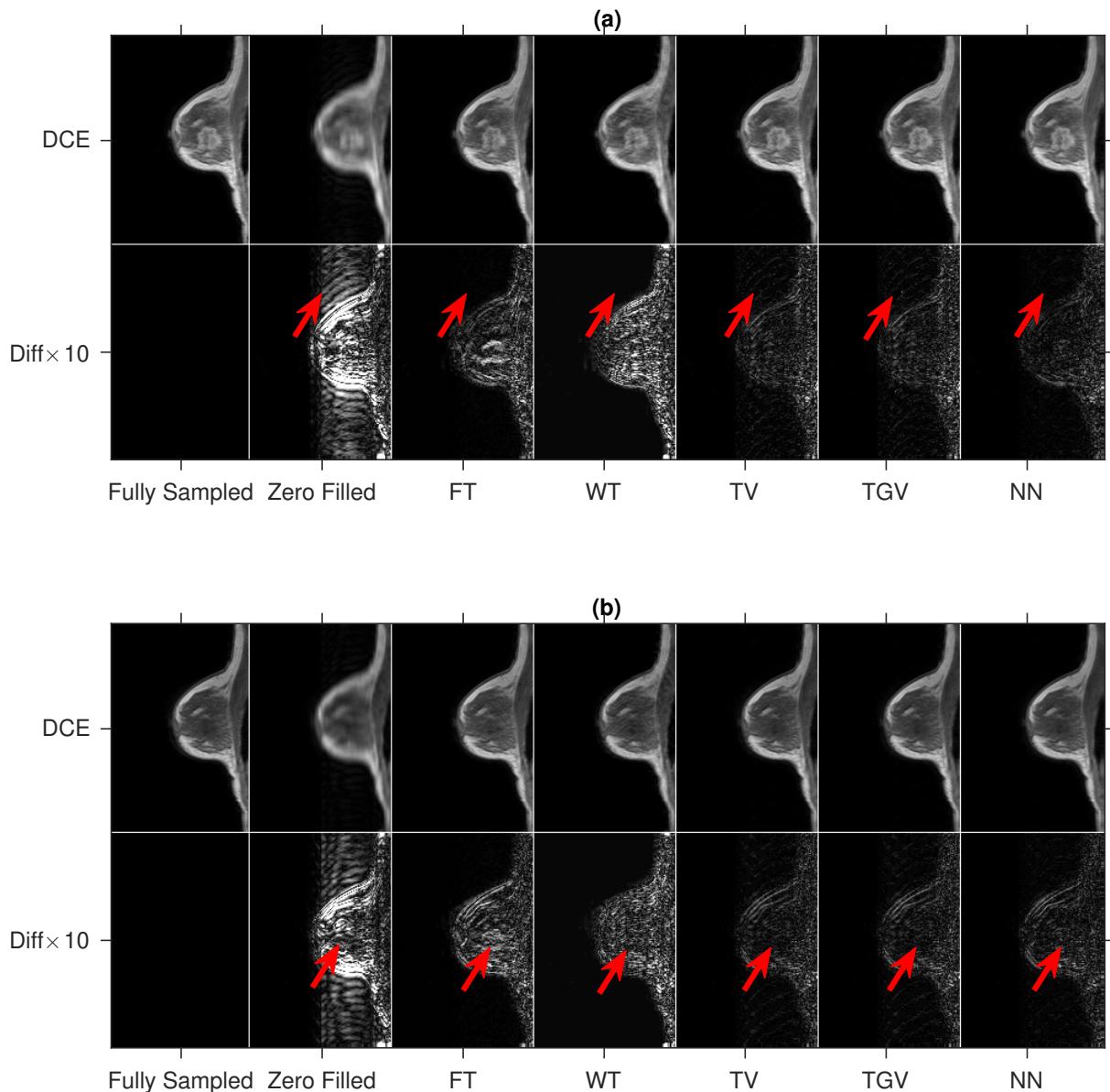


图 4.2 第一个 Cartesian 采样模式的重建结果。(a) 第 105 帧; (b) 第 1 帧。在每一个子图中, 第一行显示的分别是全采样、Zerofilled、WT、FT、TV、 TGV_α^2 和 NN 的重建图像, 第二行是相应的与全采样图像之间的差值图像。为了使得差值图像有更好的视觉效果, 我们将每一张差值图像的像素值都放大了 10 倍。图中的红色的箭头显示, 在视觉上 TV 和 TGV_α^2 模型重建肿瘤的效果是最好的, 而在 NN 去除背景伪影方面表现最好。

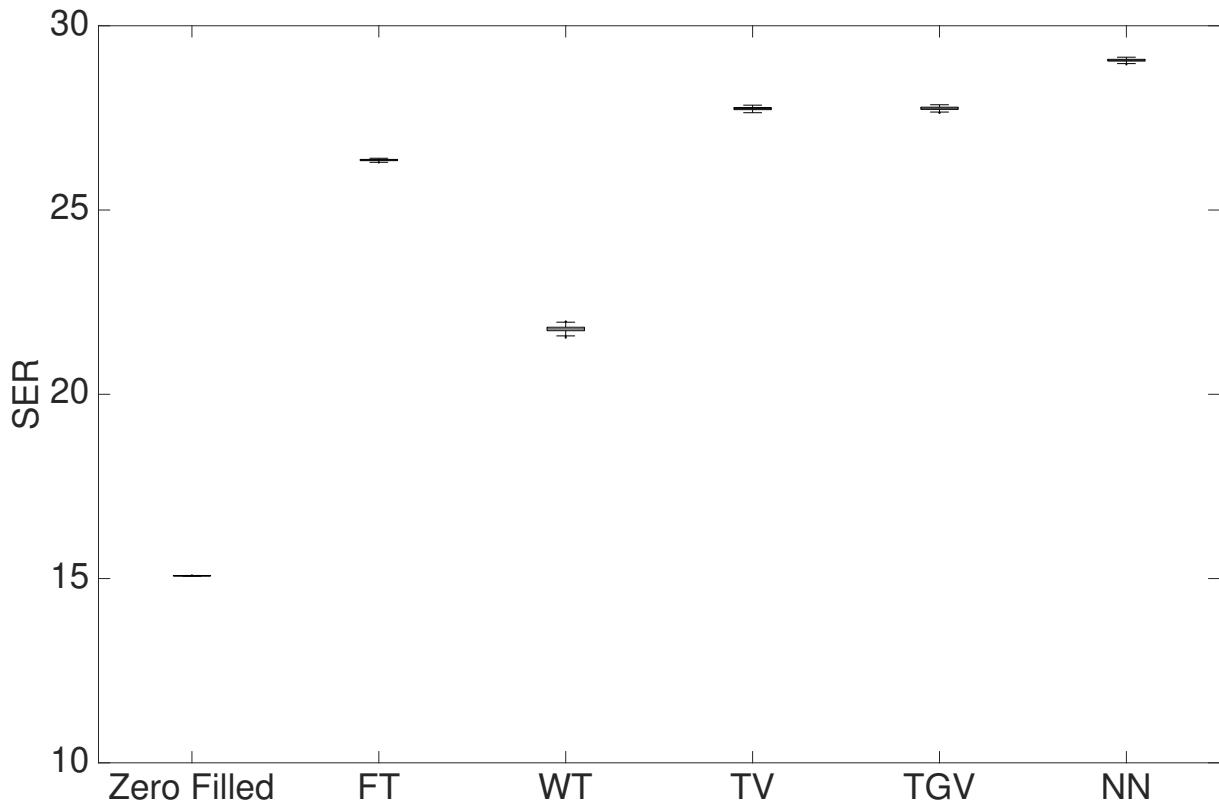


图 4.3 重建图像 SER 的箱型图。除了 TV 和 TGV_{α}^2 之间，所有结果在统计学上都有显著差异。从箱型图中可以看出，NN 的平均 SER 最高 (29.1)，而 WT 的 SER 最低 (21.8)。

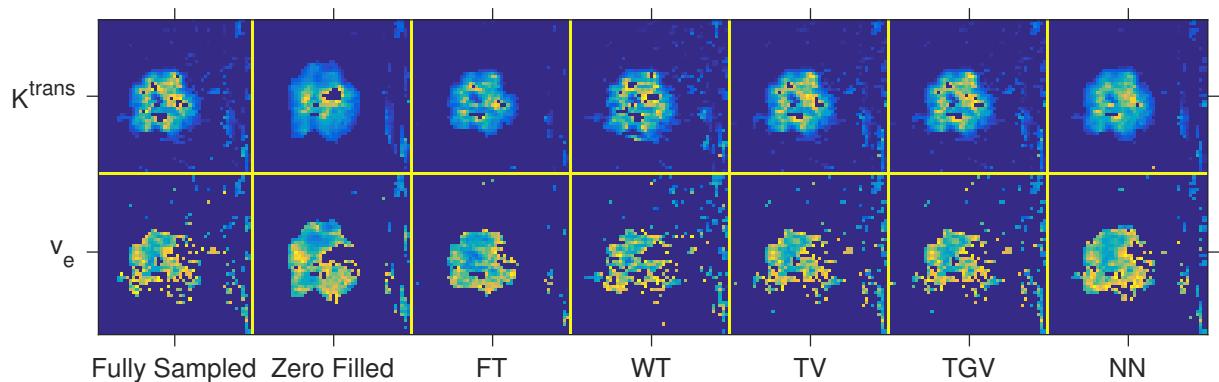


图 4.4 K^{trans} 和 v_e 参数肿瘤部分放大后的图像。第一行从左到右依次显示了全采样、Zerofilled、WT、FT、TV 和 TGV_{α}^2 的 K^{trans} 图像；第二行是相应的 v_e 图像。在视觉上，TV 和 TGV_{α}^2 与全采样的参数图像最接近。

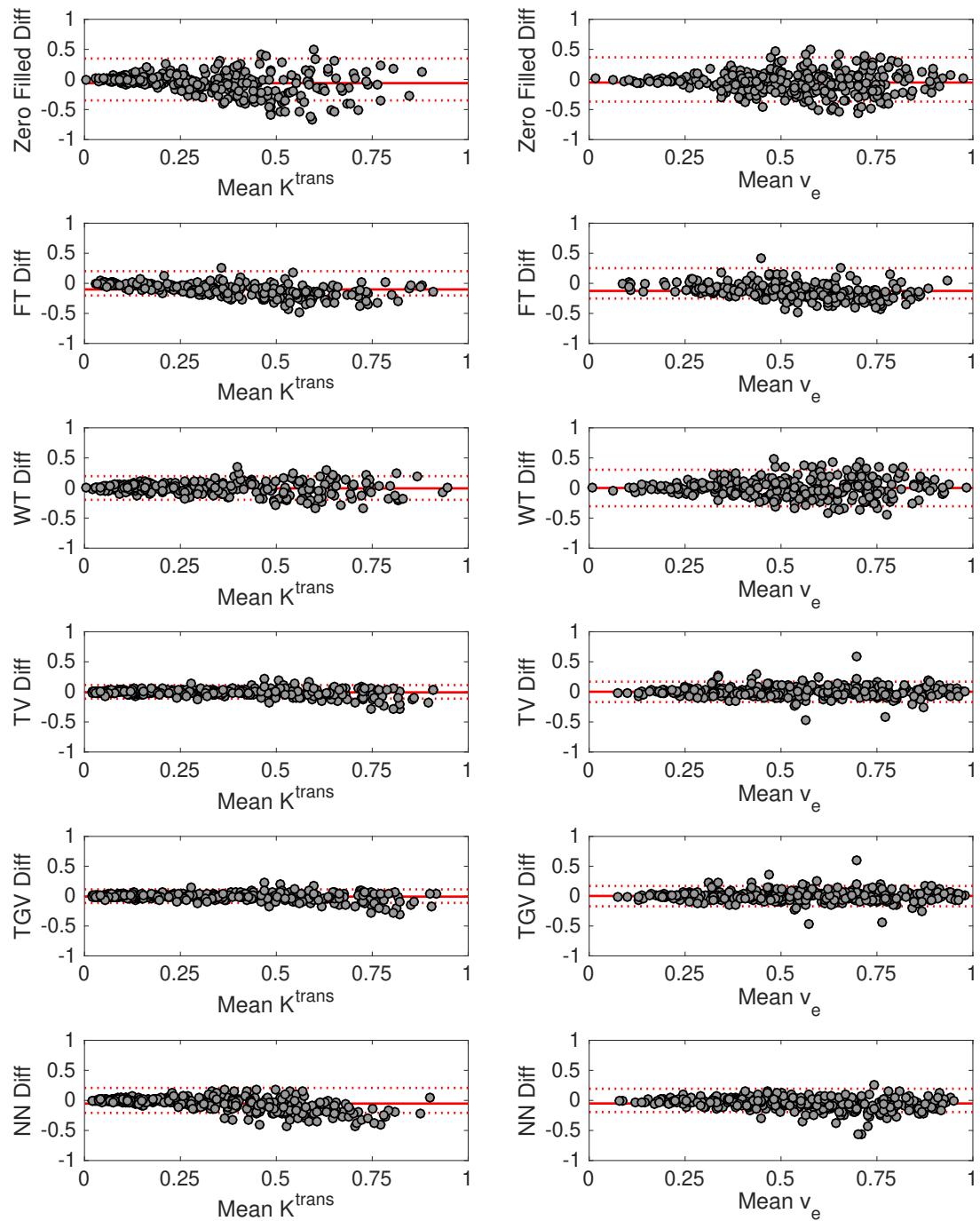


图 4.5 5 种稀疏项 K^{trans} 和 v_e 的 Bland-Altman 图。第一列为 K^{trans} , 第二列为 v_e 。从上到下依次为 WT、FT、TV、TGV 2 和 NN。TV 和 TGV 2 得到了最准确并且偏差最小的参数。FT 的偏差值最大, WT 的参数最不准确。NN 虽然低估了参数值, 但其偏差值较小。

线。从图4.6中可以看出, Zero-filled 图像最初 5 帧的重建结果是准确的, 但是当注入 CA 之后, 像素值就被低估了, 导致 SER、 K^{trans} 和 v_e 的减小。WT 低估了所有帧的肿瘤区域的平均像素值。在图4.6 (a) 中, FT 的最初 5 帧和最后 5 帧的拟合效果很差, 而且整体体现出了振荡的效应。这一点从图4.6 (b) 中可以更明显地观察到。 TGV_α^2 和 TV 最好地拟合了肿瘤区域平均像素值和单个像素值的时间曲线。在最初的几帧中, NN 高估了平均像素值, 这也影响了其最终 K^{trans} 和 v_e 的准确性。

由于在胸部 DCE-MRI 中, 医生更关心参数定量分析的准确性, 和 SER 一样, 我们也展示了 K^{trans} 和 v_e 的箱型图, 如图4.7所示。在所有的箱式图中, 四分位的范围都很小, 说明对于胸部 DCE-MRI, 基于 Cartesian 采样模式的压缩感知重建有可预测的精度。图4.7的前两个箱式图可以看出, 除了 TV 和 TGV_α^2 , 所有结果在统计上都有显著差异。与图4.6中的发现一致, TV 和 TGV_α^2 得到了最准确的 K^{trans} 和 v_e , 其 CCC 均值高方差小, 并且重建图像也最接近于全采样图像。同样, FT 在肿瘤均值与 CCC 方面最不准确, 而且肿瘤均值 K^{trans} 和 v_e 也与真实值相距甚远。虽然 WT 也得到了相对准确的肿瘤均值 K^{trans} 和 v_e , 但其方差较大, 说明其可预测性低。

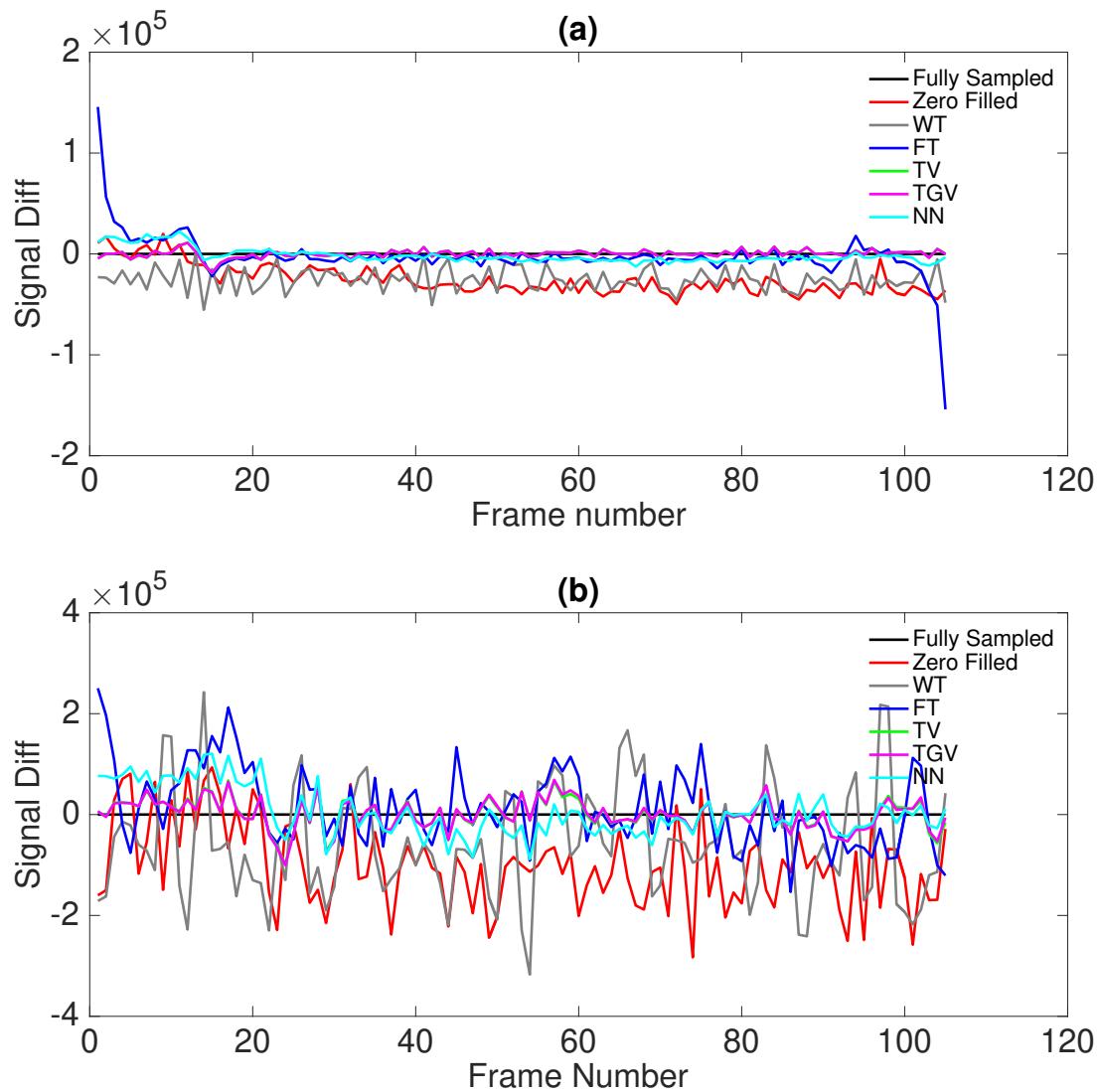


图 4.6 全采样图像与重建图像肿瘤区域像素时间曲线的比较图。(a) 肿瘤区域平均像素值的时间曲线比较; (b) 肿瘤中心像素 (96,96) 的时间曲线比较。图 (a) 显示 Zerofilled 图像在注入 CA 之后低估了肿瘤区域的平均像素值。WT 在整体上低估了平均像素值。FT 在最初和最后的 5 帧的时间曲线拟合很差, 而且整体上表现出了振荡的效应。TV 和 TGV_α^2 最好地拟合了时间曲线。NN 在注入 CA 之前高估了平均像素值, 并且在注入 CA 之后略微低估了平均像素值。图 (b) 表明 Zerofilled 图像在 CA 注入之后低估了肿瘤中心的像素值。FT 和 WT 均表现出振荡的效应, 使得曲线的拟合变得不准确。TV 和 TGV_α^2 的时间曲线表现平稳, 而 NN 在注入 CA 之前高估了肿瘤中心像素值。

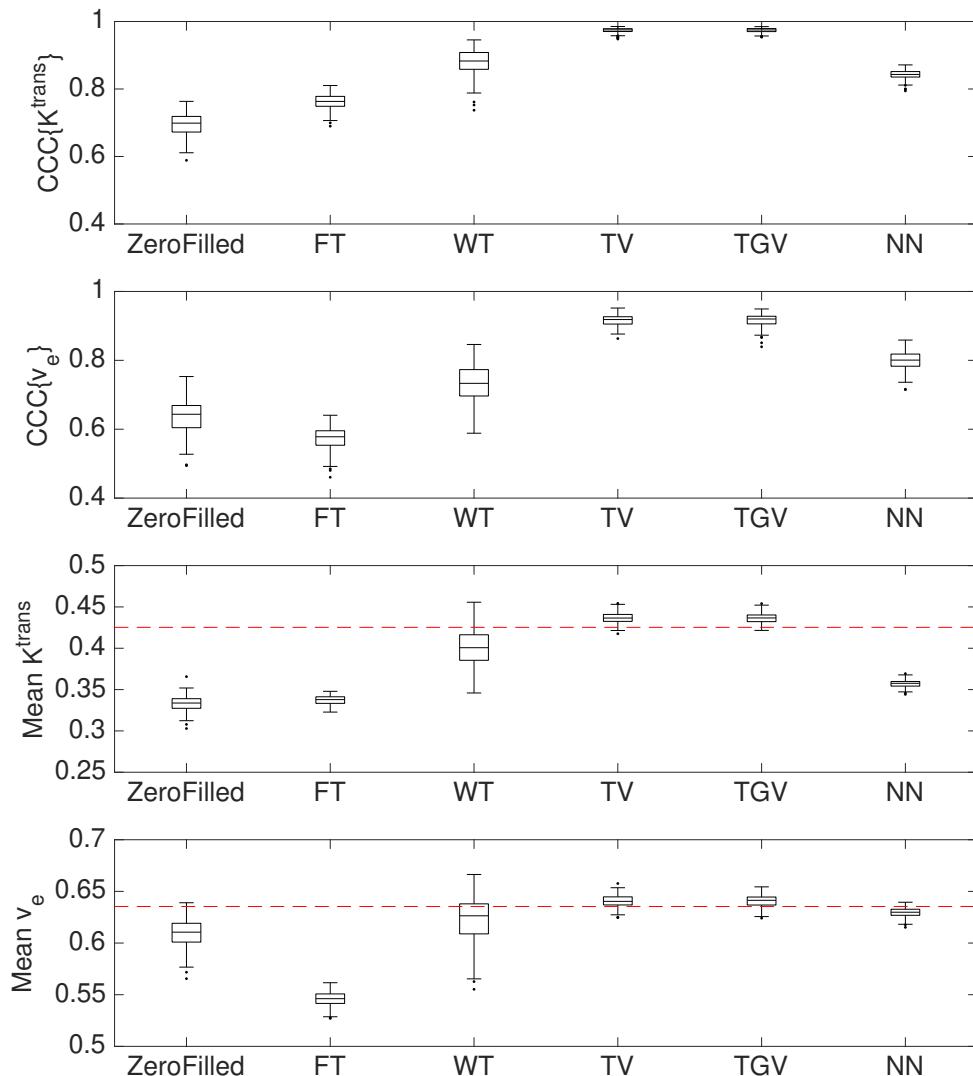


图 4.7 200 个不同 Cartesian 采样模式下 CCC 与肿瘤均值的箱式图。在前两个箱式图中，除了 TV 和 TGV_{α}^2 ，其他稀疏项的重建结果都在统计学上都有显著差异。我们可以观察到，TV 和 TGV_{α}^2 得到了最高的 CCC，这也和表3.1的结果一致。在后两个箱式图中，红色的水平线代表肿瘤平均 K^{trans} 和 v_e 的真实值，分别为 0.425 和 0.635。与表3.1的结果一致，TV 和 TGV_{α}^2 得到了最准确的肿瘤平均 K^{trans} 和 v_e ，也与真实值最接近。有趣的是，NN 倾向于低估 K^{trans} 的值但是却相对准确地得到了 v_e 。

4.5 数值实验分析

定量分析表明了 5 种时间方向的稀疏项在胸部 DCE-MRI 的压缩感知重建中表现的不同。其中, TV 和 TGV_{α}^2 得到了最准确的 K^{trans} 和 v_e , 并且它们在肿瘤区域的时间曲线也最贴合真实值。这表明拟合像素值时间曲线的误差可以预测最终量化参数的精度。如果是这样, 在没有真实值的前瞻性成像实验中, 我们可以用拟合的残差来估计参数的精度。

图4.7中的箱式图表明, TV 和 TGV_{α}^2 之间没有显著性差异, 所以很难区分这两种稀疏项在实际中的应用。但由于本文只针对胸部 DCE-MRI 图像, TV 和 TGV_{α}^2 可能在其他部位的图像上有着不同的表现, 尤其是当图像存在线性变化的梯度而不是分片常数时。

本文给出的定量参数分析结果应当告知临床并于研究影响重建方法。对于那些需要得到更加精确整体质量的应用, 最好的稀疏项选择是 NN。而另一方面, 如果研究者需要得到更精确的参数分析, 最好的选择是 TV 和 TGV_{α}^2 。

在本文之前, 没有工作量化地比较和研究不同时间方向的稀疏项对胸部 DCE-MRI 压缩感知重建的影响。本文填补了这个空白, 但是有三点主要事项。第一, 我们在数值实验中仅仅使用了胸部 DCE-MRI 数据, 如果使用其他数据, 数值实验的结果和结论可能会有变化, 但预计不会很显著。第二, 我们在数值实验中只使用了 200 个 Cartesian 采样模式, 如果使用 non-Cartesian 采样模式, 数值实验结果也可能不同。第三, FISTA 的调参过程是一个不精确的过程, 如果使用不同的参数, 数值实验结果也可能会有轻微的不同, 但我们认为整体的结论是不变的。

4.6 本章小结

在这一章节里, 我们针对胸部 DCE-MRI 图像, 量化地比较了五种时间方向的稀疏项在压缩感知重建模型中的表现。我们发现对于胸部 DCE-MRI 图像, FT 是最不适合的稀疏项, 因为 DCE-MRI 图像在时间上没有周期性。WT 虽然在参数分析上结果比较精确, 但统计上方差较大, 而且其重建图像的信噪比最低。NN 最好地抑制了由下采样产生的伪影, 因此有着最高的信噪比, 但是在参数分析上不太精确。TV 和 TGV_{α}^2 得到了最精确的参数分析, 且二者的结果在统计上没有明显差异。

5 基于 GPU 的实时 MRF 字典生成与参数图重建

磁共振指纹 (MRF) 是定量 MRI 技术的新方法, 可以在单次扫描中同时获得多个组织参数^[3,165,166], 因此 MRF 有着加速 MR 成像的潜力。MRF 主要包括三个部分, 分别为信号采集、字典生成和模式识别。其中信号采集的速度取决于 MR 自身扫描的速度, 而字典生成和模式识别则取决于算法和硬件。因此一个快速开源的 MRF 重建程序可以增加 MRF 临床应用的可行性和准确率。

5.1 MRF 重建参数图的主要问题

MRF 重建参数图的一个重要过程是生成包含不同组织参数模拟信号的字典。MRF 重建的准确性很大程度上取决于模拟信号的准确性。在 MRF 最早的文章中, Ma 等^[3] 使用了 Bloch 方程来模拟单个等色子在 MR 序列中的演化。模型假设指纹中的每个体素都由单一等色子构成, 可以很好地模拟 bSSFP 序列的演化过程。但是当处理 uSSFP 序列时, Bloch 模型必须将所有等色子考虑进去。这是因为在 uSSFP 中, 由于非平衡梯度场的存在, 指纹中的每个体素都是多个等色子的平均。因此 Bloch 方程在处理 uSSFP 时的效率不高。

扩展相图是用来模拟 MR 成像的另一个模型, 曾经被用在^[126] 中。有关 EPG 的详细信息请参考第2.2小节。EPG 利用 Fourier 变换将自旋系统描述为多个离散的配置状态。相对于 Bloch 方程, EPG 可以提供更准确的信号演化, 尤其是当自旋系统受到非均匀磁场的影响时, 比如在 uSSFP 中常用的破坏梯度场。但是 EPG 模型的计算很耗时, 因为为了确保结果的准确性, 对于每个模拟信号我们都需要追踪 100 个甚至更多的配置状态。这在程序中体现为一个维度很高的矩阵, 对计算速度和计算机存储都是很大的负担。如果只在 CPU 上运行, 在我们自己的硬件上生成一个和 Ma 等^[3] 使用的同样大小的字典需要几个小时。所以增加模型准确性的代价是巨大的。

除了字典的生成, 模式识别也非常耗时。模式识别是将采集到的指纹数据和字典进行匹配, Ma 等^[3] 使用了模板匹配的方法来进行参数图的重建。记 $X = \{x_n \in \mathbb{C}^L\}, n = 1, \dots, N$ 为采集到的指纹数据, $\mathcal{D} = \{d_k \in \mathbb{C}^L\}, k = 1, \dots, K$ 为生成的字典。那么模板匹配即为从字典 \mathcal{D} 中选取和 x_n 内积最大的元素:

$$\hat{k}_n = \arg \max_k |\langle \mathbf{d}_k, \mathbf{x}_n \rangle|. \quad (5.1)$$

并且质子密度也可以同时计算:

$$\hat{\rho}_n = |\langle \mathbf{d}_{\hat{k}_n}, \mathbf{x}_n \rangle|. \quad (5.2)$$

不难看出, 模板匹配的运行时间与字典大小 K , 时间点的个数 L 以及指纹数据中体素的个数有关, 计算复杂度为 $O(NL^2K)$ 。当字典很大, 指纹数据很多的时候, 重建参数

图所消耗的时间是许多快速成像方法无法容忍的。比如，Ma 等在^[3] 中使用生成了一个大小为 566,783 的字典，其中时间点为 1,000，并将其和 128×128 个体素进行匹配。在他们自己的硬件上，使用 MATLAB 程序运行，字典生成花费了 399 s，而匹配花费了大约 3 分钟。

在第1.4节中，我么回顾了 MRF 加速的常用方法。其中一大类是基于降维的方法，通过降低字典或者数据的维度，从而达到加速计算的目的。常用的策略是压缩字典在时间方向的维度，比如 MaGiney^[131] 利用奇异值分解来压缩字典在时间方向的维度，使得模板匹配得到了 3.4–4.8 倍的加速。但是大矩阵奇异值的计算本身计算消耗就很大，而且在依然需要先生成一个字典。Cauley 等^[130] 提出了分组的思想来处理字典，先将字典中的元素聚类，再用主成分分析来压缩字典在时间和参数方向的维度。这两种方法均没有加速字典生成，而且得到的结果只是模板匹配的近似，因而存在误差。

另一种加速方法是使用神经网络来学习重建参数图^[137]，使用神经网络对参数图直接推断比在整个字典上进行模板匹配快得多。但是基于学习的方法有一个缺陷就是，学习和训练的效果很大程度上取决于样本的好坏，即网络的训练需要生成字典和其匹配的结果。如果成像序列发生了轻微的变化，比如偏转角、 T_R 或者 T_E 发生了变化，我们必须重新生成字典并且重新训练网络。理想的解决方式是每当采样序列发生变化时，都能够快速甚至即时的生成新字典。如果能将字典生成和匹配的程序适当地并行运行，这样的解决的方式是可以达到的。

5.2 基于 GPU 的 MRF 重建

目前的方法均从模型和算法的角度来提高 MRF 重建参数图的速度，而另一个思路是通过合理利用计算机硬件，编写并优化程序来加速 MRF 重建。Malik 等^[167] 发布了一款基于 MATLAB 语言的程序 EPG-X，可以为不同的 MR 序列进行建模。但是由于他们的程序运行在 CPU 上，速度很慢，尤其是字典生成的速度。同时，EPG-X 不能够处理变化回波时间（variable T_E ）或者变化重复时间（variable T_R ）的序列，而这两种序列在 MRF 中经常遇到。此外，EPG-X 也不能处理 B_1^+ 效应，我们可以从后面的数值实验中看到，考虑 B_1^+ 可以提高参数图的准确性。Cloos^[133] 等编写了一款基于 C++ 语言的程序 PnP-MRF，可以模拟整个 MRF 重建的过程。PnP-MRF 生成字典的速度虽然较快，但模板匹配的速度很慢。而且 PnP-MRF 只使用了 Bloch 方程来生成字典，不能实现 EPG 模型，并且和 EPG-X 一样，PnP-MRF 不能够实现 B_1^+ 校正。

最近几年，图形处理单元（GPU）逐渐成为加速科学计算的主流方法，并在医学成像领域有着广泛的应用，详细的综述请参考文献^[168]。GPU 中集成了大规模并行计算单元，在允许计算复杂模型的同时，保持合理的执行时间。GPU 有以下两个主要特点：第一，与传统 CPU 芯片不同，GPU 中用于浮点运算的空间要远远多于指令分配，因此具有强大的科学计算潜力；第二，可编程的特性，即用户可以通过应用程序接口直接在 GPU 上进行编程，体现了很好的交互能力。这两个特点奠定了 GPU 并行计算的基石。

目前应用最广泛的应用程序接口是 Nvidia 公司的 CUDA 框架，其提供的编程能力在显著地发展，程序的复杂度也在慢慢地减少。基于 CUDA 框架的 GPU 编程需要注意以下三点。第一，需要对 GPU 硬件有足够的了解，这样才能充分发挥 GPU 的潜力。第二，合理地为程序中的变量分配显存，使得线程尽量可以访问连续的显存空间，这样可以大大减少线程访问显存的时间。第三，避免条件竞争的发生，即避免不同线程同时写入同一显存位置，从而造成不确定的结果。简而言之，GPU 编程是一个根据使用的硬件不断调试和优化的过程，其通常能够使得程序运算速度提高 10–100 倍。

目前，GPU 并行计算已经被广泛地应用于医学成像领域，如 CT^[169,170]、MRI^[171–175]、US^[176,177]、PET^[178,179] 等。但是据我们所知，目前没有程序在 CUDA 的框架下，利用 GPU 来实现 MRF 的字典生成和模板匹配。另外，也有一些文献中也表明了对缺少科学的研究的可重复性的担忧^[180–182]。针对以上问题，我们开发了一款基于 GPU 的开源 MRF 重建程序 snapMRF，用于并行计算字典生成和模板匹配。snapMRF 将字典生成和模板匹配全部放在 GPU 上并行计算，并且给出了一系列单元测试来保证程序的可靠性和准确性。snapMRF 既可以用 Bloch 方程也可以用 EPG 模型来进行字典生成，保证了程序的普适性。

5.3 snapMRF 程序设计与单元测试

5.3.1 snapMRF 程序流程

snapMRF 实现 MRF 重建参数图的过程总体一共可以分为 5 个步骤，如图5.1所示。其中前两个步骤为指纹数据与成像参数的输入，后面的两个步骤为字典的生成与匹配，最后一个步骤为结果的输出。这样的设计框架由 GPU 计算的原理所决定的。一般来说，利用 GPU 进行并行计算需要首先将数据从主机（host）拷贝到设备（device）上，在设备上进行计算之后，再将处理之后的数据从设备拷贝回主机上进行存储或其他处理。这样做是为了尽可能避免数据在主机和设备之间进行多次拷贝，减少时间消耗。snapMRF 并行计算的部分主要发生字典生成和匹配两部分，下面我们进行详细的介绍。注意下面变量前面的 * 表示 C 语言中的向量指针，其中主机向量指针以字母 h 开头，设备向量指针以字母 d 开头。

- 从 CSV 文件中读取 MR 序列的成像参数，将其存储在主机向量 `*h_mrf`，并拷贝到设备向量 `*d_mrf` 中。这个 CSV 文件中包含了 MR 序列的信息，即偏转角 α ，相位角 Φ ，回波时间 T_E 和重复时间 T_R 。时间点的个数储存在变量 `nreps` 中。
- 从命令行读取字典的配置信息 (T_1, T_2, B_0, B_1^+)。使用函数 `parse_length()` 解析命令行，并使用函数 `parse_params()` 将其分别储存在主机向量 `*h_t1, *h_t2, *h_b0` 和 `*h_b1` 中。为了之后处理方便，这四个主机向量被函数 `trans_params()` 合并成一个向量，并拷贝到设备向量 `*d_params`。在 `*d_params` 中，配置信息的存储顺序是 T_1 优先的，即 T_1 的指标变化最快，然后是 T_2 ，然后是 B_0, B_1^+ 的指

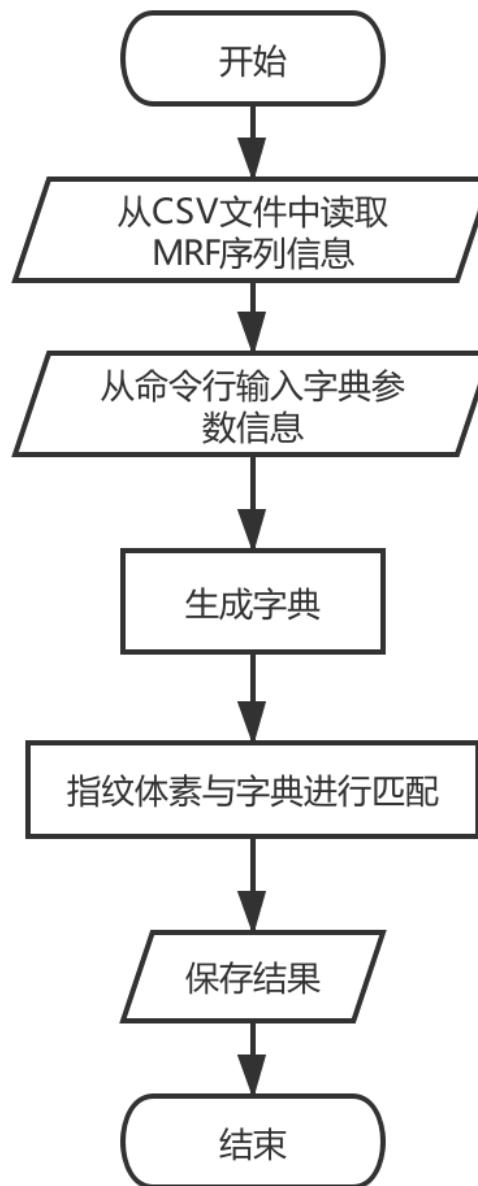


图 5.1 snapMRF 程序流程图。

标变化最慢。配置信息的个数分别存放在变量 `l_t1`, `l_t2`, `l_b0` 和 `l_b1` 中。字典中元素的个数通过函数 `compute_natoms()` 计算得到, 存储在变量 `natoms` 中。由于在实际情况中, T_1 的值一般远大于 T_2 的值, 所以 $T_1 \leq T_2$ 的情况被移除了。

- **生成字典**, 使用的函数为 `MRF_dict()`, 并将结果储存在设备向量 `*d_atoms` 中。这是程序设计中最有难度的一个步骤。对于每个 B_1^+ 和每个时间点, 设备函数关于字典中元素的指标是并行计算的。也就是说, 字典中所有元素的信号演化是并行计算的, 即我们先计算所有元素在第一个时间点的信号大小, 更新每个元素的状态矩阵 `*d_w`, 之后计算下一个时间点的信号大小, 直至所有时间点都计算结束。其中状态矩阵 `*d_w` 包含了字典中所有元素的状态矩阵, 其维度为 $(3 \times \text{nstates}) \times \text{natoms}$, 其中每个元素的状态矩阵的维度为 $3 \times \text{nstates}$ 。由于状态矩阵的维度通常很大, 这会给显存带来很大的负担, 并且造成计算资源的浪费。因此, 字典生成结束后, 我们需要将 `*d_w` 释放。字典中元素所包含的参数信息存储在 `*d_params` 中, 顺序与字典 `*d_atoms` 相同。
- **将指纹体素与字典匹配**。从 RawArray^[183] 文件中读取采集好的指纹数据, 存储在向量 `*d_img` 中。函数 `MRF_match()` 分两步计算参数图。第一步, 对于每个包含不同 B_1^+ 的子字典, 计算指纹数据在子字典下的参数图, 并将其储存在 `*d_maps` 中。当字典过大时, GPU 的显存可能有限。为了最大程度地利用 GPU 的显存, 我们首先计算程序运行到此处时 GPU 所剩显存, 并用函数 `compute_nsplits()` 根据所剩显存将指纹数据分为大小相等的组, 在每个小组上进行匹配。函数 `MRF_minimatch()` 用来计算每个小组的指纹数据的参数图。同时为了优化匹配速度, 我们将指纹数据看做一个矩阵, 则指纹与字典的内积则可以由指纹与字典的共轭转置相乘得到。这样做的目的是为了使用高度优化的 cuBLAS 软件库中矩阵相乘的函数。然后利用函数 `generate_maps()` 从内积矩阵中生成参数图。第二步, 使用 `merge_maps()` 将不同 B_1^+ 生成的参数图融合起来, 生成最终的参数图。这个设备函数关于体素是并行计算的。对于每个体素, 由于在第一步中产生了 `l_b1` 组参数, 选择与校正 B_1^+ 最接近的那个 B_1^+ 的值, 并将其相对应的 T_1 , T_2 和 B_0 作为最后的参数, 最接近校正值的 B_1^+ 也被保存了下来。最后的参数图在 `*d_maps` 中, 以 T_1 , T_2 , B_0 , B_1^+ 和质子密度的顺序存储。
- **保存结果**。将字典元素 `*d_atoms` 和参数图 `*d_maps` 从 GPU 拷贝回 CPU, 并存储成为 RawArray 格式。释放内存。

5.3.2 snapMRF 核函数的功能

在 Nvidia 的计算框架 CUDA 中, 将在 GPU 上并行计算的特殊函数称为核函数 (kernel), 将该函数运行的一个实例称为一个线程 (thread)。因此 GPU 并行计算编程的一个原则是将复杂的算法分解, 使得算法的每一步都可以在 GPU 上通过核函数并行运行。我们下边简单介绍 snapMRF 中核函数的主要功能, 如表5.1所示。

- `init_rf_pulse()` 默认初始化磁化向量大小为状态 $[0, 0, 1]^T$ 。如果使用的序列为 uSSFP，对初始磁化进行反转衰减 (inversion decay)，衰减时间为 TI ms。核的输入为字典大小 `natoms` 和状态的个数 `nstates`，输出为状态矩阵 `*d_w`。
- `fill_transition_matrix()` 根据偏转角和相位角计算转移矩阵 `*d_T_m` 中的每一个元素，其中偏转角和相位角储存在向量 `*d_mrf` 中。转移矩阵 `*d_T_m` 对应 EPG 算法中的 T 算子，它的作用是模拟射频场对磁化的影响。
- `apply_rf_pulse()` 将转移矩阵 `*d_T_m` 作用于状态矩阵 `*d_w` 上。这里单精度复值矩阵乘法使用 CUDA 自带的 cuBLAS 函数库中的函数 `cublasCgemm()` 实现。
- `shift_phase()` 磁化的水平分量相位平移作用在状态矩阵 `*d_w` 上。这个核只有在 bSSFP 序列中用到。
- `save_atom()` 从状态矩阵 `*d_w` 中读取最近的回波强度，并储存在字典 `*d_atoms` 中。输入为序列中时间点的个数 `nreps` 和当前回波所在的时间点 `index`。
- `decay_signal()` 将 T_1 和 T_2 衰减作用在状态矩阵 `*d_w` 中并把结果保存在状态矩阵中。这里衰减的时间有四种不同的选择，分别为： T_E (`type=1`)， $T_R - T_E$ (`type=2`)， $T_E/2$ (`type=3`) 和 T_R (`type=4`)。 T_E 和 T_R 存储在向量 `*d_mrf` 中，而 T_1 和 T_2 存储在向量 `*d_params`。这个核相当于 EPG 模型中的 E 算子。
- `dephase_gradients()` 将失相效果作用于状态矩阵 `*d_w` 上，并将结果保存回状态矩阵。这个核相当于 EPG 模型中的 H 算子。

因此，结合流程图5.1和如上核函数的功能，snapMRF 字典生成与匹配的详细流程如算法7所示。

5.3.3 snapMRF 并行计算的难点与创新

在这一小节，我们讨论 snapMRF 程序设计中遇到的难点。一般来说，为了能够在 GPU 上进行并行计算，我们需要认真考虑以下两个方面。一方面是内存合并 (memory coalescence)，其包括排列所需处理的数据，使得给定组内的线程可以访问连续的显存

表 5.1 snapMRF 中的核函数

核函数	输入	输出	算子
<code>init_rf_pulse()</code>	<code>natoms, nstates</code>	<code>*d_w</code>	-
<code>fill_transition_matrix()</code>	<code>*d_mrf</code>	<code>*d_T_m</code>	T
<code>apply_rf_pulse()</code>	<code>*d_w, *d_T_m</code>	<code>*d_w</code>	-
<code>shift_phase()</code>	<code>*d_w</code>	<code>*d_w</code>	-
<code>save_atom()</code>	<code>*d_w, nreps, index</code>	<code>*d_atoms</code>	-
<code>decay_signal()</code>	<code>*d_w, *d_params</code>	<code>*d_w</code>	E
<code>dephase_gradients()</code>	<code>*d_w</code>	<code>*d_w</code>	H

算法 7 snapMRF 生成字典与匹配详细流程。

输入: *d_mrf, *d_params, *d_img

输出: *d_atoms, *d_maps

- 01: 从 CSV 文件中读取 MRF 序列信息, 存入 *d_mrf;
 - 02: 从命令行输入字典参数信息, 存入 *d_params;
 - 03: 初始化状态矩阵 *d_w;
 - 04: **迭代:** 从第 1 个时刻到第 L 个时刻, 并行计算所有体素的信号
 - 05: 使用函数 `fill_transition_matrix()` 构造转移矩阵 T ;
 - 06: 使用函数 `apply_rf_pulse()` 将射频场作用在 *d_w 上;
 - 07: 使用函数 `decay_signal()` 将 T_1 和 T_2 衰减作用在 *d_w 上;
 - 08: 使用函数 `save_atoms()` 将原子的信号保存在 *d_atoms 中;
 - 09: 使用函数 `dephase_gradients()` 将梯度场作用在 *d_w 上;
 - 10: 使用函数 `decay_signal()` 将 T_1 和 T_2 衰减作用在 *d_w 上;
 - 11: **终止迭代;**
 - 12: 释放 *d_w;
 - 13: 从 RawArray 文件中读取指纹数据, 存入 *d_img;
 - 14: 计算剩余显存大小, 并根据剩余显存, 将 *d_img 分为 G 组;
 - 15: **迭代:** 从第 1 组到第 G 组, 在每一组内并行计算所有体素的参数
 - 16: 使用函数 `MRF_minimatch()` 进行匹配;
 - 17: 使用函数 `generate_maps()` 生成参数图;
 - 18: **终止迭代;**
 - 19: 将 *d_atoms 和 *d_maps 保存为 RawArray 文件;
 - 20: 释放所有显存和内存。
-

位置。这样做可以加快线程访问显存的速度，节省计算的时间。另一方面是在 GPU 进行并行计算时要避免条件竞争（race condition），其意思为两个或多个的线程同时写入同一个显存位置，造成不可预测的结果。为了实现以上两点，在 snapMRF 程序设计中，我们遇到了以下几方面挑战：

第一，确定多维向量指针在显存中的储存结构，比如指纹数据 `*d_img`、字典 `*d_atoms`、成像参数 `*d_params` 等。这意味着我们需要确定一个指标变化的优先级，即多维向量指针指标变化的快慢。在 snapMRF 中，我们的处理方式如下：对于包含时间维度的向量指针，如 `*d_atoms`、`*d_mrf`、`*d_img` 等，我们选取时间维度优先，即时间方向的指标变化最快。对于包含参数维度的向量指针，如 `*d_maps`、`*d_params` 等，我们选取参数维度优先，即参数方向的指标变化最快。这样处理的原因是数据通常是在时间或者参数方向上连续变化的。

第二，计算多维向量在显存中对应的一维指标。字典、指纹等数据虽然为多维数据，但其在显存中的储存和访问只能是一维线性的。因此，多维数据的指标和其在显存中的一维指标需要一一对应起来。这个过程并不容易且至关重要，因为一旦指标计算错误，最后的结果会变得无法预测且这种错误很难被发现。

第三，将 EPG 算法生成字典的过程分裂为小的 GPU 核函数。如第 2.2 小节所述，EPG 算法包含了多个作用矩阵，如转移矩阵 T 、衰减矩阵 E 、平移矩阵 H 等。在 snapMRF 中，我们将 EPG 算法以作用矩阵为最小单位分成了多个 GPU 核函数，即每个作用矩阵对应一个核函数，这样使得程度有了更好的可读性。

第四，确定一个合适的并行方案。由于 EPG 模型模拟了信号随着时间的演化，因此在 snapMRF 中，并行是相对于字典中原子而言的。即在进行字典生成时，在每一个时间点，并行计算所有原子的信号演化，这个过程一直持续到所有时间点计算完成为止。在进行字典匹配时，并行是相对于指纹数据的体素而言。即在进行字典匹配时，并行计算所有体素的参数并生成参数图。

第五，管理显存。为了使得 GPU 的显存得到充分的利用，我们需要时刻注意剩余显存的大小，在不超过显存上限的前提下，最大限度的利用显存。snapMRF 中，在进行字典匹配之前，我们首先释放了 `*d_w` 的显存，并且计算了此时剩余显存的大小，然后根据剩余显存的大小来为后面的数据分配显存。换而言之，我们将根据剩余的显存，将指纹数据均匀地分成了几组，在组内进行字典匹配。这样做的原因是字典和指纹数据通常需要占据很大的显存，因此通常没有足够的显存同时计算所有的体素参数图。

第六，避免数据在 GPU 与 CPU 之间多次拷贝。snapMRF 中主机与设备之间的数据拷贝只发生两次，分别为将 MR 成像序列和指纹数据从 CPU 拷贝到 GPU 和将计算结果从 GPU 拷贝到 CPU，其余所有计算均在 GPU 上进行，最大限度地减少了数据拷贝的时间开销。

5.3.4 单元测试

单元测试是简单的情形的测试，目的是为验证和确保程序的正确性。单元测试的正常结果是全部通过。如果有一个或者多个单元测试失败，那么意味着程序或者算法有错误。snapMRF 中给出了 8 组单元测试，每组单元测试对应一个简单序列的字典生成或者模板匹配，并与其相对应的解析解进行比较。所有的单元测试都包含在源文件 `test.cu` 中。Table 5.2 显示了不同序列所用到的参数。单元测试共使用了 6 个不同的序列，每个序列都有不同的参数。 $[a, b]$ 意思是参数是在区间 a 与 b 中随机选择。为了简化计算，对于所有的序列， $B_0 = 0$ 并且 $B_1 = 1$ 。序列的详细信息请参考文献^[129] 和^[184]。

第一个单元测试的序列是一个简单的自旋回波，由一个 90° (偏转角 $\alpha = \pi/2$, 相位角 $\Phi = \pi/2$) 和一个 180° (偏转角 $\alpha = \pi$, 相位角 $\Phi = 0$) 射频脉冲构成。其中 $T_1 = 600$ ms, $T_2 = 100$ ms, $T_E = 25$ ms, $T_R = 1000$ ms。通过简单计算可以得知，这个回波产生的信号为 $(e^{-T_E/T_2})^2 = [\exp(25/100)]^2 = 0.6065$ 。snapMRF 中相应的函数为 `epg_se()`。注意这里 T_E 的含义与传统自旋回波不同。在传统自选回波中， T_E 是指 90° 射频脉冲到读取回波的时间。但是在 snapMRF 中，由于程序是由稳态序列优化，这里的回波时间是指 180° 射频脉冲到读取回波的时间。如果想把这里的 T_E 转换为传统的自旋回波 T_E ，只需要将其除以 2。同样的情况也是用与快速自旋回波。

第二个单元测试是饱和恢复序列 (saturation recovery)，由 60° 相隔时间相同 (偏转角 $\alpha = \pi/3$, 相位角 $\Phi = \pi/2$) 的射频脉冲构成。这里其中 $T_1 = 600$ ms, $T_2 = 100$ ms, $T_E = 1$ ms, $T_R = 500$ ms。经计算得到解析解为 $[0.857, 0.674, 0.631, 0.622, 0.620, 0.620, 0.620, 0.620, 0.620]$ 。snapMRF 中相应的函数为 `epg_sr()`。可以看出，在经历了几个脉冲序列之后，回波强度达到了稳态。

第三个单元测试是无驰豫快速自旋回波序列 (fast spin echo)，由一个 90° (偏转角 $\alpha = \pi/2$, 相位角 $\Phi = \pi/2$) 和一系列 120° (偏转角 $\alpha = 2\pi/3$, 相位角 $\Phi = 0$) 射频脉冲构成。既然驰豫被忽略，在理论上 T_1 和 T_2 的值应为正无穷。因此，在程序中 T_1 和 T_2 并未被真正用到。我们只计算了三个回波，最后回波的强度为 $[0.750, 0.938, 0.844]$ 。

表 5.2 单元测试中使用的参数

序列类型	α (deg)	Φ (deg)	T_1 (ms)	T_2 (ms)	T_E (ms)	T_R (ms)
SE	$\frac{\pi}{2} \rightarrow \pi$	$\frac{\pi}{2} \rightarrow 0$	600	100	25	1000
SR	$\frac{\pi}{3} \rightarrow \dots$	$\frac{\pi}{2} \rightarrow \dots$	600	100	1	500
FSE	$\frac{\pi}{2} \rightarrow \frac{2\pi}{3} \rightarrow \dots$	$\frac{\pi}{2} \rightarrow 0 \rightarrow \dots$	-	-	0	0
FSE+relax	$\frac{\pi}{2} \rightarrow \pi \rightarrow \dots$	$\frac{\pi}{2} \rightarrow 0 \rightarrow \dots$	600	100	25	50
FISP	$\frac{\pi}{6} \rightarrow \dots$	$0 \rightarrow \dots$	1000	100	5	10
SSFP	$[0, \pi/3]$	$0 \rightarrow \dots$	100,200	20,40	$[3.5, 7.5]$	16

snapMRF 中相应的函数为 `epg_tse()`。

第四个单元测试是有驰豫快速自旋回波序列，由一个 90° （偏转角 $\alpha = \pi/2$, 相位角 $\Phi = \pi/2$ ）和 8 个 180° （偏转角 $\alpha = \pi$, 相位角 $\Phi = 0$ ）射频脉冲构成。这里 $T_1 = 600$ ms, $T_2 = 100$ ms, $T_E = 25$ ms, $T_R = 50$ ms。最后的回波为 $[0.607, 0.368, 0.223, 0.235, 0.082, 0.0500, 0.0302, 0.0183, 0.0111, 0.0068]$ 。snapMRF 中相应的函数为 `epg_tse()`。

第五个单元测试是 F_0 -type SSFP 序列（也成为 FISP），由一系列 30° （偏转角 $\alpha = \pi/6$, 相位角 $\Phi = 0$ ）射频脉冲构成。这里 $T_1 = 1000$ ms, $T_2 = 100$ ms, $T_E = 5$ ms, $T_R = 10$ ms。最后的回波为 $[-0.4756i, -0.4125i, -0.3358i]$ 。snapMRF 中相应的函数为 `epg_fisp()`。

第六个单元测试是翻转恢复平衡自由进动（IR-bSSFP）序列，使用 Bloch 方程来生成模拟信号。我们生成了一个含有 4 个元素的字典，字典中的每个元素有 4 个时间点。snapMRF 中相应的函数为 `roa_ssfp()`。

第七个单元测试是翻转恢复平衡自由进动（IR-bSSFP）序列，使用 EPG 模型来生成模拟信号。我们生成了一个含有 4 个元素的字典，字典中的每个元素有 4 个时间点。snapMRF 中相应的函数为 `epg_ssfp()`。

第八个单元测试是使用 EPG 生成的字典来进行自我匹配。我们从字典中随机选择 10 个体素作为指纹数据，并将其和字典进行匹配。对匹配的元素的验证确保了参数的正确性。snapMRF 中相应的函数为 `MRF_match()`。

5.4 数值实验数据与评价方法

我们进行了三组数值实验，用于比较 snapMRF 与其他 MRF 开源程序，如 EPG-X 和 PnP-MRF。数值实验所使用的数据以及评价方法如下。

5.4.1 体模数据

我们使用 32 通道的头部线圈对 MRI 系统体模^[185]（HPD, Boulder, Colorado）在 3.0 T（Philips Ingenia, Philips Healthcare, The Netherlands）进行冠状成像。体模在单个切片上成像，通过一系列不同 $MnCl_2$ 浓度的对比度球得到不同的 T_1 和 T_2 值。体模的温度为 20° 。k-space 数据使用 32 倍均匀螺旋下采样获得^[186]。图像的视野大小，平面内分辨率、平面外分辨率和螺旋采集时间为 $240\text{ mm} \times 240\text{ mm}$, $1\text{ mm} \times 1\text{ mm}$, 10 mm 和 4.9 ms 。原始指纹数据使用 BART^[187] 进行网格化和线圈组合重建，并且使用 BART 中的 eSPIRIT 估计线圈灵敏度图。样本密度使用^[188] 校正。重建后指纹数据的维度为 $1000 \times 240 \times 240$ 。体模数据所使用的 MRF 序列改变自将 uSSFP 用于 MRF 的第一篇文章^[126]。在这个序列中，绝热反转时间为 40 ms ，射频场为 sinc-gauss 型脉冲信号，并将时间带宽乘以 10 以最小化切片轮廓中 B_1^+ 的异质性。重复时间 T_R 和偏转角 α 是变化的，其中 $T_R \in [16, 18.7]$, $\alpha \in [0, 60]$ 。回波时间 T_E 固定在 4.65 ms ，总的扫描时间为 17.5 s 。

为了比较程序的运行时间，我们将 T_1 的值设置为 100 ms 到 3000 ms 之间， T_2 的值设置为 20 ms 到 2000 ms 之间。为了研究运行时间与字典大小的关系，我们选择了 10 组不同数目组合的 T_1 和 T_2 ，共形成 10 个大小不同的字典。我们将 snapMRF 的运行时间与 EPG-X^[167] 和 PnP-MRF^[189] 比较，并画出运行时间随字典大小变化曲线。

为了评估参数精度，我们比较了 snapMRF 与 EPG-X 生成 T_1 和 T_2 参数图的精度。在这组数值对比实验中，我们依然使用了体模数据。但是由于 EPG-X 只能处理固定 T_R 的序列，我们使用固定 T_R 的 MRF 序列对同样的体模数据进行成像。具体来说， T_R 固定在 16 ms，其他参数均与上一组实验相同。原始指纹数据的重建也与上一组实验相同，最终重建后指纹的数据为 $1500 \times 240 \times 240$ 。为了对比 snapMRF 与 EPG-X，我们首先比较二者在固定 T_R 序列上重建参数图的精度。其次，我们比较 snapMRF 在固定 T_R 和变化 T_R 序列上表现的差异。最后我们研究 B_1^+ 校正对参数图精度的影响。在所有比较中，我们均将字典大小选为 100,000 左右。不含 B_1^+ 校正时， T_1 的值从 50 ms 到 2500 ms，增量为 5 ms， T_2 的值从 5 ms 到 600 ms，增量为 2.5 ms，最终字典大小为 105,028。当考虑 B_1^+ 校正时， T_1 的值从 50 ms 到 2500 ms，增量为 5 ms， T_2 的值从 5 ms 到 600 ms，增量为 2.5 ms， B_1^+ 的值从 0.8 到 1.2，增量为 0.1，最终字典的大小为 105,655。参数图的精度使用相对误差来度量：

$$\text{err} = \frac{\|T_m - T_{gt}\|_2}{\|T_{gt}\|_2} \times 100\%, \quad (5.3)$$

其中 T_m 是由 snapMRF 或者 EPG-X 生成的参数图， T_{gt} 是真实值。

5.4.2 活体人脑数据

为了评估图像质量，在评估委员会同意以后，我们在一位志愿者知情的情况下，使用 32 通道头部线圈在 3.0 T (Philips Ingenia, Philips Healthcare, The Netherlands) 对其大脑横向切片成像。与体模数据相同，k-space 数据使用 32 倍均匀螺旋下采样获得^[186]。图像的视野大小，平面内分辨率、平面外分辨率和螺旋采集时间为 $240 \text{ mm} \times 240 \text{ mm}$, $1 \text{ mm} \times 1 \text{ mm}$, 5 mm 和 5.1 ms。原始 MRF 的重建算法和 MRF 的序列与体模数据均相同，最终形成指纹数据的维度为 $1000 \times 240 \times 240$ 。

在本组数值实验中，我们从视觉上比较 snapMRF 与 EPG-X 生成参数图的好坏。与体模数据类似，我们也比较四种情形重建的参数图，即固定 T_R EPG-X，固定 T_R snapMRF，变化 T_R snapMRF，变化 T_R snapMRF 加 B_1^+ 校正。在所有比较中，我们均将字典大小选为 100,000 左右。不含 B_1^+ 校正时， T_1 的值从 100 ms 到 4000 ms，增量为 10 ms， T_2 的值从 20 ms 到 2000 ms，增量为 5.5 ms，最终字典大小为 108,056。当考虑 B_1^+ 校正时， T_1 的值从 100 ms 到 4000 ms，增量为 20 ms， T_2 的值从 20 ms 到 2000 ms，增量为 14.5 ms， B_1^+ 的值从 0.8 到 1.2，增量为 0.1，最终字典的大小为 102,830。

所有程序均在 dual 10-core Intel Xeon E5-2630 2.20 GHz 的 CPU, 256 GB 内存和 Nvidia TITAN V 的 GPU 的工作站上运行。snapMRF 详细程序请参考网址<https://github.com/chixindebaoyu/snapMRF>，版本为 v0.0.1。

5.5 数值实验结果

5.5.1 程序运行时间

图5.2展示了 snapMRF 和 EPG-X（上）与 PnP-MRF（下）字典生成与模板匹配的运行时间的比较。注意图像中的曲线为对数尺度下的曲线。从图像中可以看出，所有时间曲线如设想一样随着字典大小呈线性变化。整体来看，snapMRF 的运行速度是 EPG-X 与 PnP-MRF 的 10–100×。具体来说，与 EGP-X 相比，snapMRF 在字典生成方面（上，实线）的速度提高了 60–1700×，在模板匹配（上，虚线）上速度提高了 2–20×。与 PnP-MRF 相比，snapMRF 在字典生成方面（下，实线）的速度提高了 2–100×，在模板匹配（下，虚线）上速度提高了 60–500×。因此，snapMRF 在整体上字典生成的速度提高了 10–1000×，而模板匹配的速度提高了 10–100×。注意当字典比较大时，snapMRF 的速度比其他两个程序快得多。

5.5.2 体模数据的重建参数图

snapMRF 和 EPG-X 字典生成和模板匹配的运行时间在表5.3中。可以看出，对于不同的序列，snapMRF 可以在 15 s 内生成大小为 100,000 的字典并进行模板匹配，速度远远大于 EPG-X。

表5.4和表5.5分别展示了 snapMRF 和 EPG-X 在体模数据重建生成的 T_1 和 T_2 与对应的相对误差。两个表中的第一列分别为 T_1 和 T_2 的真实值。从表5.4和表5.5中可以看出，对于固定 T_R 序列，snapMRF 与 EPG-X 的精度是几乎一样的。这是自然的，因为 snapMRF 与 EPG-X 都是用 EPG 模型生成字典，重建的方法也都使用了模板匹配。对于变化 T_R 序列， T_1 的误差从 5.0% 降低到了 2.6%，而 T_2 的误差保持不变。当我们考虑 B_1^+ 的影响时， T_1 的误差保持很低， T_2 的误差降低到 9.3%。结果表明变化 T_R 序列和 B_1^+ 校正有助于提高参数图的准确性。

图5.3显示了 snapMRF 和 EPG-X 生成的 T_1 ， T_2 和质子密度参数图。为了更好地观察感兴趣区域，我们用 Otsu 方法从质子密度图像中生成了一个掩膜，并应用到 T_1 和 T_2 参数图中。这样做可以去除背景中的高信号（研究对象之间用于填充的水），使得感兴趣区域凸现出来。另外，我们也简单地去除了外围的伪影。在绘制参数图时， T_1 图显

表 5.3 snapMRF 与 EPG-X 在字典生成和模板匹配上运行时间的比较

运行时间 (s)	EPG-X 固定 T_R	snapMRF 固定 T_R	snapMRF 变化 T_R	snapMRF 变化 $T_R+B_1^+$
体模/字典生成	17797.05	11.00	7.42	9.39
体模/模板匹配	137.13	5.97	4.14	4.88
脑部/字典生成	18629.82	11.29	7.63	8.72
脑部/模板匹配	143.55	6.13	4.23	4.63

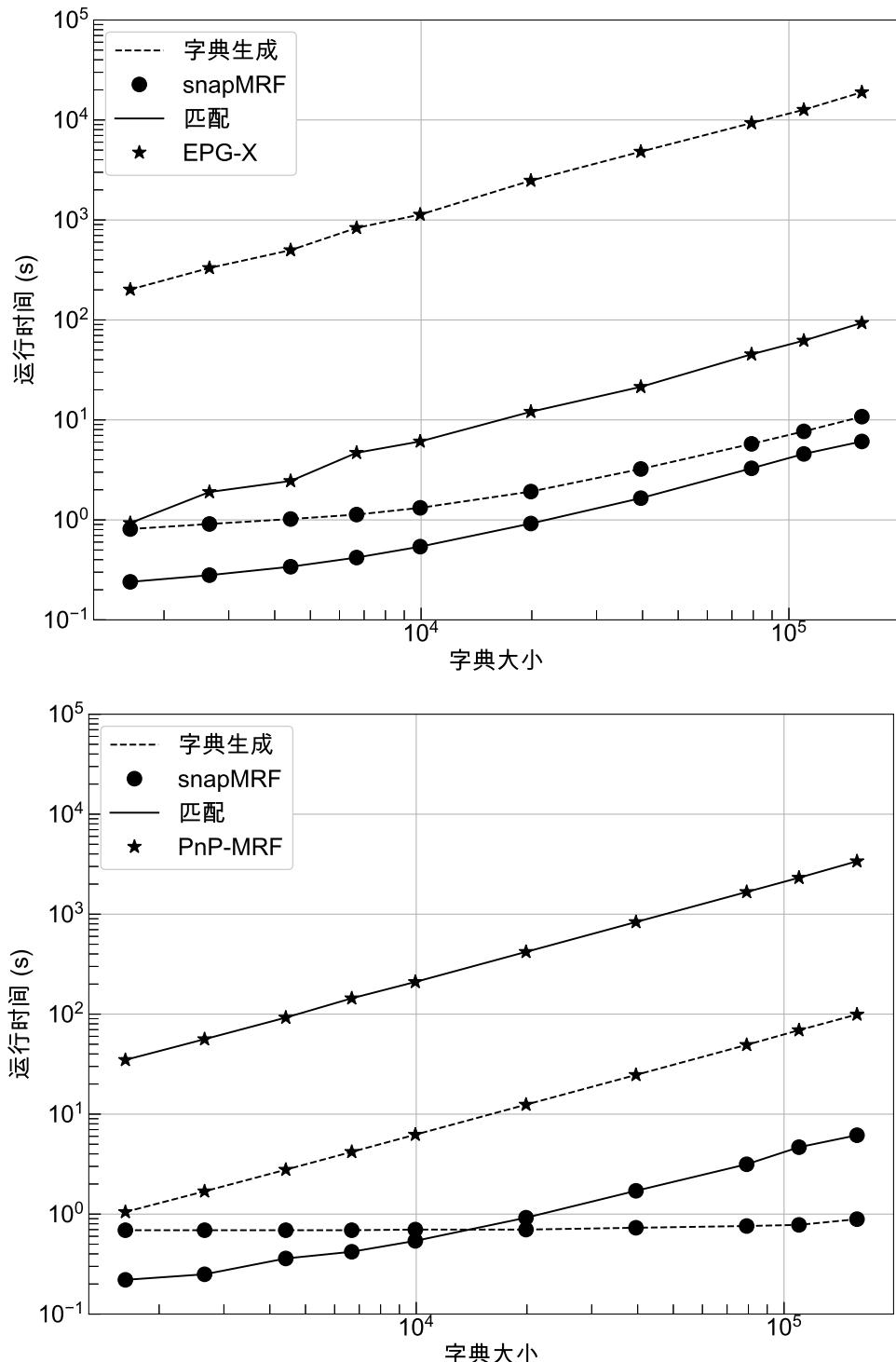


图 5.2 字典生成和模板匹配运行时间比较。（上）snapMRF 与 EPG-X；（下）snapMRF 与 PnP-MRF。注意曲线为对数尺度。运行时间随着字典大小线性变化，说明了有效的并行化。在我们的例子中，对于 240×240 个体素的指纹数据，模板匹配的时间比字典生成少得多。

表 5.4 snapMRF 与 EPG-X 在体模数据上 MRF 生成 T_1 参数图的准确性比较。表中的第一列为 T_1 的真实值。snapMRF 可以生成相当精确的 T_1 值，即使用了未经优化的序列和字典。

真实 T_1 (ms)	EPG-X 固定 T_R	snapMRF 固定 T_R	snapMRF 变化 T_R	snapMRF 变化 $T_R+B_1^+$
90.9	128.5	127.7	111.5	94.2
126.9	155.0	155.0	146.5	127.9
176.6	173.1	173.1	172.3	153.8
244.2	280.4	280.4	265.0	225.0
336.5	342.7	342.7	326.5	319.2
458.4	471.2	471.2	471.9	468.3
608.6	602.7	601.2	625.0	622.1
801.7	771.5	770.4	818.5	813.5
1044.0	945.0	943.1	1032.3	1026.0
1332.0	1262.7	1263.1	1310.8	1306.7
1604.0	1568.8	1568.1	1607.3	1593.3
1907.0	1861.2	1861.9	1854.2	1828.8
2173.0	2043.1	2043.1	2091.9	2094.2
2480.0	2366.5	2366.2	2434.6	2416.3
err (%)	4.9	5.0	2.6	3.0

表 5.5 snapMRF 与 EPG-X 在体模数据上 MRF 生成 T_2 参数图的准确性比较。表中的第一列为 T_2 的真实值。snapMRF 可以生成相当精确的 T_2 值，即使用了未经优化的序列和字典。

真实 T_2 (ms)	EPG-X 固定 T_R	snapMRF 固定 T_R	snapMRF 变化 T_R	snapMRF 变化 $T_R+B_1^+$
5.6	6.9	6.9	9.4	12.3
7.9	11.5	11.2	10.0	13.5
11.2	13.3	13.3	11.2	13.5
15.8	13.7	13.5	11.3	20.4
22.6	21.2	21.2	23.3	30.0
32.0	32.3	32.3	38.3	47.3
46.4	45.0	44.8	50.4	60.0
64.1	64.2	64.2	70.2	83.8
96.9	84.6	84.4	90.8	104.6
133.3	144.0	143.8	146.9	170.8
190.9	175.4	175.4	185.2	213.8
278.1	266.5	266.5	255.4	290.0
403.5	323.3	323.5	343.7	407.7
581.3	474.0	474.2	453.5	531.5
err (%)	16.9	16.9	17.9	9.3

示在 0–2500 ms 的范围, T_2 图显示在 0–800 ms 范围。从图5.3中可以看出, 对于固定 T_R 序列, snapMRF 和 EPG-X 所生成的参数图在视觉上几乎一致。对于变化 T_R 序列, T_1 参数图有所提高, 而 T_2 参数图保持不变。当考虑 B_1^+ 校正时, T_2 参数图也有了明显的提升。

5.5.3 活体人脑数据的重建参数图

图5.4显示了 snapMRF 在活体人脑数据中的表现以检测 snapMRF 是否可以生成干净的活体参数图。为了更好地视觉效果, 与体模数据相同, 我们使用 Otsu 方法生成了一个掩膜, 并将其作用在 T_1 和 T_2 参数图上。在绘制参数图时, T_2 参数图显示在 0–300 ms 的范围, 质子密度图像显示在 0–0.25 的范围。程序运行时间在表5.3中。

从图5.4中可以看出, 对于所有序列, snapMRF 都可以生成高质量的参数图。与体模数据的情况类似, snapMRF 和 EPG-X 在固定 T_R 序列时生成了视觉上一致的参数图。对于变化 T_R 序列, T_1 参数图更加平滑并且质子密度参数图也比固定 T_R 时更暗一些。当考虑 B_1^+ 时校正时, T_2 参数图也变得更平滑, 尤其是在灰质区域。

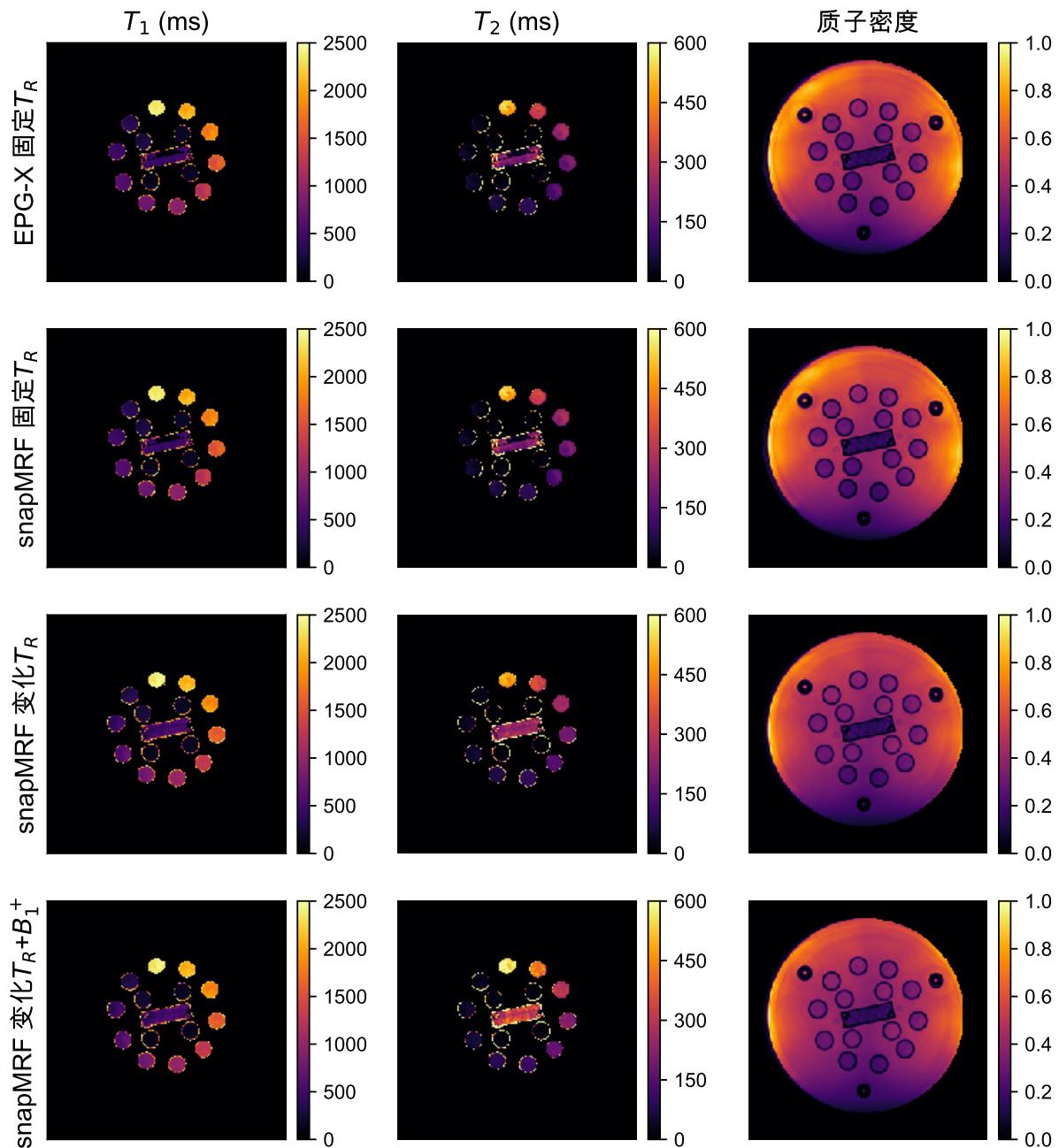


图 5.3 snapMRF 和 EPG-X 在体模数据上参数准确度的比较。第一行：EPG-X 使用固定 T_R 生成的参数图。从左到右分别为 T_1 , T_2 和质子密度。第二行到第四行为 snapMRF 的重建结果，使用的序列分别为固定 T_R 、变化 T_R 和考虑 B_1^+ 的变化 T_R 。 T_1 和 T_2 的定量估计在5.4和表5.5中。运行时间显示在表5.3中。在所有的 T_1 和 T_2 参数图中，为了更好的视觉效果，用于填充体模的背景中的水信号被人为的抑制。

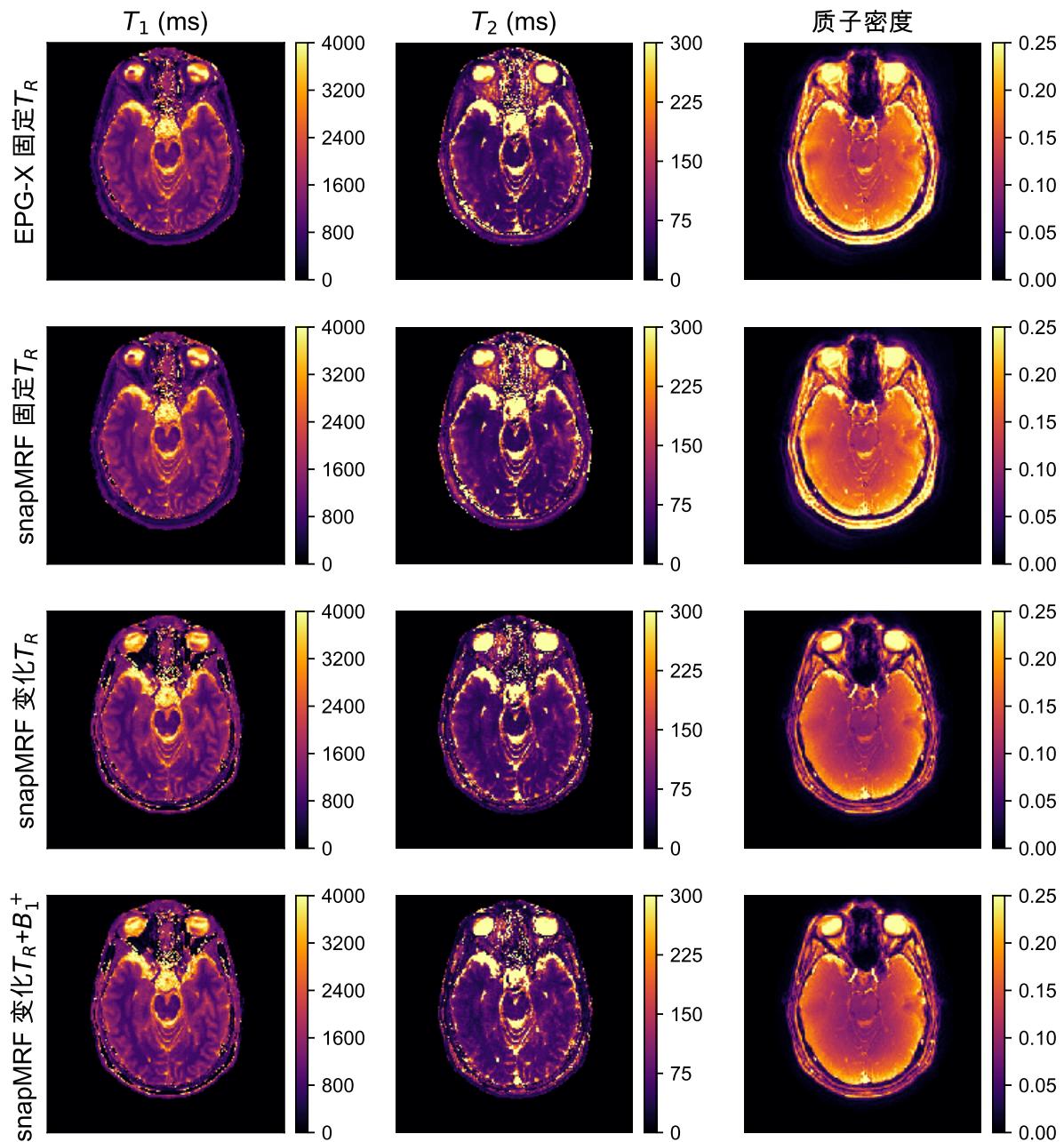


图 5.4 snapMRF 和 EPG-X 在活体人脑数据上参数准确度的比较。第一行：EPG-X 使用固定 T_R 生成的参数图。从左到右分别为 T_1 , T_2 和质子密度。第二行到第四行为 snapMRF 的重建结果，使用的序列分别为固定 T_R 、变化 T_R 和考虑 B_1^+ 的变化 T_R 。运行时间显示在表5.3中。所有的序列下，snapMRF 都可以得到质量较好的图像。

5.6 数值实验分析

在本文中，我们基于 MRF 字典生成与模板匹配运行时间慢的缺点，推出了一款新的基于 GPU 的程序 snapMRF，用于 MRF 字典生成与模板匹配。snapMRF 可以快速并准确地生成参数图，当字典较小时，可以达到实时效果。与其他开源的程序相比，字典生成的速度提高了 100–1000 倍，模板匹配的速度提高了 10–100 倍。snapMRF 的另一个优势是它适用于不同类型的 MRF 序列，包括在 MRF 中经常使用的变化 T_R 序列与 B_1^+ 校正。

对于基于数学模型和神经网络的 MRF 重建方法，利用 snapMRF 可以在不降低重建参数精度的情况下提高整体的计算效率。一方面数学模型通常需要预生成字典，而这一过程在 CPU 上十分耗时。snapMRF 可以实现 100–1000 倍的加速，因此提高了重建的效率。另一方面，snapMRF 可以达到神经网络的推断速度，并且运行时间与序列的选择无关。而且，如果研究人员仍然希望通过训练神经网络来进行参数图的重建，snapMRF 可以用来进行训练集，即字典的生成。另外，snapMRF 的快速度可以用来减少多个经历不同 B_1^+ 振幅的位置的计算时间。这适用于图像射频场的变化，也适用于切片轮廓^[190] 的变化。

snapMRF 中选择了 EPG 模型来进行字典的生成，主要原因除了 EPG 可以表征失相场这个特征之外，EPG 模型可以比较方便的加入其它因素的影响，比如化学位移 (chemical shift)、磁化转移 (magnetization transfer)^[191]、运动、扩散等。目前 snapMRF 可以实现的使用变化 T_R 序列对 B_1^+ 效应进行建模。之后我们会把前面提到因素加入到 snapMRF 中。

snapMRF 全部在单一 GPU 上运行，但是 snapMRF 在未来也可以被设置为在多个 GPU 上运行，如文献^[173] 中所用。理论上来说，snapMRF 的加速率会和 GPU 的个数成正比，因为在 snapMRF 中字典生成的并行计算是关于每个字典元素的，而模板匹配是关于每个指纹体素的。

5.7 本章小结

在本章中，我们首先介绍了磁共振指纹研究的背景意义和所存在的问题。针对 MRF 中字典生成与模板匹配速度慢的问题，我们开发了一款基于 GPU 的程序 snapMRF。我们给出了 snapMRF 的编程思路和主要流程，并简单介绍了核函数。我们将 snapMRF 与开源程序 EPG-X 和 PnP-MRF 进行了比较。运行时间上，snapMRF 整体提高了 MRF 生成参数图的速度，其中字典生成的速度提高了 100–1000 倍，模板匹配的速度提高了 10–100 倍。参数图精度上，由于可以处理多种 MRF 序列，snapMRF 可以生成更加精确地参数图。未来我们会在 snapMRF 中加入更多的功能，如切片轮廓等。

6 总结与展望

6.1 本文工作总结

核磁共振成像是现代医学最常用的成像手段，可以非侵入性地提供人体内组织的图像。提高核磁共振的成像速度一直是研究人员关心的问题，而利用压缩感知理论加速 MR 成像的方法已经被广泛地应用于临床和研究。在压缩感知的模型中，为动态 MR 图像选择合适的稀疏项是一个重要的问题，直接影响了重建图像的质量。核磁共振指纹是定量磁共振成像的新方法，可以在单次数据采集中同时获得多种组织参数。但是由于 MRF 的字典生成速度和模板匹配的速度一般十分耗时，对计算机的内存和存储都是巨大的负担。因此如何快速并精确地生成字典并进行字典匹配是一个亟待解决的问题。本文针对以上问题，研究的内容和创新点主要体现在以下几个方面：

1. 针对动态 MR 图像，利用压缩感知和图像分解的思想，提出了基于二阶时空 TGV 和核范数的重建模型。模型将图像分解为低秩部分和稀疏部分，其中低秩部分用核范数约束，稀疏部分用二阶时空 TGV 约束。核范数用来建模动态图像中时间方向高度相关的背景部分，并且可以很好地去除空间伪影；而 TGV 泛函用来表示图像中的光滑部分，可以在保证重建图像边界清晰的同时减少重建图像中的阶梯效应。文章利用 Primal-Dual 算法来求解模型，并给出了保证算法收敛的范数估计。为了减少算法的计算时间，文章也利用了图形处理单元来加速程序。文章针对体模、心脏灌注与胸部 DCE-MRI 图像，对比了四种最前沿的针对动态 MR 重建模型在不同采样模式和不同采样率下的表现。实验结果表明，相对于其他四个模型，我们提出的模型在不同采样模式和不同采样率下，可以更好地消除空间伪影并且保证图像边缘的清晰，对于胸部 DCE-MRI 图像的效果尤其显著。

2. 如何选择胸部 DCE-MRI 压缩感知模型中时间方向的稀疏项一直是未知的，为此我们针对胸部 DCE-MRI 图像，比较了五种不同的时间方向的稀疏项，并对重建结果进行了定量分析。这五种稀疏项分别为 Fourier 变换、Haar 小波变换、TV、二阶 TGV 和核范数。所有模型都使用 FISTA 算法进行求解。实验结果表明，核范数可以得到最高的信噪比，而 TV/二阶 TGV 可以得到最精确地定量分析。因此，对于胸部 DCE-MRI，选择 TV/二阶 TGV 作为稀疏项可以更好地重建病灶部分，而选择核范数则可以整体提高图像的信噪比。

3. 针对 MRF 中字典生成和模板匹配速度慢的缺点，我们利用图形处理单元进行 MRF 字典的生成和匹配，并开发了一款开源程序 snapMRF。程序基于 NVIDIA 公司的底层框架 CUDA，使用 C 语言编写。snapMRF 可以快速并准确地重建参数图，并且适用于不同的 MR 序列。相比于其他 MRF 开源程序，如 EPG-X（基于 CPU 的 MATLAB

语言编写) 和 PnP-MRF (基于 CPU 的 C 语言编写), snapMRF 的字典生成速度提高了 10–1000 倍, 字典匹配速度提高了 10–100 倍。对于较小的字典, 可以达到实时的效果。

6.2 下一步工作展望

1. 第三章中提出的基于二阶时空 TGV 和核范数的模型在动态 MR 图像的结果上优于已有的模型, 但由于模型的复杂性, 需要调节的参数较多。在数值实验中, 参数是通过穷举搜索的方式进行调节的, 耗费的时间较多。因此, 设计自适应调节参数的算法是我们今后的重点之一。
2. 第四章中所使用到的程序均用 MATLAB 编写, 并且在 CPU 上运行。因此程序运行速度较慢, 通常重建一例图像需要 10–20 分钟。可以考虑使用 GPU 来进行加速, 由于我们在第三章和第五章已经有了使用 GPU 编程的经验, 这一点应该是很容易实现的。
3. 第三章和第四章中的算法基本都是一阶的 FISTA 和 Primal-Dual 算法, 当模型复杂度高或者数据维度大时, 计算速度较慢。除了使用 GPU 加速, 也可以考虑设计二阶的算法, 提高模型的计算效率。
4. 第五章 snapMRF 程序可以处理射频场、梯度场的影响, 在以后的工作中, 我们会考虑加入其他因素的影响, 如磁化转移、切片轮廓等。
4. 本论文重点研究了压缩感知在动态 MR 图像重建中的应用, 没有考虑其他类型的医学图像, 如计算机断层扫描 (CT)、超声 (US)、正电子成像 (PET) 等。我们接下来的工作是将本文的模型应用到上述医学成像中。
5. 近几年, 深度学习在图像处理领域中变得十分火热, 并且以后研究将其应用在压缩感知和 MRF 的参数图重建中。所以如何利用深度神经网络来进行压缩感知的重建和 MRF 参数图的重建也是一个十分有趣的问题。

参考文献

- [1] 祖栋林. 核磁共振成像学 [M]. 北京: 高等教育出版社, 2004.
- [2] Yankelev T E, Gore J C. Dynamic contrast enhanced magnetic resonance imaging in oncology: Theory, data acquisition, analysis, and examples.[J]. *Curr Med Imaging Rev*, 2009, 3(2):91–107.
- [3] Ma D, Gulani V, Seiberlich N, et al. Magnetic resonance fingerprinting[J]. *Nature*, 2013, 495(7440):187–192.
- [4] Shannon C. Communication in the presence of noise[J]. *Proceedings of the IRE*, 1949, 86(2):10–21.
- [5] Candès E J, Tao T. Near-optimal signal recovery from random projections: Universal encoding strategies?[J]. *IEEE transactions on information theory*, 2006, 52(12):5406–5425.
- [6] Candès E J, et al. Compressive sampling[C]//*Proceedings of the International Congress of Mathematicians*. Madrid, Spain: EMS, 2006:1433–1452.
- [7] Candès E J, Romberg J. Quantitative robust uncertainty principles and optimally sparse decompositions[J]. *Foundations of Computational Mathematics*, 2006, 6(2):227–254.
- [8] Candès E J, Romberg J K, Tao T. Stable signal recovery from incomplete and inaccurate measurements[J]. *Communications on Pure and Applied Mathematics*, 2006, 59(8):1207–1223.
- [9] Donoho D L. Compressed sensing[J]. *IEEE Transactions on Information Theory*, 2006, 52(4):1289–1306.
- [10] Golbabae M, Vandergheynst P. Hyperspectral image compressed sensing via low-rank and joint-sparse matrix recovery[C]//*2012 IEEE International Conference on Acoustics, Speech and Signal Processing*. Kyoto, Japan: IEEE, 2012:2741–2744.
- [11] Quinsac C, Basarab A, Girault J M, et al. Compressed sensing of ultrasound images: Sampling of spatial and frequency domains[C]//*2010 IEEE Workshop On Signal Processing Systems*. San Francisco, CA, USA: IEEE, 2010:231–236.
- [12] Xu Q, Yu H, Mou X, et al. Low-dose X-ray CT reconstruction via dictionary learning[J]. *IEEE Transactions on Medical Imaging*, 2012, 31(9):1682–1697.
- [13] Lustig M, Santos J M, Donoho D L, et al. kt SPARSE: High frame rate dynamic

- MRI exploiting spatio-temporal sparsity[C]//Proceedings of the 13th Annual Meeting of ISMRM. Seattle, WA, USA: ISMRM, 2006:2420.
- [14] Wright J, Yang A Y, Ganesh A, et al. Robust face recognition via sparse representation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2008, 31(2):210–227.
- [15] Choi K, Wang J, Zhu L, et al. Compressed sensing based cone-beam computed tomography reconstruction with a first-order method[J]. Medical Physics, 2010, 37(9):5113–5125.
- [16] Needell D, Tropp J. CoSaMP: Iterative signal recovery from incomplete and inaccurate samples.[J]. Applied and Computational Harmonic Analysis, 2008, 26(3):301–321.
- [17] Tropp J A. Just relax: Convex programming methods for identifying sparse signals in noise[J]. IEEE Transactions on Information Theory, 2006, 52(3):1030–1051.
- [18] Gorodnitsky I F, Rao B D. Sparse signal reconstruction from limited data using FOCUSS: A re-weighted minimum norm algorithm[J]. IEEE Transactions on signal processing, 1997, 45(3):600–616.
- [19] Candès E J, Wakin M B, Boyd S P. Enhancing sparsity by reweighted l1 minimization[J]. Journal of Fourier Analysis and Applications, 2008, 14(5-6):877–905.
- [20] Chen S S, Donoho D L, Saunders M A. Atomic decomposition by basis pursuit[J]. SIAM Review, 2001, 43(1):129–159.
- [21] Chen S, Donoho D. Basis pursuit[C]//Proceedings of 1994 28th Asilomar Conference on Signals, Systems and Computers. Pacific Grove, CA, USA: IEEE, 1994:41–44.
- [22] Eldar Y C, Kutyniok G. Compressed sensing: theory and applications[M]. New York: Cambridge University Press, 2012.
- [23] H. Bauschke H, Combettes P. Convex analysis and monotone operator theory in Hilbert space[M]. New York: Springer, 2011.
- [24] Combettes P, R. Wajs V. Signal recovery by proximal forward-backward splitting[J]. Multiscale Modeling & Simulation, 2005, 4(4):1168–1200.
- [25] Goldstein T, Osher S. The split Bregman method for L1-regularized problems[J]. SIAM J. Imaging Sciences, 2009, 2(2):323–343.
- [26] Wang Y, Yang J, Yin W, et al. A new alternating minimization algorithm for total variation image reconstruction[J]. SIAM J. Imaging Sciences, 2008, 1(3):248–272.
- [27] Ramani S, Fessler J A. Parallel MR image reconstruction using augmented La-

- grangian methods[J]. IEEE Transactions on Medical Imaging, 2010, 30(3):694–706.
- [28] Boyd S, Parikh N, Chu E, et al. Distributed optimization and statistical learning via the alternating direction method of multipliers[J]. Foundations and Trends® in Machine Learning, 2011, 3(1):1–122.
- [29] Yang J, Zhang Y. Alternating direction algorithms for l1-problems in compressive sensing[J]. SIAM Journal on Scientific Computing, 2011, 33(1):250–278.
- [30] Combettes P, Pesquet J C. A Douglas–Rachford splitting approach to nonsmooth convex variational signal recovery[J]. IEEE Journal of Selected Topics in Signal Processing, 2008, 1(4):564–574.
- [31] Beck A, Teboulle M. A fast iterative shrinkage-thresholding algorithm for linear inverse problems[J]. SIAM J. Imaging Sciences, 2009, 2(1):183–202.
- [32] Chambolle A, Pock T. A first-order primal-dual algorithm for convex problems with applications to imaging[J]. Journal of Mathematical Imaging and Vision, 2011, 40(1):120–145.
- [33] Powell M J D. Restart procedures for the conjugate gradient method[J]. Mathematical Programming, 1977, 12(1):241–254.
- [34] Beck A, Teboulle M. Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems[J]. IEEE Transactions on Image Processing, 2009, 18(11):2419–2434.
- [35] Ji S, Ye J. An accelerated gradient method for trace norm minimization[C]//Proceedings of the 26th Annual International Conference on Machine Learning. Montreal, Quebec, Canada: ACM, 2009:457–464.
- [36] Knoll F, Bredies K, Pock T, et al. Second order total generalized variation (TGV) for MRI[J]. Magnetic Resonance in Medicine, 2011, 65(2):480–91.
- [37] Easley G, Labate D, Lim W Q. Sparse directional image representations using the discrete shearlet transform[J]. Applied and Computational Harmonic Analysis, 2008, 25(1):25–46.
- [38] Aharon M, Elad M, Bruckstein A. K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation[J]. IEEE Transactions on Signal Processing, 2006, 54(11):4311–4322.
- [39] P. Pruessmann K, Weiger M, Scheidegger M, et al. SENSE: Sensitivity encoding for fast MRI[J]. Magnetic Resonance in Medicine, 1999, 42(5):952–962.
- [40] Sodickson D, J. Manning W. Simultaneous acquisition of spatial harmonics (SMASH): Fast imaging with radiofrequency coil arrays[J]. Magnetic Resonance

- in Medicine, 1997, 38(4):591–603.
- [41] Griswold M, Jakob P, Heidemann R, et al. Generalized autocalibrating partially parallel acquisitions (GRAPPA)[J]. Magnetic Resonance in Medicine, 2002, 47(6):1202–1210.
- [42] Block K T, Uecker M, Frahm J. Undersampled radial MRI with multiple coils. Iterative image reconstruction using a total variation constraint[J]. Magnetic Resonance in Medicine, 2007, 57(6):1086–1098.
- [43] Liang D, Liu B, Wang J, et al. Accelerating SENSE using compressed sensing[J]. Magnetic Resonance in Medicine, 2009, 62(6):1574–1584.
- [44] Feng L, Grimm R, Tobias Block K, et al. Golden-angle radial sparse parallel MRI: Combination of compressed sensing, parallel imaging, and golden-angle radial sampling for fast and flexible dynamic volumetric MRI[J]. Magnetic Resonance in Medicine, 2014, 72(3):707–717.
- [45] Hong J, Kyunghyun S, Nayak K S, et al. k-t FOCUSS: A general compressed sensing framework for high resolution dynamic MRI.[J]. Magnetic Resonance in Medicine, 2010, 61(1):103–116.
- [46] Glover G H, Pauly J M. Projection reconstruction techniques for reduction of motion effects in MRI[J]. Magnetic Resonance in Medicine, 1992, 28(2):275–289.
- [47] Meyer C H, Hu B S, Nishimura D G, et al. Fast spiral coronary artery imaging[J]. Magnetic Resonance in Medicine, 1992, 28(2):202–213.
- [48] Gai N, Axel L. Correction of motion artifacts in linogram and projection reconstruction MRI using geometry and consistency constraints[J]. Medical Physics, 1996, 23(2):251–262.
- [49] Sajan Goud L, Yue H, Edward D B, et al. Accelerated dynamic MRI exploiting sparsity and low-rank structure: k-t SLR[J]. IEEE Transactions on Medical Imaging, 2011, 30(5):1042–1054.
- [50] Smith D, Li X, V Gambrell J, et al. Robustness of quantitative compressive sensing MRI: The effect of random undersampling patterns on derived parameters for DCE- and DSC-MRI[J]. IEEE Transactions on Medical Imaging, 2012, 31(2):504–11.
- [51] Holland D, Liu C, Song X, et al. Compressed sensing reconstruction improves sensitivity of variable density spiral fMRI[J]. Magnetic Resonance in Medicine, 2013, 70(6):1634–1643.
- [52] Kressler B, Spincemaille P, Nguyen T D, et al. Three-dimensional cine imaging using variable-density spiral trajectories and SSFP with application to coronary artery angiography[J]. Magnetic Resonance in Medicine, 2007, 58(3):535–543.

- [53] Karampinos D C, Van A T, Olivero W C, et al. High-resolution diffusion tensor imaging of the human pons with a reduced field-of-view, multishot, variable-density, spiral acquisition at 3T[J]. *Magnetic Resonance in Medicine*, 2009, 62(4):1007–1016.
- [54] Santos J M, Cunningham C H, Lustig M, et al. Single breath-hold whole-heart MRA using variable-density spirals at 3T[J]. *Magnetic Resonance in Medicine*, 2006, 55(2):371–379.
- [55] Ravishankar S, Ye J C, Fessler J A. Image reconstruction: From sparsity to data-adaptive methods and machine learning[J]. arXiv, 2019:1904.02816.
- [56] Ye J C. Compressed sensing MRI: a review from signal processing perspective[J]. *BMC Biomedical Engineering*, 2019, 1(8):1–17.
- [57] Ma S, Yin W, Zhang Y, et al. An efficient algorithm for compressed MR imaging using total variation and wavelets[C]//Proceedings IEEE Conference on Computer Vision Pattern Recognition. Anchorage, AK, USA: IEEE, 2008:1–8.
- [58] Yang J, Yin Z, Yin W. A fast alternating direction method for TVL1-L2 signal reconstruction from partial Fourier data[J]. *IEEE Journal of Selected Topics in Signal Processing*, 2010, 4(2):288–297.
- [59] Huang J, Zhang S, Metaxas D. Efficient MR image reconstruction for compressed MR imaging[J]. *Medical Image Analysis*, 2011, 15(5):670–679.
- [60] Bredies K, Kunisch K, Pock T. Total generalized variation[J]. *SIAM Journal on Imaging Sciences*, 2010, 3(3):492–526.
- [61] Schloegl M, Holler M, Schwarzl A, et al. Infimal convolution of total generalized variation functionals for dynamic MRI[J]. *Magnetic Resonance in Medicine*, 2017, 78(1):142–155.
- [62] Stanković R S, Falkowski B J. The Haar wavelet transform: Its status and achievements[J]. *Computers & Electrical Engineering*, 2003, 29(1):25–44.
- [63] Candès E J, Donoho D L. Continuous curvelet transform: I. Resolution of the wavefront set[J]. *Applied and Computational Harmonic Analysis*, 2005, 19(2):162–197.
- [64] Do M N, Vetterli M. The contourlet transform: An efficient directional multiresolution image representation[J]. *IEEE Transactions on Image Processing*, 2005, 14(12):2091–2106.
- [65] Qu X, Cao X, Guo D, et al. Combined sparsifying transforms for compressed sensing MRI[J]. *Electronics Letters*, 2010, 46(2):121–123.
- [66] Qu X, Zhang W, Guo D, et al. Iterative thresholding compressed sensing MRI

- based on contourlet transform[J]. *Inverse Problems in Science and Engineering*, 2010, 18(6):737–758.
- [67] Chen C, Huang J. Compressive sensing MRI with wavelet tree sparsity[C]//*Advances in Neural Information Processing Systems*. Nevada: NIPS, 2012:1115–1123.
- [68] Chen C, Huang J. Exploiting the wavelet structure in compressed sensing MRI[J]. *Magnetic Resonance Imaging*, 2014, 32(10):1377–1389.
- [69] Peyré G, Bougleux S, Cohen L. Non-local regularization of inverse problems[C]//*European Conference on Computer Vision*. Marseille, France: Springer, 2008:57–68.
- [70] Lou Y, Zhang X, Osher S, et al. Image recovery via nonlocal operators[J]. *Journal of Scientific Computing*, 2010, 42(2):185–197.
- [71] Gilboa G, Osher S. Nonlocal operators with applications to image processing[J]. *Multiscale Modeling & Simulation*, 2008, 7(3):1005–1028.
- [72] Akçakaya M, Basha T A, Goddu B, et al. Low-dimensional-structure self-learning and thresholding: Regularization beyond compressed sensing for MRI reconstruction[J]. *Magnetic Resonance in Medicine*, 2011, 66(3):756–767.
- [73] Liang D, Wang H, Chang Y, et al. Sensitivity encoding reconstruction with non-local total variation regularization[J]. *Magnetic Resonance in Medicine*, 2011, 65(5):1384–1392.
- [74] Liu S, Cao J, Liu H, et al. MRI reconstruction using a joint constraint in patch-based total variational framework[J]. *Journal of Visual Communication and Image Representation*, 2017, 46:150–164.
- [75] Qu X, Hou Y, Lam F, et al. Magnetic resonance image reconstruction from undersampled measurements using a patch-based nonlocal operator[J]. *Medical Image Analysis*, 2014, 18(6):843–856.
- [76] Dabov K, Foi A, Katkovnik V, et al. Image denoising by sparse 3-D transform-domain collaborative filtering[J]. *IEEE Transactions on Image Processing*, 2007, 16(8):2080–2095.
- [77] Eksioglu E M. Decoupled algorithm for MRI reconstruction using nonlocal block matching model: BM3D-MRI[J]. *Journal of Mathematical Imaging and Vision*, 2016, 56(3):430–440.
- [78] Gleichman S, Eldar Y C. Blind compressed sensing[J]. *IEEE Transactions on Information Theory*, 2011, 57(10):6958–6975.

- [79] Saiprasad R, Yoram B. MR image reconstruction from highly undersampled k-space data by dictionary learning[J]. IEEE Transactions on Medical Imaging, 2011, 30(5):1028–1041.
- [80] Mairal J, Bach F, Ponce J, et al. Online learning for matrix factorization and sparse coding[J]. Journal of Machine Learning Research, 2010, 11(1):19–60.
- [81] Rubinstein R, Zibulevsky M, Elad M. Double sparsity: Learning sparse dictionaries for sparse signal approximation[J]. IEEE Transactions on Signal Processing, 2009, 58(3):1553–1564.
- [82] Ravishankar S, Nadakuditi R R, Fessler J A. Efficient sum of outer products dictionary learning (SOUP-DIL) and its application to inverse problems[J]. IEEE Transactions on Computational Imaging, 2017, 3(4):694–709.
- [83] Bruckstein A M, Donoho D L, Elad M. From sparse solutions of systems of equations to sparse modeling of signals and images[J]. SIAM Review, 2009, 51(1):34–81.
- [84] Ravishankar S, Bresler Y. Efficient blind compressed sensing using sparsifying transforms with convergence guarantees and application to magnetic resonance imaging[J]. SIAM Journal on Imaging Sciences, 2015, 8(4):2519–2557.
- [85] Ravishankar S, Bresler Y. Data-driven learning of a union of sparsifying transforms model for blind compressed sensing[J]. IEEE Transactions on Computational Imaging, 2016, 2(3):294–309.
- [86] Wen B, Ravishankar S, Bresler Y. FRIST—flipping and rotation invariant sparsifying transform learning and applications[J]. Inverse Problems, 2017, 33(7):074007.
- [87] Wen B, Li Y, Bresler Y. The power of complementary regularizers: Image recovery via transform learning and low-rank modeling[J]. arXiv, 2018:1808.01316.
- [88] Wen B, Ravishankar S, Pfister L, et al. Transform learning for magnetic resonance image reconstruction: From model-based learning to building neural networks[J]. arXiv, 2019:1903.11431.
- [89] LeCun Y, Bengio Y, Hinton G. Deep learning[J]. Nature, 2015, 521(7553):436–444.
- [90] Schmidhuber J. Deep learning in neural networks: An overview[J]. Neural Networks, 2015, 61:85–117.
- [91] Jin K H, McCann M T, Froustey E, et al. Deep convolutional neural network for inverse problems in imaging[J]. IEEE Transactions on Image Processing, 2017, 26(9):4509–4522.
- [92] Yang Y, Sun J, Li H, et al. ADMM-Net: A deep learning approach for compressive sensing MRI[J]. arXiv, 2017:1705.06869.

- [93] Wang S, Su Z, Ying L, et al. Accelerating magnetic resonance imaging via deep learning[C]//2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI). Prague, Czech Republic: IEEE, 2016:514–517.
- [94] Lee D, Yoo J, Tak S, et al. Deep residual learning for accelerated MRI using magnitude and phase networks[J]. IEEE Transactions on Biomedical Engineering, 2018, 65(9):1985–1995.
- [95] Han Y, Yoo J, Kim H H, et al. Deep learning with domain adaptation for accelerated projection-reconstruction MR[J]. Magnetic Resonance in Medicine, 2018, 80(3):1189–1205.
- [96] Zhu B, Liu J Z, Cauley S F, et al. Image reconstruction by domain-transform manifold learning[J]. Nature, 2018, 555(7697):487–492.
- [97] Hyun C, Kim H, Min Lee S, et al. Deep learning for undersampled MRI reconstruction[J]. Physics in Medicine and Biology, 2017, 63(13):135007.
- [98] Hammernik K, Klatzer T, Kobler E, et al. Learning a variational network for reconstruction of accelerated MRI data[J]. Magnetic Resonance in Medicine, 2017, 79(6):3055–3071.
- [99] Quan T M, Nguyen-Duc T, Jeong W K. Compressed sensing MRI reconstruction using a generative adversarial network with a cyclic loss[J]. IEEE Transactions on Medical Imaging, 2018, 37(6):1488–1497.
- [100] Gregor K, LeCun Y. Learning fast approximations of sparse coding[C]//Proceedings of the 27th International Conference on International Conference on Machine Learning. Haifa, Israel: Omnipress, 2010:399–406.
- [101] Adler J, Öktem O. Learned primal-dual reconstruction[J]. IEEE Transactions on Medical Imaging, 2018, 37(6):1322–1332.
- [102] Gupta H, Jin K H, Nguyen H Q, et al. CNN-based projected gradient descent for consistent CT image reconstruction[J]. IEEE Transactions on Medical Imaging, 2018, 37(6):1440–1453.
- [103] Han Y, Ye J C. k-Space Deep learning for accelerated MRI[J]. arXiv, 2018:1805.03779.
- [104] Lee J, Han Y, Ryu J K, et al. k-Space deep learning for reference-free EPI ghost correction[J]. Magnetic Resonance in Medicine, 2019, 82(6):2299–2313.
- [105] Jeffrey T, Peter B, Pruessmann K P. k-t BLAST and k-t SENSE: Dynamic MRI with high frame rate exploiting spatiotemporal correlations.[J]. Magnetic Resonance in Medicine, 2003, 50(5):1031–1042.

- [106] Jung H, Ye J C, Kim E Y. Improved k-t BLAST and k-t SENSE using FOCUSS[J]. Physics in Medicine and Biology, 2007, 52(11):3201–3226.
- [107] Candès E J, Recht B. Exact matrix completion via convex optimization[J]. Foundations of Computational Mathematics, 2009, 9(6):717–772.
- [108] Zhao B, Haldar J P, Christodoulou A G, et al. Image reconstruction from highly undersampled (k, t)-space data with joint partial separability and sparsity constraints[J]. IEEE Transactions on Medical Imaging, 2012, 31(9):1809–1820.
- [109] Ricardo O, Emmanuel C, Sodickson D K. Low-rank plus sparse matrix decomposition for accelerated dynamic MRI with separation of background and dynamic components[J]. Magnetic Resonance in Medicine, 2015, 73(3):1125–1136.
- [110] Candès E J, Li X, Ma Y, et al. Robust principal component analysis?[J]. Journal of the ACM, 2011, 58(3):1–37.
- [111] Trémoulhéac B, Dikaios N, Atkinson D, et al. Dynamic MR image reconstruction–separation from undersampled (k,t)-space via low-rank plus sparse prior[J]. IEEE Transactions on Medical Imaging, 2014, 33(8):1689–1701.
- [112] Shin P J, Larson P E, Ohliger M A, et al. Calibrationless parallel imaging reconstruction based on structured low-rank matrix completion[J]. Magnetic Resonance in Medicine, 2014, 72(4):959–970.
- [113] Haldar J P. Low-rank modeling of local k -space neighborhoods (LORAKS) for constrained MRI[J]. IEEE Transactions on Medical Imaging, 2013, 33(3):668–681.
- [114] Jin K H, Lee D, Ye J C. A general framework for compressed sensing and parallel MRI using annihilating filter based low-rank Hankel matrix[J]. IEEE Transactions on Computational Imaging, 2016, 2(4):480–495.
- [115] Lee D, Jin K H, Kim E Y, et al. Acceleration of MR parameter mapping using annihilating filter-based low rank Hankel matrix (ALOHA)[J]. Magnetic Resonance in Medicine, 2016, 76(6):1848–1864.
- [116] Ongie G, Jacob M. Off-the-grid recovery of piecewise constant images from few Fourier samples[J]. SIAM Journal on Imaging Sciences, 2016, 9(3):1004–1041.
- [117] Lingala S G, Jacob M. Blind compressive sensing dynamic MRI[J]. IEEE Transactions on Medical Imaging, 2013, 32(6):1132–1145.
- [118] Wang Y, Ying L. Compressed sensing dynamic cardiac cine MRI using learned spatiotemporal dictionary[J]. IEEE Transactions on Biomedical Engineering, 2013, 61(4):1109–1120.
- [119] Caballero J, Price A N, Rueckert D, et al. Dictionary learning and time sparsity

- for dynamic MR data reconstruction[J]. IEEE Transactions on Medical Imaging, 2014, 33(4):979–994.
- [120] Schlemper J, Caballero J, Hajnal J V, et al. A deep cascade of convolutional neural networks for dynamic MR image reconstruction[J]. IEEE Transactions on Medical Imaging, 2017, 37(2):491–503.
- [121] Qin C, Schlemper J, Caballero J, et al. Convolutional recurrent neural networks for dynamic MR image reconstruction[J]. IEEE Transactions on Medical Imaging, 2018, 38(1):280–290.
- [122] Wang S, Ke Z, Cheng H, et al. Dimension: Dynamic MR imaging with both k-space and spatial prior knowledge obtained via multi-supervised network training[J]. arXiv, 2018:1810.00302.
- [123] Ke Z, Wang S, Cheng H, et al. CRDN: Cascaded residual dense networks for dynamic MR imaging with edge-enhanced loss constraint[J]. arXiv, 2019:1901.06111.
- [124] Panda A, B. Mehta B, Coppo S, et al. Magnetic resonance fingerprinting - An overview[J]. Current Opinion in Biomedical Engineering, 2017, 3:56–66.
- [125] D Elster A. Gradient-echo MR imaging: Techniques and acronyms[J]. Radiology, 1993, 186(1):1–8.
- [126] Jiang Y, Ma D, Seiberlich N, et al. MR fingerprinting using fast imaging with steady state precession (FISP) with spiral readout[J]. Magnetic Resonance in Medicine, 2015, 74(6):1621–1631.
- [127] Rieger B, Zimmer F, Zapp J, et al. Magnetic resonance fingerprinting using echo-planar imaging: Joint quantification of T1 and T2* relaxation times[J]. Magnetic Resonance in Medicine, 2016, 78(5):1724–1733.
- [128] Jiang Y, Ma D, Jerecic R, et al. MR fingerprinting using the quick echo splitting NMR imaging technique[J]. Magnetic Resonance in Medicine, 2016, 77(3):979–988.
- [129] Weigel M. Extended phase graphs: Dephasing, RF pulses, and echoes—pure and simple[J]. Journal of Magnetic Resonance Imaging, 2015, 41(2):266–295.
- [130] Cauley S F, Setsompop K, Ma D, et al. Fast group matching for MR fingerprinting reconstruction[J]. Magnetic Resonance in Medicine, 2015, 74(2):523–528.
- [131] Yang M, Ma D, Jiang Y, et al. Low rank approximation methods for MR fingerprinting with large scale dictionaries: Large scale low rank MR fingerprinting[J]. Magnetic Resonance in Medicine, 2017, 79(4):2392–2400.
- [132] Davies M, Puy G, Vandergheynst P, et al. A compressed sensing framework for magnetic resonance fingerprinting[J]. SIAM Journal on Imaging Sciences, 2014,

7(4):2623–2656.

- [133] Cline C C, Chen X, Mailhe B, et al. AIR-MRF: Accelerated iterative reconstruction for magnetic resonance fingerprinting[J]. Magnetic Resonance Imaging, 2017, 41:29–40.
- [134] Pierre E Y, Ma D, Chen Y, et al. Multiscale reconstruction for MR fingerprinting[J]. Magnetic Resonance in Medicine, 2016, 75(6):2481–2492.
- [135] Wang Z, Li H, Zhang Q, et al. Magnetic resonance fingerprinting with compressed sensing and distance metric learning[J]. Neurocomputing, 2016, 174:560–570.
- [136] Zhao B, Setsompop K, Ye H, et al. Maximum likelihood reconstruction for magnetic resonance fingerprinting[J]. IEEE Transactions on Medical Imaging, 2016, 35(8):1812–1823.
- [137] Cohen O, Zhu B, Rosen M S. MR fingerprinting deep reconstruction network (DRONE)[J]. Magnetic Resonance in Medicine, 2018, 80(3):885–894.
- [138] Hoppe E, Körzdörfer G, Würfl T, et al. Deep learning for magnetic resonance fingerprinting: A new approach for predicting quantitative parameter values from time series.[J]. Studies in Health Technology and Informatics, 2017, 243:202–206.
- [139] Balsiger F, Konar A S, Chikop S, et al. Magnetic resonance fingerprinting reconstruction via spatiotemporal convolutional neural networks[C]//International Workshop on Machine Learning for Medical Image Reconstruction. Granada, Spain: Springer, 2018:39–46.
- [140] 乔梁, 涂光忠. NMR 核磁共振 [M]. 北京: 化学工业出版社, 2009.
- [141] Haacke E M, Brown R W, Thompson M R, et al. Magnetic resonance imaging: physical principles and sequence design[M]. New York: Wiley-Liss, 1999.
- [142] Sprawls P. Magnetic resonance imaging: principles, methods, and techniques[M]. Wisconsin, USA: Medical Physics Publishing, 2000.
- [143] Bloch F. Nuclear induction[J]. Physical Review, 1946, 70(7-8):460–474.
- [144] Haidekker M A. Medical imaging technology[M]. New York, USA: Springer, 2013.
- [145] van Vaals J J, Brummer M E, Dixon W T, et al. Keyhole method for accelerating imaging of contrast agent uptake[J]. Journal of Magnetic Resonance Imaging, 1993, 3(4):671–675.
- [146] Candès E J, Romberg J, Tao T. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information[J]. IEEE Transactions on Information Theory, 2006, 52(2):489–509.
- [147] Zhao B, P. Haldar J, Christodoulou A, et al. Image reconstruction from highly

- undersampled (k , t)-space data with joint partial separability and sparsity constraints[J]. *IEEE Transactions on Medical Imaging*, 2012, 31(9):1809–20.
- [148] Gamper U, Boesiger P, Kozerke S. Compressed sensing in dynamic MRI[J]. *Magnetic Resonance in Medicine*, 2008, 59(2):365–73.
- [149] Qiu C, Lu W, Vaswani N. Real-time dynamic MR image reconstruction using Kalman filtered compressed sensing[C]//ICASSP. Washington, DC, USA: IEEE Computer Society, 2009:393–396.
- [150] Han S, Paulsen J, Zhu G, et al. Temporal/spatial resolution improvement of in vivo DCE-MRI with compressed sensing-optimized FLASH[J]. *Magnetic Resonance Imaging*, 2012, 30(6):741–52.
- [151] Lustig M, Donoho D L, Santos J M, et al. Compressed sensing MRI[J]. *IEEE Signal Processing Magazine*, 2008, 25(2):72–82.
- [152] Segars W P, Tsui B M. Study of the efficacy of respiratory gating in myocardial SPECT using the new 4-D NCAT phantom[J]. *IEEE Transactions on Nuclear Science*, 2002, 49(3):675–679.
- [153] Sharif B, Bresler Y. Adaptive real-time cardiac MRI using PARADISE: Validation by the physiologically improved NCAT phantom[C]//IEEE International Symposium on Biomedical Imaging: From Nano to Macro. Arlington, TX, USA: IEEE, 2007:1020–1023.
- [154] Li X, Abramson R G, Arlinghaus L R, et al. Multiparametric magnetic resonance imaging for predicting pathological response after the first cycle of neoadjuvant chemotherapy in breast cancer.[J]. *Investigative Radiology*, 2015, 50(4):195–204.
- [155] Wang Z, Bovik A C, Sheikh H R, et al. Image quality assessment: from error visibility to structural similarity[J]. *IEEE Transactions on Image Processing*, 2004, 13(4):600–612.
- [156] Smith D, Li X, Abramson R, et al. Potential of compressed sensing in quantitative MR imaging of cancer[J]. *Cancer Imaging*, 2013, 13(4):633–44.
- [157] Chen L, Schabel M, Dibella E. Reconstruction of dynamic contrast enhanced magnetic resonance imaging of the breast with temporal constraints[J]. *Magnetic Resonance Imaging*, 2010, 28(5):637–45.
- [158] Ji J, Lang T. Dynamic MRI with compressed sensing imaging using temporal correlations[C]//IEEE International Symposium on Biomedical Imaging: From Nano to Macro. Paris, France: IEEE, 2008:1613–1616.
- [159] Smith D, Welch E, Li X, et al. Quantitative effects of using compressed sensing in dynamic contrast enhanced MRI[J]. *Physics in Medicine and Biology*, 2011,

- 56(15):4933–46.
- [160] Lustig M, Donoho D, Pauly J M. Sparse MRI: The application of compressed sensing for rapid MR imaging[J]. *Magnetic Resonance in Medicine*, 2010, 58(6):1182–1195.
- [161] I. Rudin L, Osher S, Fatemi E. Nonlinear total variation based noise removal algorithms[J]. *Physica D: Nonlinear Phenomena*, 1992, 60(1-4):259–268.
- [162] Li X, Abramson R, Arlinghaus L, et al. Multiparametric magnetic resonance imaging for predicting pathological response after the first cycle of neoadjuvant chemotherapy in breast cancer[J]. *Investigative Radiology*, 2014, 50(4):195–204.
- [163] Tofts P, Brix G, L. Buckley D, et al. Estimating kinetic parameters from dynamic contrast-enhanced T1- weighted MRI of a diffusible tracer: Standardized quantities and symbols[J]. *Journal of Magnetic Resonance Imaging*, 1999, 10(3):223–32.
- [164] Smith D, Li X, Arlinghaus L, et al. DCEMRI.jl: A fast, validated, open source toolkit for dynamic contrast enhanced MRI analysis[J]. *PeerJ*, 2015, 3(909):1–19.
- [165] European Society of Radiology (ESR). Magnetic resonance fingerprinting—A promising new approach to obtain standardized imaging biomarkers from MRI[J]. *Insights into Imaging*, 2015, 6(2):163–165.
- [166] Bipin Mehta B, Coppo S, McGivney D, et al. Magnetic resonance fingerprinting: A technical review[J]. *Magnetic Resonance in Medicine*, 2019, 81(1):25–46.
- [167] Malik S J, Teixeira R P A, Hajnal J V. Extended phase graph formalism for systems with magnetization transfer and exchange: EPG-X: Extended phase graphs with exchange[J]. *Magnetic Resonance in Medicine*, 2018, 80(2):767–779.
- [168] Després P, Jia X. A review of GPU-based medical image reconstruction[J]. *Physica Medica*, 2017, 42:76–92.
- [169] Biguri A, Dosanjh M, Hancock S, et al. TIGRE: A MATLAB-GPU toolbox for CBCT image reconstruction[J]. *Biomedical Physics & Engineering Express*, 2016, 2(5):055010.
- [170] Palenstijn W, Batenburg K, Sijbers J. Performance improvements for iterative electron tomography reconstruction using graphics processing units (GPUs)[J]. *Journal of Structural Biology*, 2011, 176(2):250–253.
- [171] Knoll F, Unger M, Diwoky C, et al. Fast reduction of undersampling artifacts in radial MR angiography with 3D total variation on graphics hardware[J]. *Magnetic Resonance Materials in Physics, Biology and Medicine*, 2010, 23(2):103–114.
- [172] Sørensen T S, Schaeffter T, Noe K Ø, et al. Accelerating the nonequispaced fast

- Fourier transform on commodity graphics hardware[J]. IEEE Transactions in Medical Imaging, 2008, 27(4):538–547.
- [173] Smith D S, Sengupta S, Smith S A, et al. Trajectory optimized NUFFT: Faster non-Cartesian MRI reconstruction through prior knowledge and parallel architectures[J]. Magnetic Resonance in Medicine, 2019, 81(3):2064–2071.
- [174] Eklund A, Dufort P, Villani M, et al. BROCCOLI: Software for fast fMRI analysis on many-core CPUs and GPUs[J]. Frontiers in Neuroinformatics, 2014, 8(24):1–19.
- [175] Xanthis C G, Venetis I E, Chalkias A V, et al. MRISIMUL: a GPU-based parallel approach to MRI simulations[J]. IEEE Transactions on Medical Imaging, 2014, 33(3):607–617.
- [176] Jung Woo C, Amin N, Omer O, et al. GPU-based real-time volumetric ultrasound image reconstruction for a ring array[J]. IEEE Transactions on Medical Imaging, 2013, 32(7):1258–1264.
- [177] Yakang D, Jie T, Di D, et al. Real-time visualized freehand 3D ultrasound reconstruction based on GPU[J]. IEEE Transactions on Information Technology in Biomedicine, 2010, 14(6):1338–1345.
- [178] Garcia M P, Bert J, Benoit D, et al. Accelerated GPU based SPECT Monte Carlo simulations[J]. Physics in Medicine and Biology, 2016, 61(11):4001–4018.
- [179] Cui J, Pratx G, Meng B, et al. Distributed MLEM: An iterative tomographic image reconstruction algorithm for distributed memory architectures[J]. IEEE Transactions on Medical Imaging, 2013, 32(5):957–67.
- [180] Vasilevsky N A, Brush M H, Paddock H, et al. On the reproducibility of science: unique identification of research resources in the biomedical literature[J]. PeerJ, 2013, 1(148):1–22.
- [181] Collins F S, Tabak L A. Policy: NIH plans to enhance reproducibility[J]. Nature, 2014, 505(7485):612–613.
- [182] Begley C G, Ioannidis J P. Reproducibility in science[J]. Circulation Research, 2015, 116(1):116–126.
- [183] Smith, David S. RawArray file format reference implementation[OL]. <http://github.com/davidssmith/ra>. [Online; accessed 12-May-2018].
- [184] Hargreaves, B. Bloch equation simulation[OL]. <http://mrsrl.stanford.edu/~brian/bloch>. [Online; accessed 25-January-2019].
- [185] Keenan K E, Stupic K F, Boss M A, et al. Comparison of T1 measurement using ISMRM/NIST system phantom[C]//Proceedings of ISMRM 24th Annual Meeting.

Singapore: ISMRM, 2016:3209.

- [186] Pipe J G, Zwart N R. Spiral trajectory design: A flexible numerical algorithm and base analytical equations[J]. Magnetic Resonance in Medicine, 2014, 71(1):278–285.
- [187] Uecker M, Ong F, Tamir J I, et al. Berkeley advanced reconstruction toolbox[C]//Proceedings of ISMRM 23rd Annual Meeting. Toronto, Ontario, Canada: ISMRM, 2015:2486.
- [188] Zwart N R, Johnson K O, Pipe J G. Efficient sample density estimation by combining gridding and an optimized kernel[J]. Magnetic Resonance in Medicine, 2012, 67(3):701–710.
- [189] Cloos M A, Knoll F, Zhao T, et al. Multiparametric imaging with heterogeneous radiofrequency fields[J]. Nature communications, 2016, 7(12445):1–10.
- [190] Ma D, Coppo S, Chen Y, et al. Slice profile and B1 corrections in 2D magnetic resonance fingerprinting[J]. Magnetic Resonance in Medicine, 2017, 78(5):1781–1789.
- [191] Hamilton J I, Griswold M A, Seiberlich N. MR Fingerprinting with chemical exchange (MRF-X) to quantify subvoxel T1 and extracellular volume fraction[J]. Journal of Cardiovascular Magnetic Resonance, 2015, 17(35):1–3.

致 谢

值此论文完成之际，谨在此向多年来给予我关心和帮助的老师、同学、朋友和家人表示衷心的感谢！

首先，我要感谢我的导师杨孝平教授，感谢杨老师在我读博期间给予我的无私帮助和关怀。是他把我领入了图像处理中的数学问题这一领域，本论文的选题、攻关到最终完成，都凝聚着老师的心血，浸透着老师的汗水，他孜孜不倦的工作热情是我学习的榜样。杨老师工作繁忙，但每周都会安排讨论班，不仅拓展了我们的研究广度，还创造了良好的学术氛围。杨老师不仅知识渊博，高瞻远瞩，治学严谨，而且平易近人，品德谦逊，让我受益匪浅，终生难忘，从他身上我学会了许多做学问和为人处世的道理，这将激励我不断追求更高的目标。这里再次向杨老师表示衷心的感谢，师生情谊，难忘终生。

感谢美国范德比尔特大学的 Akram Aldroubi 教授和 David Smith 教授对我在美国访学期间的悉心指导和关怀，这一年的经历增加了我对科研的热爱，同时感谢组里的唐穗博士和 Amernak 博士对我在学习和生活上的帮助，帮助我解决了许多问题，也提供了很多富有建设性的想法和讨论。

感谢教研室和我一起学习的同学们：李旭、贵鹿颖、高益铭、钱春俊、谷亚楠、聂梓伟、王钦霞、张梦璐、李渊强和吴宇雳等陪我度过了快乐的博士生涯，给我留下了美好的回忆，使我在紧张的学习之余有了丰富的课余生活，也感谢他们在学习和生活上给予的帮助。

我要深深地感谢我的父母和家人，感谢他们这么多年的理解与支持，感谢他们一直默默地关心我、帮助我，正是他们始终如一的奉献，才使我顺利地完成了学业。在今后的科研工作中，我必会更加努力，来回报你们对我的爱。

最后，再次感谢所有帮助过我的老师和朋友。

附录

攻读博士学位期间已发表的论文：

- [1] Dong Wang, David S. Smith and Xiaoping Yang. Dynamic MR Image Reconstruction using TGV and Low-rank Decomposition. Magnetic Resonance in Medicine (SCI 二区). 2019. doi: 10.1002/mrm.28064.
- [2] Dong Wang, Jason Ostenson, and David S. Smith. snapMRF: GPU-Accelerated Magnetic Resonance Fingerprinting Dictionary Generation and Matching using Extended Phase Graphs. Magnetic Resonance Imaging (SCI 三区). 2019. doi: 10.1016/j.mri.2019.11.015.
- [3] Dong Wang, Lori R. Arlinghaus, Thomas E. Yankeelov, Xiaoping Yang, and David S. Smith. Quantitative Evaluation of Temporal Regularizers in Compressed Sensing Dynamic Contrast Enhanced MRI of the Breast. International Journal of Biomedical Imaging (EI). 2017. doi: 10.1155/2017/7835749.

攻读博士学位期间参加的科学研究情况：

- [1] 主持了南京理工大优秀博士研究培养基金（2013-2015）
- [2] 参与了国家自然科学基金重大研究计划培育项目“高阶非线性偏微分方程图像模型及其基础算法”(项目编号：91330101, 2014.01-2016.12, 杨孝平教授主持)