

分 类 号 TP391
收藏编号
学校代码 10386



密级 公开
学号 181120083
编号

福州大学

学术型硕士研究生学位（毕业）论文

复杂环境下基于稀疏编码的人脸识别方法

研究

学 科 专 业: 数字媒体技术
研 究 方 向: 网络与移动媒体技术
研 究 生 姓 名: 周筝
指 导 教 师 、 职 称: 陈国栋 副教授
协 助 导 师 、 职 称:
所 在 学 院: 物理与信息工程学院
答 辩 委 员 会 主 席 : 吴怡 教授

二〇二一年六月

一 遵守学术行为规范承诺

本人已熟知并愿意自觉遵守《福州大学研究生和导师学术行为规范暂行规定》和《福州大学关于加强研究生毕业与学位论文质量管理的规定》的所有内容，承诺所提交的毕业和学位论文是终稿，不存在学术造假或学术不端行为，且论文的纸质版与电子版内容完全一致。

二 独创性声明

本人声明所提交的论文是我个人在导师指导下进行的研究工作及取得的研究成果。尽我所知，除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含为获得福州大学或其他教育机构的学位或证书而使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示了谢意。本人完全意识到本声明的法律结果由本人承担。

三 关于论文使用授权的说明

本人完全了解福州大学有关保留使用学位论文的规定，即：学校有权保留递交论文的复印件，允许论文被查阅和借阅；学校可以公布论文的全部或部分内容，可以采用影印、缩印或其他复制手段保存论文。（保密的论文在解密后应遵守此规定）

本学位论文属于（必须在以下相应方框内打“√”，否则一律按“非保密论文”处理）：

1、保密论文：本学位论文属于保密，在_____年解密后适用本授权书。

2、非保密论文：本学位论文不属于保密范围，适用本授权书。

研究生本人签名：_____ 签字日期：20 年 月 日

研究生导师签名：_____ 签字日期：20 年 月 日

复杂环境下基于稀疏编码的人脸识别方法研究

中文摘要

人脸识别技术是计算机视觉领域最重要的技术之一，近年来，从理论创新到应用扩展，人脸识别技术取得了巨大的成功并得到了广泛关注。其中基于深度学习的方法被证明能实现较高的识别率，但仍然需要消耗较多计算和存储资源。稀疏表达人脸识别是一种基于压缩感知理论实现高效特征提取以及分类的方法，其具有扎实的数学理论基础，但由于其实际的计算以及时间消耗，许多基于表示的识别任务仍然受到稀疏求解器的限制。本文重点研究一种基于数据驱动的稀疏表达人脸识别方法，通过结合传统稀疏编码方法和深度学习思想实现快速的人脸稀疏表达以及人脸识别。主要研究内容包括以下几个方面：

(1) 针对当前人脸稀疏求解器速度过慢的缺点，本文重点考虑成本效益的框架模型，提出采用数据驱动的人脸稀疏表达方法，以实现给定迭代次数下尽可能低的误差。文章将数据驱动的稀疏编码方法扩展到人脸识别领域，得到理论支撑的同时能有效的加快人脸稀疏表达速度。

(2) 本文提出一种将迭代软阈值算法 (Iterative Shrinkage Thresholding Algorithm,ISTA) 展开成计算图，学习稀疏字典并自动调整步长来加速人脸稀疏表达的策略。这种策略使 ISTA 可以通过反向传播进行训练，这样不仅可以使稀疏字典适合训练数据集，而且可以提高收敛速度。为了提高其有效性使其更加适应实时环境，本文提出采用学习步长代替调整步长策略。步长改进的方法加快了稀疏表达速度，其在图像信号重建仿真和低维特征输入的人脸识别实验上较其他同类型算法优势较为突出。

(3) 本文提出构建轻量级的稀疏表达人脸识别模型，在充分调研和验证可训练稀疏编码的权重特性的前提下，文章提出采用一种既存的 tied 权重设计来重新建模稀疏表达人脸识别模型，该模型得到了有效简化并实现了计算和资源减负。图像信号重建和低维特征输入的人脸识别实验都验证了该模型的有效性。

(4) 本文首先实验性的提出当前基于动能补偿的稀疏求解器存在动能“不足”需要优化补偿的猜想，其次假设存在与数据相关的“最佳”动能，最后提出一种学习动量方案来实现“最佳”增益，以加速人脸稀疏编码器的求解。所提学习动能方案仅仅需要增加一个与索引无关的学习标量，但能同时实现高性能以及低模型复杂度，图像信号重建仿真以及人脸稀疏编码、解码工作证实了上述假设，并实现了更优的稀疏表达能力。

关键词：稀疏编码；人脸识别；LISTA 算法；步长；动能

Research on Sparse Coding-based Face Recognition under Complex Environment

Abstract

Face recognition is one of the most important technologies in computer vision. Extending from theoretical innovations to application expansions, in more recent years, face recognition has achieved great success and received widespread attention. On the one hand, deep learning-based methods have been proven to achieve a higher recognition rate, but more consumption of computation and resources are needed. On the other hand, sparse representation-based face recognition is an effective compressed sensing-based feature extraction and classification method, supported by solid mathematical theoretical foundations. However, many representation-based recognition tasks are still limited by sparse solvers due to their effective computational and time cost. This paper focuses on study a data-driven sparse representation-based face recognition method, by binding sparse coding and deep learning to achieve faster sparse representation and face recognition. The main works are listed as follows:

(1). Taking into account the slow speed of the existed face sparse solvers, this paper focuses on the cost-effective framework and proposes a data-driven sparse representation-based face recognition scheme to produce as low as possible classification error in a given number of iterations. This article extends data-driven sparse coding methods to the field of face recognition, which can accelerate the speed of sparse representation of faces effectively and is supported by theoretical analysis.

(2). In order to accelerate the speed of face representations, we propose a strategy of sparse dictionary learning together with automatic tuning of the step size in the iterative soft thresholding algorithm (ISTA) via its unfolding to a computation graph. Our strategy makes ISTA trainable by back-propagation, so as not only to fit the sparse dictionary to a training dataset but also to improve the speed of convergence. In order to make it practical in large-scale applications, we suggest learning the step sizes instead of tuning them. Our work accelerated the speed of sparse representation of faces and showed its fast speed of reconstructing images and signals as well as representing and recognizing faces compare to other same type algorithms.

(3). This paper proposes to build a lightweight sparse representation-based face recognition model. After a variety of investigations and verifications of the weights

structure designs of trainable sparse coding, the article proposes to use an existing tied weight design to remodel the sparse representation-based face recognition model. This idea can help to simplify the model effectively and reduce both the calculation and resource consumptions. The model has been verified by reconstruction simulation of images and signals as well as in real face recognition experiments which take low-dimension features as inputs.

(4). This paper experimentally conjectures that the current acceleration method-based sparse solver has "insufficient" momentum and needs to be compensated via optimization. Following the assumption of the existence of a data-dependent "optimum" in the acceleration technique implemented, we introduce a learned momentum scheme to achieve "optimal" gain. The proposed scheme only adds an additional parameter but helps to achieve both "good" performances and "low" consumption of computational resources. Extensive synthetic reconstruction experiments and face-based encoding and decoding results confirmed our hypothesis and showed greater sparse representations.

Key words: **Sparse coding; Face recognition; LISTA; Step size; Momentum**

目 录

中文摘要.....	I
Abstract.....	II
主要符号表.....	VII
第一章 绪论	1
1.1 研究背景和意义.....	1
1.2 国内外的研究与发展现状.....	2
1.3 论文主要工作.....	4
1.4 论文结构安排.....	5
第二章 稀疏表达人脸识别理论介绍及模型构建	7
2.1 稀疏表达人脸识别的理论介绍.....	7
2.1.1 人脸稀疏表达求解.....	9
2.1.1.1 传统的稀疏编码算法.....	9
2.1.1.2 可训练的稀疏编码算法.....	10
2.2 稀疏表达人脸识别的模型架构.....	10
2.2.1 稀疏表达人脸识别模型 SRC	11
2.2.2 可训练稀疏表达的人脸识别模型构建.....	12
2.3 本章小结.....	14
第三章 步长改进的稀疏表达人脸识别方法研究	15
3.1 引言	15
3.2 步长改进的稀疏编码模型.....	15
3.2.1 调整步长模型.....	16
3.2.2 学习步长模型.....	16
3.2.3 算法框架.....	17
3.2.4 仿真实验.....	18
3.2.4.1 实验设置.....	18
3.2.4.2 仿真结果分析.....	18
3.3 步长改进的稀疏表达人脸识别.....	21
3.3.1 模型求解.....	21
3.3.2 实验设置.....	22
3.3.2.1 数据集介绍.....	22
3.3.2.2 识别环境.....	24
3.3.2.3 其他设置.....	24
3.3.3 结果分析.....	24

3.3.3.1 光照条件下的人脸识别.....	25
3.3.3.2 遮挡及其他条件下的人脸识别.....	26
3.4 本章小结.....	27
第四章 稀疏表达人脸识别模型简化	29
4.1 引言	29
4.2 稀疏编码的权重结构讨论.....	29
4.2.1 基于学习迭代软阈值的权重结构.....	30
4.2.2 权重结构验证.....	30
4.2.3 绑定的权重结构及其模型.....	32
4.2.4 仿真实验.....	33
4.3 绑定的权重与其人脸识别.....	34
4.3.1 模型求解.....	34
4.3.2 结果分析.....	35
4.3.2.1 光照条件下的人脸识别.....	35
4.3.2.2 遮挡及其他条件下的人脸识别.....	37
4.4 本章小结.....	38
第五章 动能补偿的稀疏表达人脸识别	39
5.1 引言	39
5.2 稀疏编码与动能补偿.....	39
5.2.1 可训练迭代算法和 Nesterov 加速度	41
5.2.1.1 算法验证.....	41
5.2.2 可训练迭代算法与最佳动能补偿模型.....	42
5.2.2.1 学习标量 γ 的验证	43
5.2.3 算法参数比较.....	44
5.2.4 仿真实验.....	45
5.3 动能补偿的稀疏表达人脸识别.....	47
5.3.1 模型求解.....	47
5.3.2 结果分析.....	48
5.3.2.1 光照条件下的人脸识别.....	48
5.3.2.2 遮挡及其他条件下的人脸识别.....	50
5.4 本章小结.....	51
总结与展望.....	52
1 论文主要工作总结.....	52
2 本文创新点.....	53
3 未来工作展望.....	53

参考文献.....	55
致 谢.....	60
个人简历.....	61
在学期间的研究成果及发表的学术论文	62

主要符号表

CNN	Convolutional Neural Network	卷积神经网络
RNN	Recurrent Neural Network	循环神经网络
ResNets	Residual Networks	残差网络
CSIS	Center for Strategic and International Studies	战略与国际研究中心
LASSO	Least Absolute Shrinkage and Selection Operator	套索算法
SRC	Sparse Representation for Classification	稀疏表达分类方法
CoD	Coordinate Descent Algorithm	坐标下降算法
AMP	Approximate Message Passing	近似消息传递
LAMP	Learned AMP	学习的 AMP
ADMM	Alternating Direction Method of Multipliers	基于乘数的交替方向方法
LSTM	Long Short Term Memory	长短期记忆模型
SVD	Singular Value Decomposition	奇异值分解
SNR	Signal-to-Noise Ratio	信噪比
ISTA	Iterative Shrinkage Thresholding Algorithm	迭代软阈值算法
FISTA	Fast Iterative Shrinkage Thresholding Algorithm	快速迭代软阈值算法
LISTA	Learned ISTA	学习的迭代软阈值算法
LISTA-CP	LISTA with partial weight coupling	LISTA 的部分权重耦合算法
LFISTA	Learned FISTA	学习的 FISTA
TiLISTA	Tied LISTA	绑定的 LISTA
ALISTA	Analytic LISTA	分析的 LISTA
GLISTA	Gated LISTA	增益的 LISTA
TsLISTA	Tuned steps LISTA	调整步长的 LISTA
LsLISTA	Learned steps LISTA	学习步长的 LISTA
LM-FISTA	Learned Momentum FISTA	学习动量的 FISTA

-t	tied	绑定的权重结构
-ss	support selection	支持选择
-NA	Nesterov Acceleration	Nesterov 加速度
LM-	Learned Momentum	学习的动能

第一章 绪论

1.1 研究背景和意义

人脸识别是当前计算机视觉和生物识别领域被研究最多的主题之一，其经过多年的发展，已在理论及实践上具备了充分可行性。广义的人脸识别包括构建人脸识别系统的一系列相关技术，包括人脸图像采集、人脸定位、人脸识别预处理、身份确认以及身份查找等；而狭义的人脸识别特指通过人脸进行身份确认或者身份查找的技术或系统。人脸识别属于一种生物特征识别技术，其在识别深度信息、整体信息以及动态信息上具有更大优势而区别于其他身体特征识别技术，如：虹膜识别、手势识别以及声音识别技术。虹膜识别采用局部信息并对摄像头与眼睛的距离要求很高，手势识别基于指纹信息，在深度信息上采样不足容易出现伪造等情况，声音识别在动态信息上无法体现其动态性，容易造成安全信息泄漏等问题。与以上生物识别技术相比，人脸识别具有识别过程友好、操作简易、可多人同时识别等优点，目前主要应用于安防、健康管理、银行以及零售销售等领域。

基于人工设计的特征和传统机器学习技术的传统方法近来已被使用非常大型的数据集训练的深度神经网络取代，人脸识别技术快速发展，识别准确率也不断提高。但人脸识别技术依旧存在技术难点，首先其对实时性要求比较严格；其次在多种复杂环境包括强弱光、脸部妆容以及遮挡等情况下，识别能力会受到极大限制；最后，当前基于机器学习的方法由于其输入图像较大以及网络模型较为复杂，计算负担以及资源消耗负担相对传统方法比较繁重。

从学术角度来看，人脸识别方法的研究不仅能促进计算机视觉学科发展给其带来新的视野，而且也会对图像信号处理领域有一定促进作用。从工业角度看，人脸识别方法的研究会促进工业水平的发展，如今人脸识别技术广泛应用于社会的各行各业，包括安保、执法、工地、工作考勤等，针对这些不同的应用场景，人脸识别的检测也存在不同的或并发的问题难点。实时性、训练模型的复杂程度以及图像的预处理以及计算资源消耗等方面，现有的人脸识别系统会因为这些问题难点导致性能的急剧下降，极大地限制了人脸识别的应用范围，亟待科研工作者去解决。更进一步，结合当前的国内外的实际环境，新冠疫情下遮挡的人脸识别场景十分频繁，对实时性要求更高，其具有较强的应用价值。综上所述，人脸识别方法的研究具有较强的理论及实际意义。

1.2 国内外的研究与发展现状

人脸识别按照其需求及功能不同可以分为三个部分：（1）人脸检测，即对人脸的位置信息的检测，其检测方法主要包括深度学习^[1-3]和非深度学习^[4,5]两个部分，评价标准以召回率、误检数以及检测速率为主；（2）关键点检测：通过检测到的人脸对其进行定位关键点，2013年，Sun等人^[6]首次将CNN应用到人脸关键点检测，随后更多的学者将基于深度学习的方法应用到关键点检测上，比如Face++和研究^[7]等。关键点是人脸形状的稀疏表示，其在人脸跟踪，美颜等场景中应用较多；（3）人脸识别：主要是完成身份比对（如著名的Deepface^[8]和FaceNet^[9]可以确定两张被拍照的面孔是否属于同一个人）以及识别任务，识别任务是本文的研究重点，以下详述基于识别任务的人脸识别的国内外研究现状。

文章首先针对基于识别任务的人脸识别进行综述。CNN是用于面部识别的最常见类型的深度学习方法。深度学习方法^[10-12]的主要优点是可以使用大量数据对其进行训练，多层人工神经网络模型有很强的特征学习能力，深度学习模型学习得到的特征数据对原数据有更本质的代表性，这将大大便于分类和可视化问题。基于深度学习的人脸识别最初受限于网络“深度”，网络复杂度过高，深度学习网络则会出现网络退化现象，而此问题通过ResNets^[13]的引入得到了解决。研究^[14]于2014年提出了GaussianFace，获得了98.52%的面部识别准确率，但其所需的存储容量和计算时间方面消耗较多。2015年，在LFW^[15]数据集上，FaceNet的识别准确度达到了99.63%。2018年报告^[16]详述了来自45位研究者研究的127种算法的识别准确性，评估的主要结果是，从2013-2018年识别准确率得到大幅提高，人脸识别的能力大大加强，甚至当图像具有明显的老化或不合格的质量时，识别依旧能获得成功，算法容忍度得到有效提高。准确性的提高源于完全采用基于深度卷积神经网络的方法或其集成对先前方法的替代，深度学习的人脸识别取得了巨大成功。2019年该项研究还继续提供了其他人脸识别的结果^[17]。在CSIS的2020年测试中^[18]，最佳面部识别算法的错误率仅为0.08%。报告^[19]指出在最近的六年中，人脸识别的准确率在六年内提高了50倍，人工神经网络算法正在帮助人脸识别算法变得更加准确。然而人工智能算法还存在一些其他漏洞（如RealAI团队通对抗样本攻破手机人脸识别解锁^[20]）等。

Wright等人提出基于稀疏表达来提高鲁棒性的分类算法SRC^[21]并应用于人脸识别，该研究采用数学的线性表达的方法进行建模求解最小化问题得到稀疏解实现图像重构，通过计算重构图像和输入图像的最小误差来进行分类。SRC在特征提取以及提高鲁棒性方面取得了巨大成功，吸引了大量学者开始将稀疏表达用于图像识别方面的研究。基于稀疏表达的人脸识别的研究重点主要包括字典学习

以及 l_1 范数稀疏编码的求解。字典学习方法^[22-25]通过构建出较好的基向量从而更加有效的表达信号，但其往往需要较多的迭代次数才能达到期望的结果，学习额外的补偿字典依旧需要借助稀疏编码算法来进行求解稀疏解，此种形式对于存储以及计算资源方面要求都更高，资源耗损较多。 l_1 范数稀疏编码的求解主要包括传统的稀疏编码算法^[26-31]以及可训练的稀疏编码算法，可训练的稀疏编码是将传统迭代算法展开作为深度神经网络来加速求解稀疏的一种手段，近十年来，基于传统迭代算法展开的做法已经成为图像信号处理新的重要研究内容^[32-35]。加速 l_1 范数稀疏编码的求解是本文的研究重点，以下详述可训练的稀疏编码算法国内外进展。

文章其次对可训练的稀疏编码算法进行综述。Gregor 和 LeCun^[36]首先提出将 ISTA^[26]和 CoD^[37]展开成深度神经网络来加速求解稀疏，进而引入了一种学习的算法，其中基于 ISTA 的为 LISTA。遵循这种思想，近年来引入了许多这种算法的变体算法^[32-35,38-40]。Sprechmann 等人^[32]提出了一种构造前馈网络的方法，以加快稀疏编码的近似速度。研究^[33]构造可学习的追踪过程架构，该架构是从近端下降算法的迭代中得出的。Hershey 等人^[39]将迭代展开为类似于神经网络的分层结构。LISTA 的想法启发了 AMP 算法^[41]的延伸及其贝叶斯扩展^[42]。Metzler 等人^[40]研究了类似的基于学习的模型，该模型也是受到 AMP 的启发。当涉及到 l_∞ 范数最小化问题，基于 ADMM 算法，Wang 等人^[43]将原始凸最小化问题表示为前馈神经网络。Chen 等人^[44]提出了一种非共享权重的 LISTA-CP，而 Liu 等人^[45]通过一个给定的字典预算算一个“好的”权重矩阵，有效的减少了训练参数的数量，并被证明与原来的 LISTA 算法具有相同的性能。Moreau 等人^[46]提出了一种 LFISTA 算法，原始的 FISTA 中添加的“动能”项是一种对历史信息的计算，即最后两次迭代的输出之差，来实现快速逼近稀疏解的逐元素的计算的方法。在类 LISTA 的算法中，一些算法^[47-49]也是采用这种思路来实现，通过对历史信息的回顾来刺激或补偿稀疏解在估计中比预期更低的成分。研究^[47]作用于阈值算子。研究^[48]介绍了一种自适应的“动能”策略来保证参数的更新以及对历史信息的封装，该研究还是第一个将稀疏编码算法和 LSTM^[50]结合到一起形成一个新的循环神经网络模型的研究，使用 LSTM 单元而不是简单的循环神经网络单元，因此被证明更加有助于历史信息的捕获并能加快收敛速度。研究^[49]提出一种“门”策略，分别作用于软阈值收缩算子运算前的稀疏解以及运算之后的稀疏解，使其能够进一步逼近“真实”的稀疏解。此种“门”策略需要添加相关的新的计算步骤，也会增加相应的训练参数，其中在阈值算子之后添加的“门”策略是根据时间和输入变化的，但是由于额外的方程以及计算的加入，这种策略下的稀疏编码会变得更加复杂以及需要更多资源消耗。其他主要的研究内容还包括对理

论的完善^[46,44,49,51-52]以及实际应用的扩展^[53-58]等。

文章最后对以上综述内容进行总结。机器学习的方法性能较为突出，但通常图像空间输入的维度较大，相对特征提取计算消耗较大，构建的网络模型也因为会有较多的滤波器而产生较大的计算量，存储计算资源消耗较多同时对硬件的要求也较高。**SRC** 人脸识别算法基于压缩感知原理通过随机矩阵得到随机特征从而实现从图像空间到特征空间的转换，相比于一些传统的特征提取方法其从速率和效率上都得到了极大的提升，相比于机器学习中通过卷积提取特征的方法，其也能大大的减少计算复杂性以及实现高效的降维处理，但其仍需较多次的迭代才能得到有效的稀疏表达，额外的字典学习使得计算以及资源消耗相对较多。可训练的稀疏编码方法结合深度学习优异的学习能力与稀疏编码高效的表达能力能有效解决稀疏表达过慢的问题，但现阶段的算法存在不适用于无监督学习，模型框架一般较为复杂、资源消耗较多等局限性，且编码速度仍然存在提升空间。

1.3 论文主要工作

尽管当前人脸识别的研究与文献颇多，基于深度学习的人脸识别在百万级别的数据集如 **MegaFace**^[59] 等都已经达到相当高的水准，但是在真实应用场景中依旧存在对识别环境敏感且实时性不强，对图像空间以及大量滤波器的计算、资源消耗较大等问题，其仍具有较大的研究空间；基于稀疏编码的人脸识别具有高效的特征提取方式并能产生较低维的特征空间作为输入，但在最小化问题的求解上通常需要大量多次的迭代才能达到收敛，字典学习的方法也需要对字典进行复杂的构建和另外的资源消耗。本文在充分研究和分析现有算法的基础上，针对提升复杂环境中的人脸识别的实时性和识别能力以及对计算资源的减负，分别从以下几点阐述论文的主要工作：

(1) 针对复杂环境下的稀疏表达人脸识别研究，本文提出采用可训练稀疏编码算法替代传统稀疏编码算法对 l_1 最小化问题进行求解。文章将一般的稀疏表达人脸识别转化为数据驱动的稀疏表达人脸识别方法并构建其基于机器学习的识别模型，通过结合稀疏编码扎实的理论基础、高效的特征提取方式以及深度学习的快速的学习能力，能够在有效求解稀疏表达减少资源消耗的同时具备一定的理论保障。

(2) 针对快速求解人脸稀疏表达部分，本文提出采用调整和学习迭代算法的步长的策略。合适的步长对快速求解人脸识别的 l_1 最小化问题有重要影响，步长太大会导致结果不收敛、步长太小则收敛速度太慢。本文提出在训练中学习字典并同步自动化通过 SVD 计算调整步长，但 SVD 计算会产生巨大的计算与时间

开销，因此本文进一步提出采用学习步长的方法。改进步长的策略在较低维的特征空间作为输入的前提下，仅仅几次迭代就能获得较高的稀疏表达能力。

(3) 针对优化稀疏表达人脸识别的模型部分，本文提出结合改进的步长的策略和一种既存的 tied 权重结构^[64]构建更加轻便的人脸识别模型，它能帮助简化计算以及资源负担。为了验证这种权重结构的有效性，文章更进一步设计了一种权重模型证实了它的可行性，确保了本文所提人脸识别模型的合理性。仿真实验和人脸识别实验结果都表明本文添加的这种权重设计能够简化模型并更具竞争力。

(4) 为了实现更进一步加速人脸稀疏表达的求解及性能优化，本文提出一种学习最优动能的稀疏编码技术。在假设现有算法得到的稀疏表达结果不够充分且存在一种数据驱动的最优动能的前提下，文章充分调查并实验性验证了该假设的正确性。已经存在一些研究^[46,48,49]提出将“动量”项添加到基于学习的算法中的基本思想，但是，这些研究在所提出的某些设定下，添加动量项通常会导致可训练参数的数量大量增加以及模型复杂化。与这些研究不同，本文的贡献是研究动量的“最优”标量，仅仅只需要添加一个标量训练参数，却有助于同时实现“好的”性能和“低”的计算资源消耗。

1.4 论文结构安排

本文具体组织结构如图 1-1 所示，各章节研究内容如下：

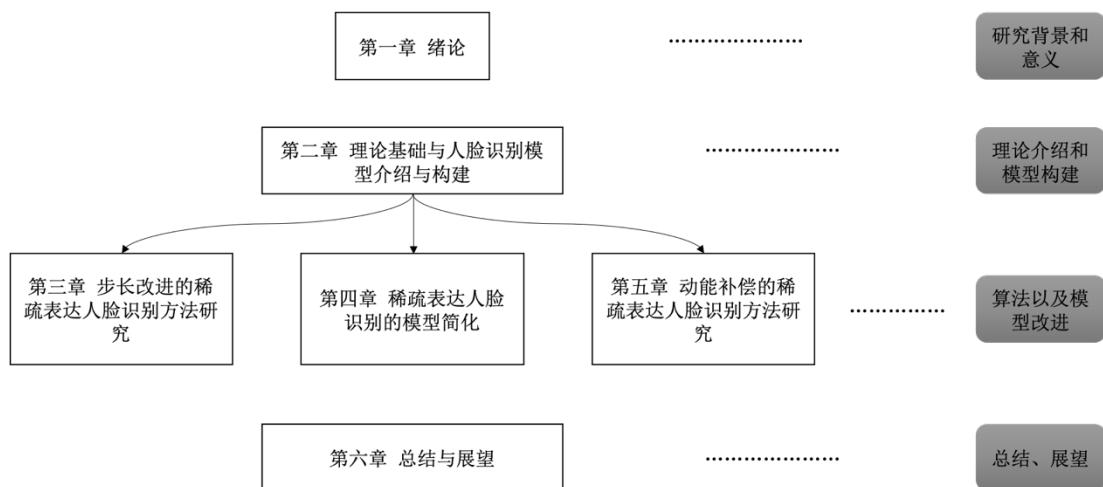


图 1-1 论文组织结构图

第一章，绪论。介绍课题的研究背景和意义、对本课题的国内外研究与发展

现状，阐述了课题的主要工作。

第二章，基本理论介绍与人脸识别模型搭建。给出了基本的图像信号处理理论并介绍了基于稀疏表达的人脸识别的算法及其模型，据此搭建了基于可训练稀疏编码的人脸识别模型框架。

第三章，步长改进的稀疏表达人脸识别方法研究。首先提出了基于步长调整和步长学习的两种稀疏编码方法，从这两方面的改进，设置多种对照实验，对其进行了仿真实验验证。随后，构建了基于这两种步长改进方法的人脸识别模型。最后，从总体上介绍了本文使用的人脸识别数据集及基本的实验设置，采用上述改进的模型对光照以及遮挡两种复杂环境下的人脸识别效率和速率进行实验验证，设置多种对照实验，并结果分析。

第四章，稀疏表达人脸识别的模型简化。本章通过一种简单的验证算法证实了一种既存的 tied 权重结构的有效性，多组仿真实验验证了简化模型的性能，通过将这种权重结构替换第三章所提稀疏表达人脸识别模型从而达到模型简化的效果。文章设计多组实验验证不同复杂环境下的改进权值的人脸识别的识别能力并进行实验结果分析。

第五章，动能补偿的稀疏表达人脸识别方法研究。本章首先提出稀疏表达需要进行动能补偿的假设，随后实验性的验证了假设内容，提出学习最优动能的策略，对该“最优”动能结果进行分析并通过仿真实验验证算法的可行性。其次，构建基于动能补偿的人脸识别模型，设计多组实验验证不同复杂环境下的动能补偿后的人脸识别的效率和速率并进行实验结果分析。

最后是总结与展望。对本文主要研究内容进行总结，指出了本文所提算法的优点和不足，并对今后需要进一步深入研究的工作进行展望。

第二章 稀疏表达人脸识别理论介绍及模型构建

2.1 稀疏表达人脸识别的理论介绍

对象识别中的一个基本问题是如何使用来自 s 个不同对象类别的标记训练样本来正确确定新测试样本所属的类别。稀疏解的估计和实现分类是稀疏表达人脸识别研究的两个主要问题，而其核心是如何得到更加精准的稀疏表达从而重建信号实现分类。

Wright 等人^[21]首次将稀疏表达应用于人脸识别，提出了基于稀疏表达的分类算法 SRC。该研究将第 i 类给出的 n_i 个训练样本安排为矩阵 $A_i = [v_{i,1}, v_{i,2}, \dots, v_{i,n_i}] \in \mathbb{R}^{m \times n_i}$ 的列，其中， m 为矩阵 A_i 的行， $m \in \mathbb{R}$ ， n_i 为第 i 类的列数， $n_i \in \mathbb{R}$ ， $v_{i,j}$ 是训练样本中第 i 类的第 j 个灰度图像($w \times h$)展开成向量 $v_{i,j} \in \mathbb{R}^m$ 得到的， $1 \leq i \leq s$, $1 \leq j \leq n_i$ 。于是，任意一个来自第 i 类的新的测试样本 $b_i \in \mathbb{R}^m$ 可以被线性表示为：

$$b_i = a_{i,1}v_{i,1} + a_{i,2}v_{i,2} + \dots + a_{i,n_i}v_{i,n_i}, \quad (2-1)$$

其中， $a_{i,1}$ 为对应向量 $v_{i,j}$ 的系数。进而，矩阵 A 可以用 s 个不同对象的训练样本表示为 $A = [A_1, A_2, \dots, A_s] \in \mathbb{R}^{m \times n}$ ，则任意的属于 $k \in \mathbb{R}$ 个不同对象的测试样本 $b \in \mathbb{R}^m$ 可以被线性表示为：

$$b = A\hat{x}, \quad (2-2)$$

其中， $\hat{x} = [0.0, \dots, a_{i,1}, a_{i,2}, \dots, a_{i,n_i}, \dots, 0.0]^T \in \mathbb{R}^n$ ，其是一个大部分元素都为零的稀疏向量。上述问题可以数学建模为一个 l_1 范数的最小化问题的求解：

$$\hat{x} = \arg \min_x \|x\|_1 \quad \text{s. t. } Ax = b, \quad (2-3)$$

其中， $x \in \mathbb{R}^n$ 。然而在大部分自然场景中，人脸识别往往受到多种复杂环境的影响，比如：光照、遮挡以及低像素模糊等。本文统称由于这种复杂环境造成的影响为“噪音项” $\varepsilon \in \mathbb{R}$ 。当假定“噪音项” ε 较小且均匀分布时，这个问题的求解会转变成：

$$\hat{x} = \arg \min_x \|x\|_1 \quad \text{s.t. } \|b - Ax\|_2 \leq \varepsilon. \quad (2-4)$$

由于 SRC 构造的感知矩阵 A 来源于训练样本，这意味着解 \hat{x} 会自然而然成为一个稀疏解，从而这是一个求解输入样本稀疏表达的问题。将其应用于人脸识别问题则可以表示为首先求解人脸的稀疏表达 \hat{x} ，其次重构图像 $\hat{b}_i = A\hat{x}_i$, $\hat{b}_i \in \mathbb{R}^m$, 其中 $\hat{x}_i = [a_{i,1}, a_{i,2}, \dots, a_{i,n_i}]^T \in \mathbb{R}^{n_i}$ ，最后通过计算 \hat{b}_i 与给定的 b 之差的最小值来确定类别 i 实现分类：

$$\min_i r_i(b) = \|b - A\hat{x}_i\|_2. \quad (2-5)$$

上述人脸识别的模型是基于一般的图像空间所提出，即直接采用图像转化为向量输入计算，而现实中的图像数据一般比较大计算较为复杂，计算量较大，当人脸识别从图像空间转化为特征空间的求解时，输入即为提取的特征转化为向量进行输入，这种求解方案计算会相对简单且在一定条件下识别能力也不会削弱：

$$\tilde{b} := Rb = RA\hat{x}, \quad (2-6)$$

其中， $R \in \mathbb{R}^{d \times m}$, $d \ll m$, $d \in \mathbb{R}$, $\tilde{b} \in \mathbb{R}^d$ 。综上可知，基于特征空间的最小化问题可以进一步转化为：

$$\hat{x} = \arg \min_x \|x\|_1 \quad \text{s.t. } RAx = \tilde{b}, \quad (2-7)$$

或

$$\hat{x} = \arg \min_x \|x\|_1 \quad \text{s.t. } \|\tilde{b} - RAx\|_2 \leq \varepsilon. \quad (2-8)$$

对于这两种基于图像空间和特征空间的最小化问题，一般采用稀疏编码中的迭代算法进行求解，而迭代算法求解的一个问题就是通常需要计算的迭代次数太多，计算消耗巨大。

2.1.1 人脸稀疏表达求解

从图像信号处理的角度，考虑一个有约束的最小化问题：其中 $f: \mathbb{R}^n \rightarrow \mathbb{R}$ 是光滑的函数， $g: \mathbb{R}^n \rightarrow \mathbb{R}$ 是一个可能不可微的凸函数：

$$\min_x f(x) + g(x), \quad (2-9)$$

如果 $f(x) = \frac{1}{2} \|b - Ax\|_2^2$ 并且 $g(x) = \lambda \|x\|_1$ ， λ 是一个正尺度可以调节稀疏强度。则式子(2-9)中的最小化问题称为 LASSO 回归^[60]，而此回归问题可以直接对应到基于特征空间的人脸稀疏表达求解的 l_1 最小化问题。

2.1.1.1 传统的稀疏编码算法

对于上述回归问题，有一种经典的迭代算法 ISTA^[26]可以对其进行求解：

$$x^{(k+1)} = SF(x^{(k)} + \alpha A^T(b - Ax^{(k)}), \alpha \lambda), \quad (2-10)$$

此处，SF是一个收缩算子，被定义为： $SF(x, \theta) = \text{sign}(x)\max(|x|-\theta, 0)$.正常数 α 是步长，步长的最佳值通常被分析得到是 $1/L$ ， L 是 ∇f 的最小的利普希茨常数，其值为 $A^T A$ 的最大特征值。与 ISTA 相比，FISTA^[27]的收缩运算是针对稀疏解 $z^{(k)} \in \mathbb{R}^n$ 起作用，其对 ISTA 添加“Nesterov”加速度，即用一个特殊的线性组合乘以软阈值收缩运算的最后两个输出之差，FISTA 可以写成：

$$z^{(1)} = x^{(0)}, t_{(1)} = 1, \quad (2-11)$$

$$x^{(k)} = SF(z^{(k)} + \alpha A^T(b - Az^{(k)}), \alpha \lambda), \quad (2-12)$$

$$t_{(k+1)} = \frac{1}{2} \left(1 + \sqrt{1 + 4t_{(k)}^2} \right), \quad (2-13)$$

$$z^{(k+1)} = z^{(k)} + \frac{t_{(k)} - 1}{t_{(k+1)}} (x^{(k)} - x^{(k-1)}). \quad (2-14)$$

其研究表明，ISTA 的复杂度为 $O(1/k)$ ，而 FISTA 在 ISTA 的基础上得到了有效改进，其复杂度为 $O(1/k^2)$ 。FISTA 在理论与实际中都被证明比 ISTA 更加有效。

2.1.1.2 可训练的稀疏编码算法

上文求解稀疏解的传统方法主要是基于一种反复迭代逼近“真实”的稀疏解的数学迭代模型来实现，通常需要较多的次数才能得到理想的结果。而一种使用编码器进行稀疏代码预测的基本思想^[61-62]，以及将这种编码器转变成深度学习模型的特殊形式和进行特殊参数化的策略^[36]，实现了仅需少数几次迭代就能快速逼近稀疏解的功能。同时，这种将迭代算法展开转化成深度神经网络的方法结合了深度学习能够快速学习的特点，还让深度学习这种“黑匣子”式的训练方式在稀疏编码器中得到有效的解释。

一种可学习的迭代算法 LISTA^[36]就显示了其只需要很少的迭代次数就能快速得到更优稀疏解的能力，该研究重写 ISTA^[26]为：

$$x^{(k+1)} = \text{SF}(W_e b + Sx^{(k)}, \theta), \quad (2-15)$$

其中， $W_e \in \mathbb{R}^{n \times m}$, $S \in \mathbb{R}^{n \times n}$, $\theta \in \mathbb{R}$, 并被初始化为 $W_e = A^T/L, S = I - (A^T A)/L, \theta = \lambda/L$ ，该算法的训练参数为 $\Theta = \{(W_e, S, \theta)\}$ 。该方法将 ISTA 看作是一个循环神经网络，其框图如图 2-1(a)所示，通过将 ISTA 迭代层展开作为深度神经网络层，该研究进而展示了 LISTA 的共享权重的结构框图（如图 2-1(b)所示），同时对坐标下降算法^[37]也采取同样的策略，并取得了显著的实验结果。

这种可训练并高效的算法对信号与图像处理方面的研究具有重大意义，主要体现在图像信号去噪、恢复与重建等方面。其新颖性主要在于将一般的求解 LASSO 回归的迭代算法展开形成计算图，权重结构的设计确保可以使用基于梯度的学习方法来训练它们，并可以确保通过它们对梯度进行反向传播。其高效性吸引着国内外许多学者对其从理论到应用方面进行了更加深入的研究。本文主要基于迭代软阈值算法 ISTA，对其衍生的学习型迭代算法进行讨论与研究。

2.2 稀疏表达人脸识别的模型架构

稀疏表达人脸识别主要可以分解为两个步骤：第一是编码问题即求解稀疏解；第二是解码问题即重构图像进行分类。本文探讨的主题是人脸识别中的编码问题，从上文可知稀疏表达的人脸识别的编码问题是解决 l_1 范数最小化的问题。解决这种最小化问题，本文重点从 2.1 节所介绍的传统的稀疏编码手段以及可训练的稀疏编码手段进行考虑。针对不同的编码手段可以构建不同的稀疏表达人脸识别模型，本小节首先介绍稀疏表达人脸识别模型 SRC^[21]，其次阐述基于可训

练习稀疏编码手段的人脸识别模型。

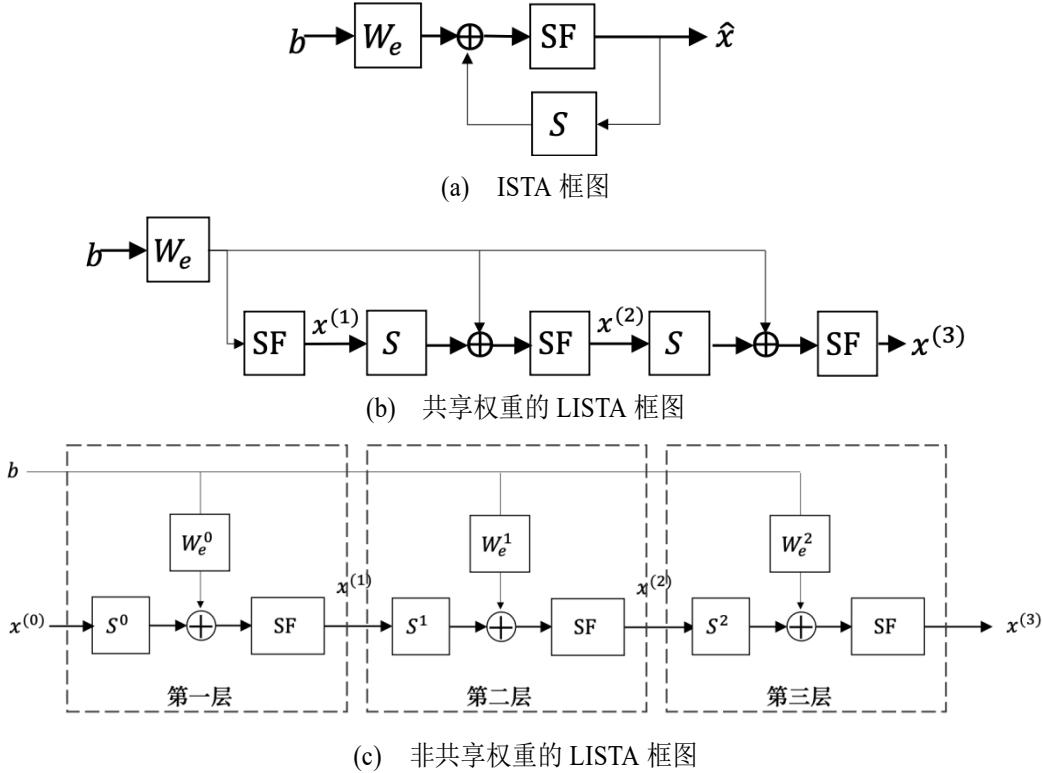


图 2-1 ISTA 和 LISTA 算法的结构框图。(b) 图和 (c) 图层数次限定为三层, $x^{(0)}$ 是一个 n 维的零向量

2.2.1 稀疏表达人脸识别模型 SRC

稀疏表达人脸识别模型的主要模块包括 (1) 从图像空间到特征空间的转化; (2) 稀疏解的估计; (3) 图像重构; (4) 图像分类。本章给出具体的算法模型 (如图 2-2) 与算法步骤 (如图 2-3), 以下分别对其进行介绍:

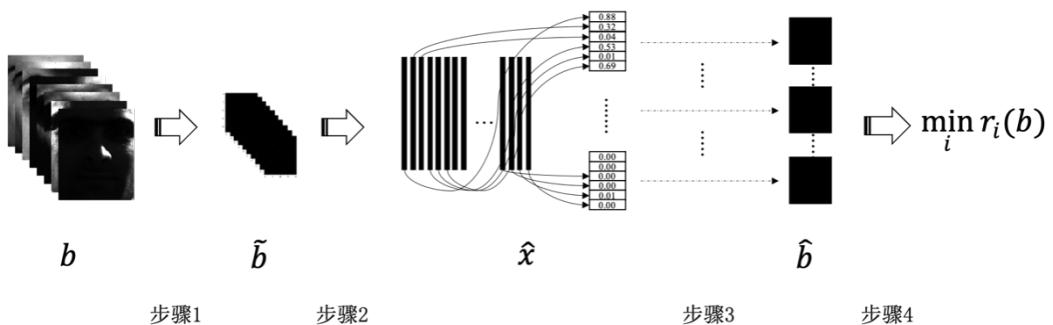


图 2-2 SRC 人脸识别模型结构图。步骤 1 为图像空间到特征空间的转化; 步骤 2 为稀疏解的估计; 步骤 3 为图像重构; 步骤 4 为图像分类

(1) 图像空间到特征空间的转化：SRC 模型的特征空间通过随机矩阵产生得到。基于压缩感知的原理，在一定条件下，高维空间对低维空间的映射可以通过随机矩阵所产生，随机矩阵一般是从标准高斯分布 $R_{ij} \sim N(0, 1/m)$ 中采样所得。采用这种随机变换 R ，可以得到人脸图像的随机特征空间。这种特征提取方式相比传统的特征提取方式以及基于深度学习的特征提取方式要更简单并具有可比的性能。

(2) 稀疏解的估计：SRC 模型稀疏解的估计是将得到的特征空间作为输入，求解 ℓ_1 范数最小化的问题。

(3) 图像重构：图像重构是根据稀疏解与字典重构图像，字典分量与稀疏解的对应分量进行线性组合得到。

(4) 图像分类：通过计算重构图像与输入图像的最小误差从而得到最小误差所指向的对象类别实现分类。

1. 输入：构建由 s 类训练样本组成的矩阵 $A = [A_1, A_2, \dots, A_s] \in \mathbb{R}^{m \times n}$,

输入图像 $b \in \mathbb{R}^m$;

2. 对矩阵 A 的每一列进行归一化；

3. 求解 ℓ_1 最小化问题：

$$\hat{x} = \arg \min_x \|x\|_1 \quad \text{s.t. } Ax = b$$

或者

$$\hat{x} = \arg \min_x \|x\|_1 \quad \text{s.t. } \|b - Ax\|_2 \leq \varepsilon$$

4. 计算误差：

$$\text{for } i=1, 2, \dots, s \text{ 执行 } \min_i r_i(b) = \|b - A\hat{x}_i\|_2;$$

5. 输出：identity(b) = $\arg \min_i r_i(b)$ 。

图 2-3 稀疏表达人脸识别 SRC 算法的步骤

2.2.2 可训练稀疏表达的人脸识别模型构建

与传统的基于稀疏表达的人脸识别算法不同，本文采用将迭代算法展开为深度神经网络的形式，即图 2-2 中所示的步骤 2、3 和 4 为深度学习的网络模型，主要分为稀疏解求解层以及图形重构分类层。本文所提出的模型的权重矩阵以及

分类方式与传统的深度学习的方式有所不同，本文是从稀疏表达的角度来进行模型构建。根据不同的迭代次数以及不同的迭代算法及其变体算法可以构建不同类型的稀疏编码人脸识别模型。本文主要探讨基于迭代软阈值算法的不同变体算法及其处于不同迭代次数下的人脸识别性能。其算法模块不变依旧为：（1）从图像空间到特征空间的转化；（2）稀疏解的估计；（3）图像重构；（4）图像分类。但其模型结构发生变化（如图 2-4 所示），具体的算法步骤如图 2-5 所示，主要体现为：

(1) 稀疏解的估计：稀疏解的估计从简单的算法的求解变成将迭代算法展开成深度神经网络层，其意味着，稀疏解的估计的结构转变成根据不同迭代次数而来的不同深度学习网络层。比如采用 2 次迭代的 LISTA 的稀疏解估计，将会变成 2 层深度神经网络层。

(2) 图像重构：从简单的矩阵与稀疏解分量的计算变成了一个深度网络神经层。

(3) 图像分类：从简单的最小二乘计算进而得到识别类别变成一个 softmax 层。通过 softmax 层得到的结果计算损失，反向传播进行优化从而得到更优结果。

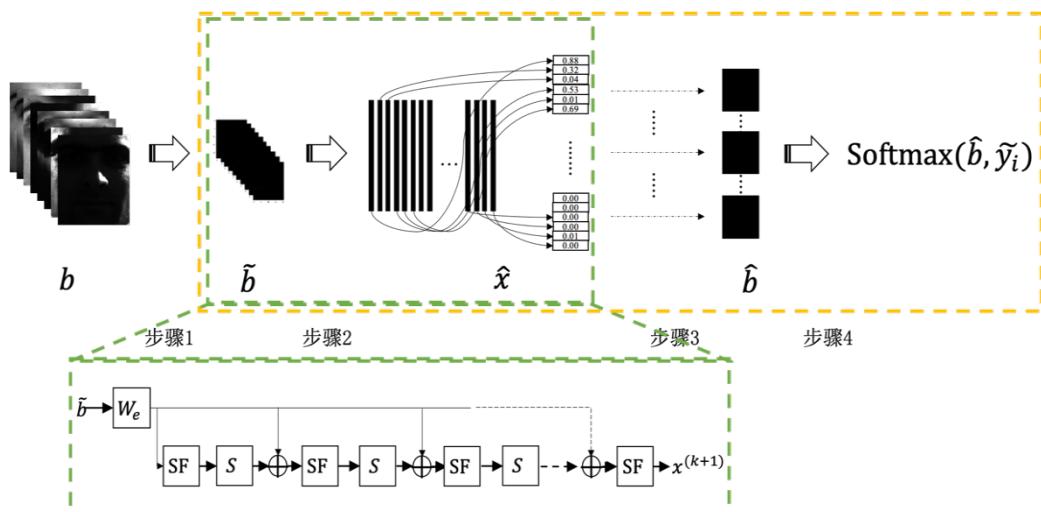


图 2-4 可训练稀疏编码人脸识别模型结构图。橙色方框为转化为深度网络层的模型结构，绿色框部分为采用基于 LISTA 的稀疏编码的展开结构。步骤 1 为图像空间到特征空间的转化；步骤 2 为稀疏解的估计层；步骤 3 为图像重构层；步骤 4 为图像分类层，其中 \tilde{y}_i 表示识别标签

损失函数设计：可训练稀疏编码的人脸识别模型需要对其损失函数进行设计，其中，实现识别的深度学习算法一般采用交叉损失熵函数作为损失函数：

$$L_1(\Theta) = - \sum_{i=0}^{N-1} \tilde{y}_i \log(b^*_i(\Theta^{(k)})), \quad (2-16)$$

其中, \tilde{y}_i 表示识别标签, $b^*_i(\cdot)$ 是预测的结果向量, 该函数能够较好的评价标签和预测结果之间的损失。但是基于稀疏表达的人脸识别算法还需对稀疏解做一定的正则化处理和稀疏性控制。因此, 本文可训练稀疏编码的人脸识别方法的损失函数设计为:

$$L(\Theta) = \frac{1}{N} (L_1(\Theta) + L_2(\Theta)), \quad (2-17)$$

$$L_2(\Theta) = \sum_{i=0}^{N-1} \|\tilde{b} - \tilde{A}x_i^k(\Theta^{(k)})\|_2^2 + \lambda \|x_i^k(\Theta^{(k)})\|_1. \quad (2-18)$$

1. 预处理:

- a) 输入: 构建由 s 类训练样本组成的矩阵 $A = [A_1, A_2, \dots, A_s] \in \mathbb{R}^{m \times n}$, 输入图像 $b \in \mathbb{R}^m$;
- b) 采用随机矩阵 $R \in \mathbb{R}^{d \times m}$, 将图像空间转化为随机特征空间 $\tilde{A} = RA$, $\tilde{b} = Rb$;
- c) 对矩阵 \tilde{A} 的每一列进行归一化;

2. 求解 L_1 最小化问题:

$$\hat{x} = \arg \min_x \|x\|_1 \quad \text{s. t. } \tilde{A}x = \tilde{b}; \text{ 或者 } \hat{x} = \arg \min_x \|x\|_1 \quad \text{s. t. } \|\tilde{b} - \tilde{A}x\|_2 \leq \varepsilon.$$

采用基于可训练的稀疏编码方法解决上述问题 (前向传播) :

for $i=1,2,\dots,k$ (迭代次数)

$$x^{(k+1)} = \text{SF}\left(W_e \tilde{b} + Sx^{(k)}, \theta\right)$$

for $i=1,2,\dots,s$ (类别数) 执行 $\min_i r_i(\tilde{b}) = \|\tilde{b} - \tilde{A}\hat{x}_i\|_2$;

图 2-5 可训练稀疏编码人脸识别算法的步骤

2.3 本章小结

本章首先介绍基于稀疏表达的人脸识别方法 SRC 的理论基础, 阐述了其算法思想; 其次, 介绍了其原有的算法模型, 在此基础上构建了本文所提出的稀疏表达的人脸识别模型。其中, SRC 算法主要分成两个部分, 稀疏解求解以及实现分类, 本章从基础理论出发, 基于本文研究核心稀疏解的估计介绍了估计稀疏解的两种手段: 传统稀疏编码迭代算法和可训练的稀疏编码迭代算法。其次, 本章在既有人脸识别模型的基础上构建改进的算法模型, 并设计该模型所需用到的损失函数。

第三章 步长改进的稀疏表达人脸识别方法研究

3.1 引言

SRC 人脸识别方法提出采用训练样本来构建感知矩阵 A , 解决 ℓ_1 最小化问题求解稀疏解, 从而通过稀疏表达重构图像, 通过计算重构图像与原始图像之间的误差最小值来达到分类目的。这种方法的主要难点在于如何高质量并快速的重构图像, 在不同的环境下, 比如光照、妆容和遮挡等都增加了图像重构的难度与速度。SRC 策略的矩阵 A 的构建较为简单, 使得环境变化较大时其会出现明显的鲁棒性不足的缺陷。许多研究^[23,25]通过对感知矩阵 A 进行改进来重建图像, 然而对感知矩阵 A 进行改进的方式会在一定程度上增加计算量与资源消耗。这些研究得到相对较好的字典学习后的矩阵 A 后进行最小化问题求解时仍然需要几十次或更多的迭代次数才能得到相对优秀的稀疏解, 这种方法对算法迭代次数要求较多。

可训练的稀疏编码是一种通过将传统迭代算法展开成深度神经网络进行学习实现快速逼近稀疏解的一种有效编码方式。研究^[27]表明, 迭代软阈值算法 ISTA 的收敛速率为 $O(1/k)$, 快速迭代阈值算法 FISTA 为 $O(1/k^2)$, 而相比于以上两种传统的迭代阈值算法, 可训练的稀疏编码算法从理论^[44,46,49,51,52]到实际应用^[53-58]都表现出其在逼近稀疏解的收敛速度相对更快。

可训练稀疏编码在近十年中得到迅速的发展和研究。通常, ISTA 的步长 α 被计算为 (最小) Lipschitz 常数 L 的倒数, L 为 $A^T A$ 的最大特征值, LISTA 算法中, 步长 α 仅仅作为初始值被包含在权重或阈值中进行训练。本文认为稀疏编码算法尤其是基于梯度下降策略的稀疏编码算法 ISTA 中的步长是逼近“真实”稀疏解的关键点。合适的步长可以有效加快稀疏解的估计速度并减少迭代次数。本章从步长调整角度出发结合字典学习改进 LISTA, 构建基于此方法的人脸识别模型, 实现快速人脸稀疏表达, 进而加快人脸识别速率并提高识别准确率。

3.2 步长改进的稀疏编码模型

本小节针对可训练稀疏编码算法的步长提出调整和学习两种不同的策略, 分别建立稀疏表达人脸识别模型并给出实验结果与分析。

3.2.1 调整步长模型

研究^[44]表明，使用部分权重耦合结构可以帮助减少展开的 ISTA 中的可训练参数。这种结构是本文接下来要考虑的结构。在该部分权重耦合结构中，每一个迭代层的训练参数为： $\Theta^k = \{W_e^{(k)}, \theta^{(k)}\}$ ，步长 α 被包括在矩阵 W_e 和阈值变量 θ 中，并且该方法分别被初始化为 $W_e = \alpha A^T$ 和 $\theta = \alpha \lambda$, $\alpha = 1/L$ 为初始值。ISTA 中的步长是快速逼近“真实”稀疏解的关键点。如果步长太小，则需要更多的迭代次数或层数来实现期望的有效稀疏解。相反，如果步长太大，或者 Lipschitz 常数 L 太小，则 ISTA 可能会不稳定并发散。事实证明，在加快稀疏解的估计速度方面，合适的 Lipschitz 常数 L 非常重要。

为了保持“最佳”步长以获得更好的性能，本文建议将步长 α 保持为分析得到的最大奇异值的平方 $\sigma_{max}^2(A^T)$ ，同时字典 A^T 是可训练的。本文在字典学习后使用 SVD 来微调步长，使正尺度 λ 得以学习来利用稀疏性。本文的新颖之处在于在展开的 ISTA 中使用反向传播的本征分解来进行步长调整。在这种方案中，本文采用 W_* 来表示矩阵 A^T , $W_* \in \mathbb{R}^{n \times m}$:

$$x^{(k+1)} = \text{SF}(x^{(k)} + \alpha W_*^{(k)}(b - Ax^{(k)}), \lambda^{(k)}\alpha), \quad (3-1)$$

该算法的可训练参数为 $\Theta = \{(W_*^{(k)}, \lambda^{(k)})\}_{k=0}^{K-1}$ ，步长 α 将在训练过程中自动计算。本文将此方案命名为 TsISTA。

3.2.2 学习步长模型

使用 TsISTA 为学习的 ISTA 计算“最佳”步长可以加快接近真实稀疏解的速度，但是，当前研究结果受到耗时的 SVD 计算的限制。研究^[63]提出了一种策略来学习展开的 ISTA 的步长，将步长设置为学习的 ISTA 中唯一的可训练的参数，并以无监督的方式训练网络：

$$x^{(k+1)} = \text{SF}(x^{(k)} + \alpha^{(k)} A^T(b - Ax^{(k)}), \alpha^{(k)}\lambda), \quad (3-2)$$

其中， $\alpha^{(k)}$ 的初始值为 $1/L$ 。该研究结果表明，在无监督学习的情况下，当稀疏解足够稀疏时，学习的步长方案比其他算法更好，能快速收敛。这意味着，如果本文充分利用正尺度 λ 来使其稀疏，则独立训练步长是有效的。为了减少训练模型产生的时间的消耗，本文建议学习步长，而不是使用 SVD 自动计算步长。因此，

学习步长版本的 TsLISTA 将变为：

$$x^{(k+1)} = \text{SF}(x^{(k)} + \alpha^{(k)} W_*^{(k)}(b - Ax^{(k)}), \lambda^{(k)} \alpha^{(k)}), \quad (3-3)$$

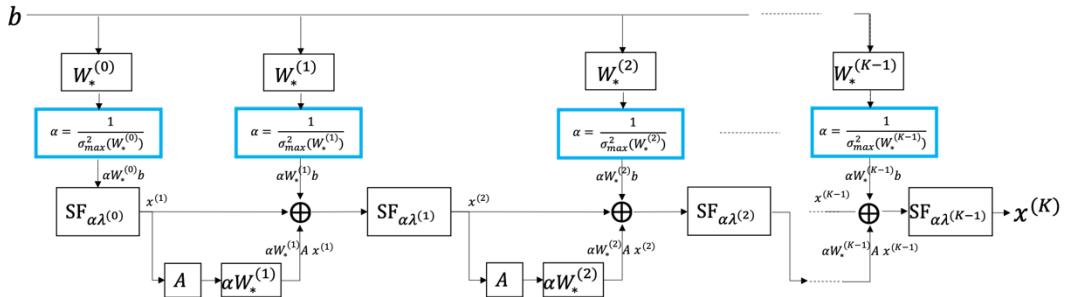
其可训练参数为 $\Theta = \{W_*^{(k)}, \alpha^{(k)}, \lambda^{(k)}\}_{k=0}^{K-1}$ 。本文将此方案命名为 LsLISTA。

研究^[63]将 LISTA 的每一层网络模型视为参数化的 $\Theta^k = \{W, \alpha, \beta\} \in \mathbb{R}^{m \times n} \times \mathbb{R}_*^+ \times \mathbb{R}_*^+$ 后的函数 $\varphi_\theta: \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^n$ ：

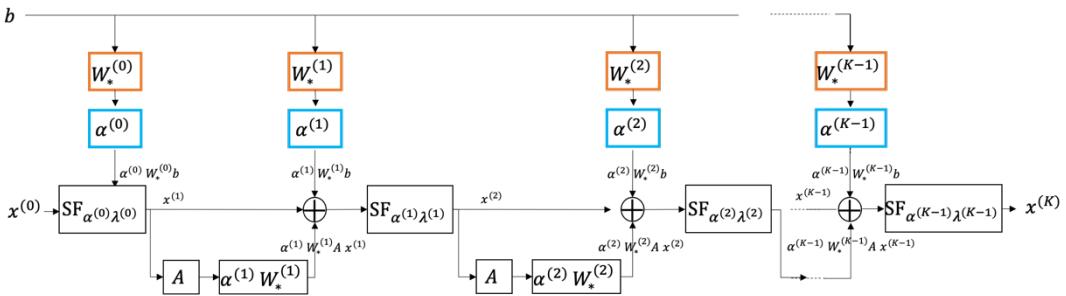
$$\varphi_\theta(x, b) = \text{SF}(x + \alpha W^T(b - Ax^{(k)}), \beta \lambda), \quad (3-4)$$

当 $W = A, \alpha = \beta = 1/L$ 时，其输出的结果与 ISTA 的一次迭代相同。在这种情况下，本文将 LISTA 层视为由 $\Theta^k = \{W, \alpha, \lambda\}$ 参数化的函数 φ_θ 。

3.2.3 算法框架



(a) K 层 TsLISTA 算法的框图。蓝色方框是使用 SVD 自动计算步长 α 的方案。 $\sigma_{\max}^2(\mathbf{W}_*)$ 是 \mathbf{W} 的最大奇异值的平方。计算出的步长也用于软收缩算子



(b) K 层 LsLISTA 算法的框图。蓝色方框是学习步长 α 的方案。橙色方框 $\mathbf{W}_*^{(k)}$ 是一个非共享的权重。学到的步长也用于软收缩算子

图 3-1 TsLISTA 与 LsLISTA 算法的结构框图

针对上述提出的两种不同基于步长的改进的稀疏编码算法，本小节给出其基本的算法框图，而此处仅为稀疏解求解的具体结构框图（如图 3-1 所示）。通过比较图 3-1（a）和（b），可以清晰看到学习步长的方案减少了 SVD 计算的成本。

3.2.4 仿真实验

本小节是为了验证本节所提算法而进行的仿真实验。仿真实验是通过假设存在真实的图像信号进而对无噪声或存在噪声的图像信号进行重建的模拟过程，其能够良好的反映算法的表达能力以及性能，同时仿真实验的结果能够体现本文算法同样适用于图像信号恢复与重建。本文首先介绍具体的仿真实验设置，接下来对本文所提算法 TsLISTA 和 LsLISTA 进行实验验证，并给出实验结果分析。

3.2.4.1 实验设置小节，对仿真实验的数据集、实验数据划分以及实验环境进行说明，本实验设置同样适用于第四章和第五章中的仿真实验设置。

3.2.4.1 实验设置

感知矩阵 A 形状为 $m = 250, n = 500$ 。 A 是从标准高斯分布 $A_{ij} \sim N(0, 1/m)$ 中采样得到的。对于基准稀疏向量 x^* ，非零项的采样个数为 n 的 10%，每个非零项都从标准高斯分布中采样，并且该样本假定所有项上的分布均匀， $\lambda = 0.4$ 。训练批尺度大小为 64，验证批尺度大小为 1000，测试数据为 1000，优化器为 Adam。应对不同的噪音设置，本文会模拟相应的测试数据，所有的算法在相应的测试数据上进行测试，来保证每个算法的实验环境一致。对于训练和验证数据，本文基于一种理想情况进行产生，每个批尺度的产生都遵循同样的分布并实时产生，训练过程遵循研究^[44,45]中以不同的学习速率训练数据的方法¹，本文的实验代码在未来会进行开源。实验环境为 TensorFlow-GPU 2.2.0，GPU 是 GeForce RTX 2080。本文使用 NMSE^[44]来评估算法：

$$\text{NMSE}(\hat{x}, x^*) = 10\log_{10} \frac{\mathbb{E}\|x - x^*\|_2^2}{\mathbb{E}\|x^*\|_2^2}. \quad (3-5)$$

本文通过两组仿真实验来进行算法评估，分别为无噪声和有噪声情况下的对照实验。无噪声时 SNR 为 ∞ ，有噪声时 SNR 分别为 20 和 10。

3.2.4.2 仿真结果分析

首先是调整步长的验证：针对步长调整算法 TsLISTA，由于 TsLISTA 具有

¹ <https://github.com/VITA-Group/LISTA-CPSS>

与 LISTA-CP 相同数量的可训练参数，因此 LISTA 和 LISTA-CP 设置为对比算法，同时传统迭代算法 ISTA 和 FISTA 同样作为对照算法。由于研究^[64]中提出 tied 结构下的 LISTA 结果更优同时本文将在下一章对此进行验证，为了体现本文算法的优势，在本实验中，本文使用 LISTA-t^[64]作为比较。结果如图 3-2 所示。

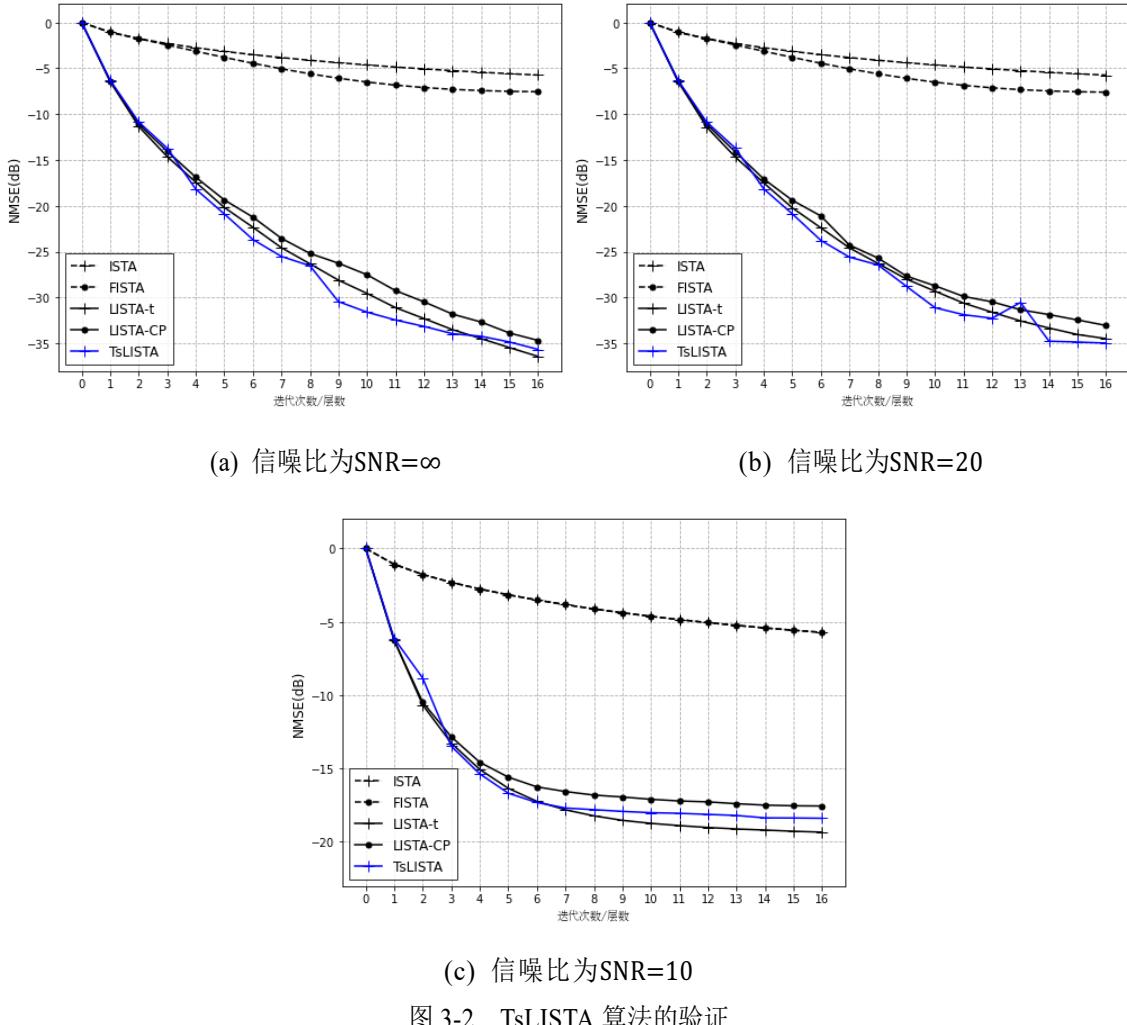


图 3-2 TsLISTA 算法的验证

实验结果分析如下：

(1) 对比传统的稀疏编码算法 ISTA、FISTA 和可训练的稀疏编码算法 LISTA、LISTA-CP 和 TsLISTA，可训练的迭代算法仅仅少数几次迭代就能快速逼近稀疏解。

(2) TsLISTA 与 LISTA-CP 可训练参数数量相同，但 TsLISTA 性能更优。对比 LISTA-t 算法和本文所提算法 TsLISTA，不同信噪比的情况下，信噪比较大的时候 TsLISTA 稍微更优；信噪比为 10 时，LISTA-t 更优，但 TsLISTA 仍然具有可比性，TsLISTA 得到直接验证。

其次是学习步长的验证：针对展开的 ISTA 的步长学习算法 LsLISTA，本文

通过设置多种不同类型的对比算法来体现本文所提算法的优势。第一、由于 LsLISTA 具有与 LISTA-CP 类似的权重结构设计，因此将 LISTA-CP 设置为主要对比算法；第二、本文主要研究的是基于 ISTA 的迭代展开算法，为了丰富对照实验的多样性，本文同时采用基于 AMP 改进的可训练迭代算法 LAMP^[64]作为对照实验；第三、为了体现本文所提学习步长策略与调整步长策略的区别性，TsLISTA 也作为对比算法。实验结果如图 3-3 所示。

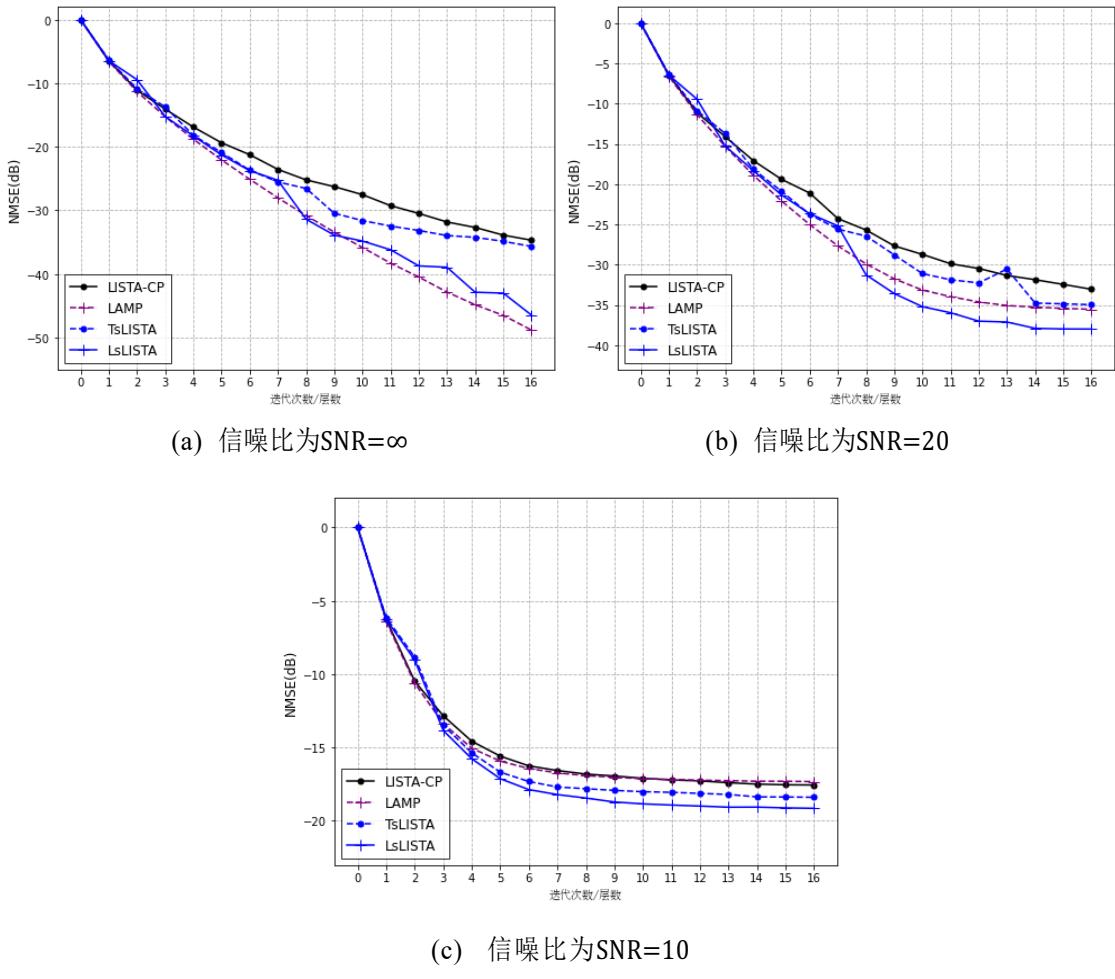


图 3-3 LsLISTA 算法的验证

实验结果分析如下：

(1) 通过对比 LISTA-CP 算法与本章所提算法 LsLISTA, LsLISTA 性能较优，据此可以得出，将步长独立训练并通过正尺度 λ 学习平衡稀疏度比直接训练阈值 θ 更加有效，分离步长的策略得以验证。

(2) 通过对比 LAMP 算法与本章所提改进步长的算法算法 TsLISTA 与 LsLISTA，本文所提两种算法在信噪比更低的时候性能稳定并较其更好，而在实际生活中大部分的场景都存在一定程度的噪音，本文所提算法符合实际应用场景。

景。

(3) 通过对比本章所提两种算法 TsLISTA 与 LsLISTA, LsLISTA 在 TsLISTA 算法的基础上进行了改进, 得到了学习步长的方案, 其性能得到提升, 同时减少了由于 SVD 计算带来的计算和存储负担, 因此学习步长方案更加有效。

3.3 步长改进的稀疏表达人脸识别

本小节基于上文中对步长改进的两种算法构建人脸识别模型并进行实验结果分析。

3.3.1 模型求解

本文专注编码部分的提升, 提高通过少数几次迭代就能快速获取人脸稀疏表达的能力。本章探讨的问题是通过步长改进实现快速求解稀疏解属于人脸识别中的编码问题, 通过对编码的改进来提高人脸识别稀疏表达性能。从上文 2.2 小节介绍的稀疏编码人脸识别模型可以得知, 基于可学习的稀疏编码进行人脸识别由四个部分组成, 本小节主要针对稀疏人脸表达的编码方面进行讨论。

首先是求解人脸稀疏表达, 本文是围绕从图像空间转化为随机特征空间的稀疏表达人脸识别进行求解, 即求解式子 (2-7) 或 (2-8) 中的最小化问题。采用调整步长的方法可以表示为:

$$x^{(k+1)} = \text{SF}\left(x^{(k)} + \alpha \tilde{W}_*^{(k)}(\tilde{b} - RAx^{(k)}), \lambda^{(k)}\alpha\right), \quad (3-6)$$

其中 $\tilde{W}_*^{(k)} = (RA)^T \in \mathbb{R}^{d \times n}$ 。

采用调整步长的方法可以表示为:

$$x^{(k+1)} = \text{SF}\left(x^{(k)} + \alpha^{(k)} \tilde{W}_*^{(k)}(\tilde{b} - RAx^{(k)}), \lambda^{(k)}\alpha^{(k)}\right), \quad (3-7)$$

这两种方案的具体的结构示意图如图 3-4 所示。可以看到学习步长的方法减少了对步长的 SVD 计算步骤, 对于大规模的应用比如本文的人脸识别应用来说会大幅度减少计算量, 模型结构也相对简单。

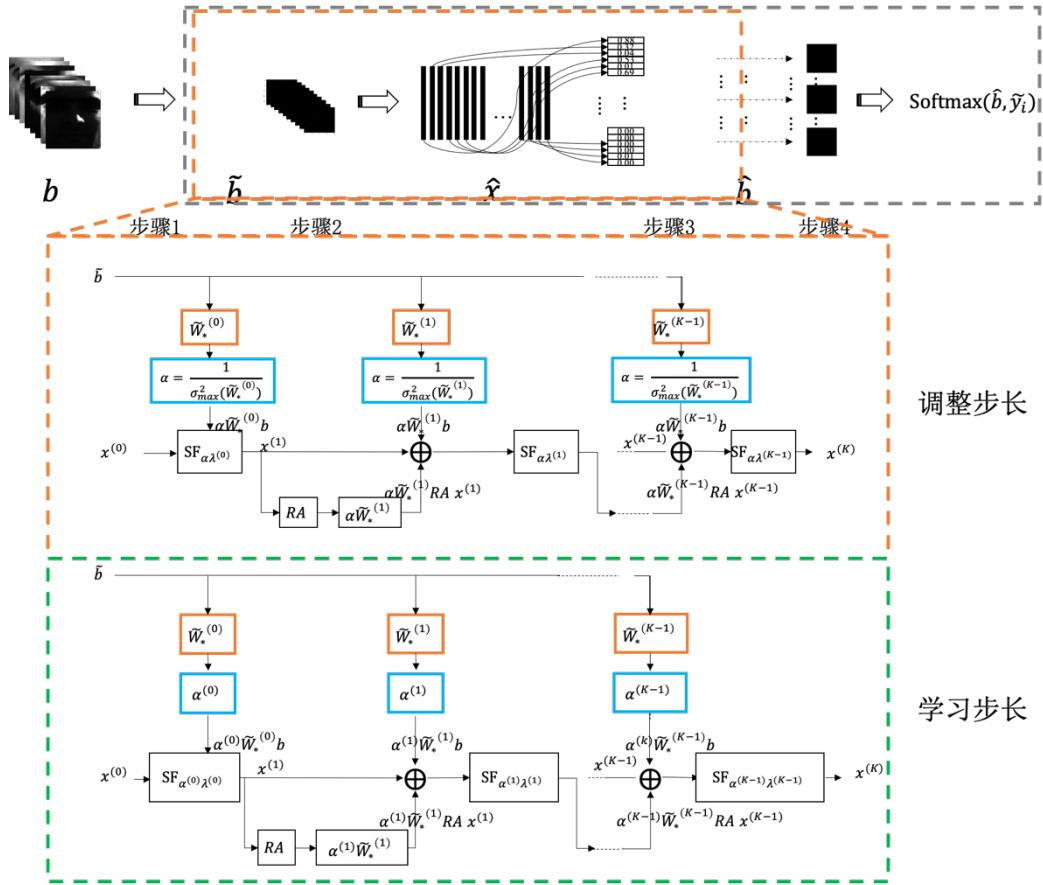


图 3-4 基于步长改进的稀疏编码人脸识别模型框架图。上半部分为流程结构图，其中灰色虚线方框内为采用可迭代算法展开为深度神经网络后的网络结构；下半部分橙色方框内为调整步长的稀疏编码方法，绿色方框内为训练步长的稀疏编码方法

3.3.2 实验设置

本小节对数据集、实验数据划分以及实验环境进行说明，本实验设置同样适用于第四章和第五章中的人脸识别实验设置。

3.3.2.1 数据集介绍

本文主要对光照以及遮挡等复合环境下的人脸识别方法进行研究，数据集选用 Extend Yale B^[65] 和 AR 人脸数据集^[66]，下面将分别对这两种数据集进行介绍并提供本文使用的样本数据。

Extend Yale B 包含 38 个个体处于 64 种不同照明条件下的 2,414 个正面人脸图像，此数据集能够良好的较好的检验不同光照环境下的人脸识别算法的性能。本文选取其中几种光照环境变化较大的正面人脸图像作为示例展现在图 3-5 中。



图 3-5 Extend Yale B 数据集示例展示。选取的示例图像来自于同一个人的不同光照环境下的图像，可以看到此数据集对光照条件要求比较严格，能较好的反映真实生活场景的环境

AR 人脸数据集包含 126 人的人脸正面图，其中包括 70 位男性和 56 位女性，总计超过 4000 张彩色图像。整个实验采集过程包括两个部分，这两个部分采集的内容一样，包括不同的面部表情，照明条件和遮挡等，时间相隔 14 天。AR 人脸数据集的图像的光照条件没有像 Extend Yale B 那样严格从不同角度进行采集，光照主要是左或右或左右同时采光拍摄的图像，但 AR 人脸数据集对人脸的妆容、面部表情以及是否佩戴眼镜等方面都没有做严格的要求，所以该数据集较为贴切真实场景，其遮挡环境主要包括太阳镜和围巾所造成的对眼睛部分以及下巴及嘴部的遮挡。其中每个个体包含 26 张图片，相比于 Extend Yale B 来说针对某一种类别的图像数量会相对较少（比如仅含光照条件的人脸正面图仅有 8 张），矩阵 A 的列数大幅减少，此外由于测试者对于佩戴眼镜以及妆容方面没有限制，这也使得识别难度也会增大。本文以光照条件的四张图像以及遮挡条件的六张图像作为示例展现在图 3-6 中，本文将图像转化为灰度图并且经过裁剪后的图像进行展示。



图 3-6 AR 人脸数据集示例展示。选取的示例图像来自于同一个人的不同光照自然状态和佩戴太阳镜以及佩戴围巾遮挡的图像，可以看到此数据集所设置环境结合光照和其他遮挡条件，能较好的反映真实生活场景的环境

3.3.2.2 识别环境

本文采用不同的数据集进行测试，复杂环境主要划分为光照和光照与遮挡以及其他环境的复合环境，数据集的划分采用随机选择方式，对数据集的随机划分能确保本文的结果和结论不依赖于对数据集的任何特殊选择。

本文首先对光照环境下的数据集划分进行解释说明，本文对光照环境下的人脸识别测试采用两种数据集分别进行测试，Extend Yale B 数据集和 AR 人脸数据集。以下分别介绍对上述两种数据集所采取的前期预处理手段。

(1) Extend Yale B 数据集适用于光照环境的人脸识别测试，下载得到的裁剪图像的尺寸大小为 192×168 ，总个体 38 人、2414 张图像，根据图像空间到特征空间的转化，本文将图像空间转化到特征空间的 56 维 (8×7)、120 维 (12×10) 和 504 维 (24×21)。数据集按照 2:1:1 的比例进行数据划分，其中得到训练集为 1205，验证集为 604，测试集为 605。

(2) 对于 AR 人脸数据集，本文主要提取其中的光照以及遮挡的数据作为本文的仿真数据。本文选取男 60 人女 50 人作为样本对象，图像被裁减到 240×200 的尺寸大小，本文将图像空间转化到特征空间 120 维 (12×10)。这 110 个样本对象的图像总量为 2860 张，光照条件的样本总量为 880 张，其中自然光条件下的有 220 张。针对光照条件下的人脸识别测试，本文选取光照条件的样本 880 张作为光照条件的数据集。数据集按照 2:1:1 的比例进行数据划分，其中得到训练集为 440，验证集为 220，测试集为 220。

本文其次对光照与遮挡以及其他环境的复合环境下的数据集划分进行解释说明，本文对遮挡以及其他环境下的人脸识别采用 AR 人脸数据集进行仿真测验。根据上文描述，在本文选取的 110 个对象的 2860 张图像中，遮挡条件的样本总量为 1320 张，其中自然光条件下的为 440 张。针对遮挡条件下的人脸识别测试，本文结合光照条件下的样本中自然光照条件的 220 张图像和遮挡条件的 1320 张图像一起作为数据集总计 1540 张图像。数据集按照 2:1:1 的比例进行数据划分，其中得到训练集为 770，验证集为 330，测试集为 440。输入特征空间为 120 维 (12×10)。

3.3.2.3 其他设置

SRC 中感知矩阵 A 的构建是通过训练数据进行构建，本文同样采用此种方法构建初始化 A 。

3.3.3 结果分析

本文主要研究如何快速求解人脸稀疏表达的 ℓ_1 最小化问题，本小节对比算法

采用传统的迭代算法 ISTA 和 FISTA 以及经典的基于学习的 LISTA。

3.3.3.1 光照条件下的人脸识别

针对光照条件下的人脸识别，本小节总共分为两个部分来对本章所提的算法进行验证。第一部分采用 Extend Yale B 数据集对算法进行验证，第二部分采用 AR 人脸数据集对算法进行验证。

Extend Yale B 数据集上的实验结果：本小节针对随机特征空间为 120 维和 56 维进行实验，迭代次数为 10 次，结果如图 3-7 所示。

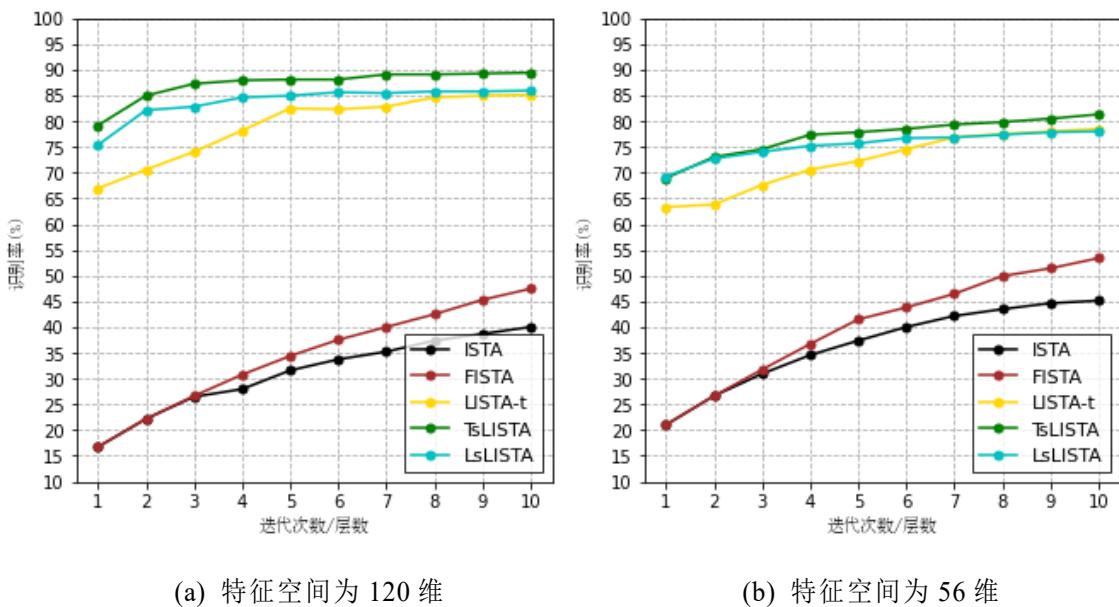


图 3-7 Extend Yale B 数据集上改进算法的验证

通过分析图 3-7 可知，本文采用可训练的稀疏编码模型替换传统的编码模型后识别效率以及速率都有极大的提升。其中，调整步长的方案 TsLISTA 相比于其他方案都得到了明显的性能提升，它仅仅需要一次迭代就可以分别在特征空间为 120 维、56 维的情况下达到 79.17% 和 68.93% 的识别率。而达到同等识别率，基准算法 LISTA 分别需要 3 和 4 次迭代左右，传统编码算法 ISTA 分别需要超过 100 次迭代以及 78 次迭代，FISTA 为 38 和 23 次迭代。学习步长的方案没有调整步长方案好，但是较其他基础算法更为优秀。通过分析表 3-1 可知，在 100 次迭代后，FISTA 可以得到 89.75% 和 85.78% 的识别率，但是在牺牲迭代次数以及系统资源的情况下。本章所提调整步长的算法 TsLISTA 在前十次迭代的最佳识别率依旧远远高于 LISTA 与 ISTA，而学习步长的算法 LsLISTA 也获得了跟 LISTA 相近的结果。本章所提算法在 Extend Yale B 数据集得到良好验证。

表 3-1 光照环境下算法在 Extend Yale B 数据集上的识别结果。其中 ISTA 与 FISTA 的识别率为 100 次迭代后的结果，其他算法为 10 次迭代中最好的识别结果

特征空间	识别率 (%)				
	ISTA	FISTA	LISTA	TsLISTA	LsLISTA
120	72.56	89.75	85.12	89.42	85.95
56	72.07	85.78	78.51	81.32	78.01

AR 人脸数据集上的实验结果：为了避免本算法存在由于特定数据集而得到特定结果的局限性，本小节针对 AR 人脸数据集随机特征空间为 120 维进行实验，迭代次数为 10 次，结果如图 3-8 所示。

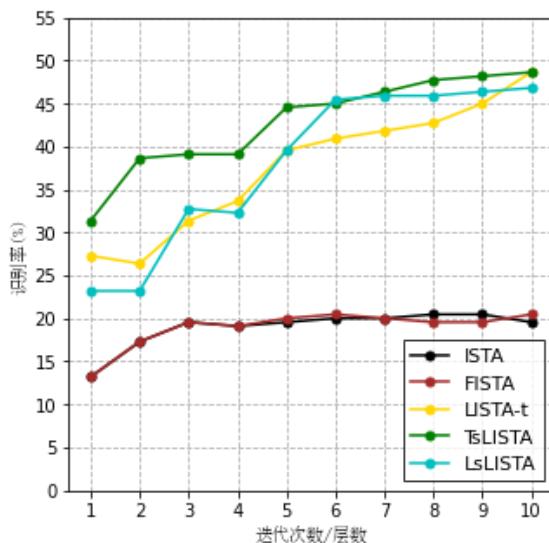


图 3-8 AR 数据集上改进算法关于光照环境的验证

通过分析图 3-8 可知，相比于 120 维随机特征在 Extend Yale B 数据集上的结果，AR 数据库上整体的识别率降低，这是因为 AR 数据库较为贴切真实人脸识别场景，训练人脸以及测试人脸中都对眼镜以及妆容不作限制，这表明环境的复杂程度以及随着字典的基的减少，识别率会降低。此外，传统算法远远不及本文所提的模型结果，总体结果符合实验预期，本文所提模型以及算法表现更优。

3.3.3.2 遮挡及其他条件下的人脸识别

针对遮挡条件下的人脸识别，本文采用 AR 人脸数据集进行验证。求解人脸稀疏表达的 ℓ_1 最小化问题，本小节针对特征空间为 120 维的输入特征进行实验并将迭代次数扩展到 16 次，结果如图 3-9 所示。

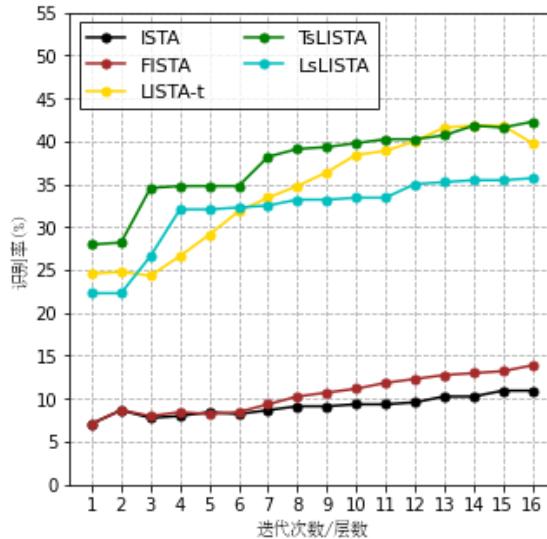


图 3-9 AR 数据集上改进算法关于遮挡及其他环境的验证

从图 3-9 和表 3-2 可知，对于存在遮挡以及光照复杂环境下的人脸识别率整体大大降低，此外随机特征 120 维也对识别效果有所限制。整体来看，本文所提出的采用可训练稀疏编码方式进行人脸识别任务所得到的识别率远比传统算法的结果要好，这种策略得到良好验证。此外，本文改进步长的算法也相比于基准算法 LISTA 要表现更优，改进步长的策略得到验证。

表 3-2 遮挡及其他环境下算法在 AR 数据集上的识别结果。其中 ISTA 与 FISTA 的识别率为 100 次迭代后的结果，其他算法为 16 次迭代后的结果

特征空间	识别率 (%)				
	ISTA	FISTA	LISTA	TsLISTA	LsLISTA
120	17.50	34.55	41.59	42.27	35.68

3.4 本章小结

能否快速求解人脸稀疏表达是基于稀疏表达的人脸识别的关键问题，而基于展开的稀疏编码算法，尤其是以渐进梯度算法作为基准方案的算法中的步长可以作为重要的突破口。本章首先提出采用反向传播的本征分析来调整步长加快人脸稀疏表达，随后，为了节省计算资源消耗，本文提出采用学习步长的策略替代调整步长的方案，优化计算的同时可以得到可比的实验结果。仿真实验的结果表明本章改进步长的算法在信号恢复、重建的实验中表现较优，能够快速的实现图像信号恢复；从人脸识别的实验结果可知，本文所提采用可训练稀疏编码方式替代

传统算法进行人脸稀疏表达能有效加快表达速度以及提高识别概率，这种策略得到有效验证；更进一步，本章所提步长改进的算法在两个数据集以及多种复杂环境中的实验结果都表明其在快速实现人脸稀疏表达方面具有明显优势。

第四章 稀疏表达人脸识别模型简化

4.1 引言

第三章从改进步长的角度研究得到了高效的可训练稀疏编码方法并实现了快速的人脸稀疏表达，其中基于调整步长的人脸识别算法由于 SVD 计算，会产生更多的资源消耗，而采用学习步长的人脸识别算法能够有效的缓解这种资源消耗。为了进一步减少计算以及资源消耗，本章主要针对迭代算法展开成深度神经网络的权重结构的设计进行讨论，并结合第三章所提算法进行权重改进。深度学习网络可以对不同网络层设置不同的权重结构，通常都伴随着大量滤波器的卷积运算等，计算以及存储方面通常耗损巨大，相应训练时间较长。迭代型算法由于其迭代的特性和一般深度学习网络有着不一样的权重设计，迭代算法展开为深度神经网络后，不同迭代之间的权重几乎一致，但是其可以以共享或不共享的方式存在，如何充分利用这一特性提出更加轻便的人脸识别模型也是本文重点研究问题之一。

合理的权重结构能有效减少训练参数，简化网络模型。本章主要从权重结构设计角度出发，构建轻便的稀疏编码人脸识别模型，加快人脸识别编码解码的实现。

4.2 稀疏编码的权重结构讨论

可训练迭代算法的权重设计与可训练参数的数量直接相关，而训练参数对深度神经网络起着关键的作用。对于快速接近“真实”稀疏解决方案具有重要意义。研究^[46]显示，学习共享层权重效率很低。在近似稀疏解中，学习非共享权重比共享权重具有更好的性能。但是，随着迭代（或层次）的增加，它具有太多可训练的参数。为了有效地减少可训练参数并获得相对较好的性能，本文已知一种既存的 tied 权重结构^[64]，在此基础上，提出可训练迭代算法的矩阵权重在每一层模型中都具有独立的可训练参数是非必要的猜想，对这种猜想进行实验验证，并通过实验验证这种既存的绑定权重结构的合理性。

本小节将通过三个部分对可训练迭代算法的权重结构进行全面剖析，验证上文所提猜想以及 tied 权重结构设计的合理性。这三个部分分别为：第一，目前权重结构研究的进展；第二，验证上文所提猜想；第三，结合第三章中学习步长的

方法与 tied 权重结构设计提出新的可训练稀疏编码算法。

4.2.1 基于学习迭代软阈值的权重结构

可训练迭代算法当前主要存在两种不同类型的权重，第一种是共享权重，第二种是非共享权重。共享权重指的是，对于 K 层模型，每个权重只有一个可训练的参数，并且该参数在每一层(或每次迭代)都得到训练。对于非共享的权重，在 K 层模型中，每个权重都有 K 个可训练参数，并且这 K 个参数在相应的层中独立训练。图 2-1 (b) 显示了具有共享权重的 LISTA 的框图。LISTA 的权重最初是作为共享的提出的^[36]，研究^[44]介绍了 LISTA 的非共享的权重图，其结构框图如图 2-1 (c) 所示，研究^[46,67]的分析显示共享权重没有非共享权重有效，进而研究^[46]在其算法 LFISTA 中使用了非共享的权重。Borgerding 和 Schniter^[64]提出 LISTA-t，其将非共享的 LISTA (训练参数为 $\Theta = \{(W_e^{(k)}, S^{(k)}, \theta^{(k)})\}_{k=0}^{K-1}$) 的权重改写为 tied 结构，其训练参数变化为 $\Theta = \{W_e^{(0)}, S^{(0)}, \{\theta^{(k)}\}_{k=0}^{K-1}\}$ ，该方式减少了训练参数并相比于非共享权重下的 LISTA 实现了更优的结果。研究^[44]中提出部分权重耦合结构，但其性能只在一定条件下优越于 LISTA，相对于 LISTA-t 时不太理想。LISTA-t 权重结构使用共享的 $W_e^{(0)}$ 和 $S^{(0)}$ ，意味着在经过 K -层 LISTA-t 模型后，

$W_e^{(0)}$ 和 $S^{(0)}$ 可以被训练得到一个“好”的 $\widetilde{W_e^{(0)}}$ 和 $\widetilde{S^{(0)}}$ ，经由此思路，ALISTA^[55]中则是通过省去训练权重矩阵的方式，直接通过预计解一个标准凸二次方程方式得到“好”的权重矩阵。此种方式能够有效减少训练参数，但是也增加了对这种“好”的权重矩阵的计算。但是以上算法都没有对 tied 权重模型结构进行验证或解析，本文将在下一节针对此问题给出具体的分析与实验验证。

4.2.2 权重结构验证

在 K 层 LISTA 模型中，共享权重具有较少的可训练参数，并且需要较少的资源消耗，但性能不如未共享权重。为了验证本节所提所有权重都采用非共享权重结构是不合理的的猜想，本文尝试在训练 K 层模型时交替使用共享权重和非共享权重。非共享权重在每次迭代中始终具有独立的可训练参数，本文引入一种方案，以使每个第 j 层 (例如 1, 第 j 层) 具有独立的可训练参数，并在接下来的几层 (例如，第 2 … (j-1) 层) 中以共享方式训练这些参数：

$$x^{(k+1)} = \text{SF}(W_e^{(i)} b + S^{(i)} x^{(k)}, \theta^{(k)}), \quad (4-1)$$

K 层模型的训练参数为 $\Theta = \{(P^0, P^1, \dots, P^{K-1}; \theta^0, \theta^1 \dots \theta^{K-1})\}$, 其中 $P^i = \{(W_e^i, S^i) | i = 0, 1 \dots K-1\}$ 为:

$$P^i = \begin{cases} P^i, & \text{如果 } i/j = 0 \\ P^{i-1}, & \text{如果 } i/j \neq 0 \end{cases}, \quad (4-2)$$

其中 j 是训练参数的调节器, $1 \leq j \leq k$ 。当 $j = 1$ 时, 此模型是非共享模型^[44], 每层的可训练参数为, $\Theta = \{(W_e^k, S^k, \theta^k)\}$; 当 $j = k$ 时, 此模型是 LISTA-t^[64], 每层的可训练参数为 $\Theta = \{(W_e^0, S^0, \theta^k)\}$ 。对于不同的 j , 为了方便描述, 本文将其加上 “-j” 来体现不同的 j 的选择, 比如当 $j=1$ 时, 为 LISTA-1。

为了验证 LISTA 中可训练参数的数量的影响, 本文通过实验采用了以下不同的可训练参数集。本文根据一定的规则命名它们, 并根据训练参数的数量的多少列出它们。可训练参数为 $\Theta = \{(W_e, S, \theta)\}$ 的为 LISTA; 每层的可训练参数为等式 (4-2) 的为 LISTA-j; 可训练参数为 $\Theta = \{(W_e^k, S^k, \theta^k)\}_{k=0}^{K-1}$ 被称为不共享的 LISTA 或 LISTA-1。为了验证其可适用性, 本文同时对基于 FISTA 迭代展开的可训练稀疏编码算法 LFISTA^[46]采取相同的措施, 同时为了简化讨论, 本文只列出可训练参数为 $\Theta = \{(W_e^k, W_g^k, W_m^k, \theta^k)\}_{k=0}^{K-1}$ 的为 LFISTA(或 LFISTA-1) 和可训练参数为 $\Theta = \{W_e, \{W_g^k, W_m^k, \theta^k\}_{k=0}^{K-1}\}$ 的 LFISTA-16。在表 1 中, 本文设置 $K = 16$, $j= 1, 2, 4, 8$ 和 16 , 本文总结了 LISTA 不同参数设置的可训练参数的数量。

表 4-1 算法的不同权重设计下的训练参数比较

算法	训练参数
LISTA ^[36]	$O(mn + n^2 + 1)$
LISTA-1 ^[44]	$O(16mn + 16n^2 + 16)$
LISTA-2	$O(8mn + 8n^2 + 16)$
LISTA-4	$O(4mn + 4n^2 + 16)$
LISTA-8	$O((2mn + 2n^2 + 16)$
LISTA-t/LISTA-16 ^[64]	$O(mn + n^2 + 16)$
LFISTA/LFISTA-1 ^[46]	$O(16mn + 16n^2 + 16n^2 + 16)$
LFISTA -t/LFISTA-16	$O(mn + 16n^2 + 16n^2 + 16)$

通过表 4-1 可以得知, 随着 j 的不同设置, LISTA 的可训练参数发生了变化。共享的 LISTA 具有最少数量的可训练参数, LISTA-1 具有最大数量。LFISTA-t

的训练参数相对于 LFISTA 较少。本文通过实验验证了 LISTA 的不同数量的可训练参数。为了简化讨论过程，该实验是不增加噪声的模拟实验，本文的实验也适用于嘈杂的环境。实验设置参照第 3.2.4.1 小节，其中， $\lambda=0.4$, $\text{SNR}=\infty$, 结果如图 4-1 所示：

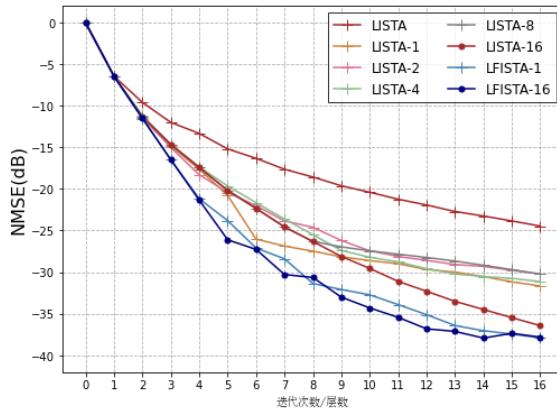


图 4-1 权重结构的验证

通过分析图 4-1 可知，共享权重结构的 LISTA 实验结果较其他结果更差，比较 LISTA 和 LISTA-t，可以得到结论训练参数 θ 的非共享式的算法更加有效，它对调节每层的不同阈值具有重要作用。通过比较 LISTA-j 的不同设置可知，当 $j=8$ 时，可以看到从第八次迭代开始，实验结果逐渐速率降低；当 $j=4$ 时，可以看到从第四次迭代开始，性能稍微下降； $j=2,4,8$ 时显示几乎相同的结果，但是它们具有不同数量的可训练参数； $j=1$ 时，LISTA 是非共享的权重结构，但除了第 6 至 8 层所示结果较好之外，其余结果与其他上述 2, 4, 8 的结果相似；tied 结构的 LISTA，随着迭代次数的增加，性能明显比其他要好；LFISTA-t 展示了同样的结果。至此，本小节可以得出结论，随着迭代层数的增加，在每个单层中使用独立的矩阵权重是没有必要的，这也验证和符合前文所做的的猜想，tied 权重结构被实验性的验证。

4.2.3 绑定的权重结构及其模型

上文中已经验证 tied 权重结构的有效性，因此，本文在第三章所提的算法 LsLISTA 中也采用了这种 tied 结构，这种结构能有效减少可训练参数的数量：

$$x^{(k+1)} = \text{SF}(x^{(k)} + \alpha^{(k)} W_* v^{(k)}, \lambda^{(k)} \alpha^{(k)}), \quad (4-3)$$

其可训练参数变为 $\Theta = \{W_*, \{\alpha^{(k)}, \lambda^{(k)}\}_{k=0}^{K-1}\}$ 。本文将此方案命名为 LsLISTA-t。本文注意到此方案与 TiLISTA^[55]具有相似的可训练参数设置，但是在 TiLISTA 中，其权重矩阵设置为 $W_e^{(k)} = \eta^{(k)} W_e$ ，其中 W_e 初始化为 A^T/L ， $\eta^{(k)}$ 为 1。本文的视角是通过步长改进以及模型简化以实现快速稀疏人脸表达的目标， W_* 初始化为 A^T ， $\alpha^{(k)}$ 初始化为 $1/\sigma_{\max}^2(A^T)$ 。 W_* 是一个在 K 层网络的每一层都得到训练的共享权重矩阵。LsLISTA-t 方法的结构框图如图 4-2 所示。

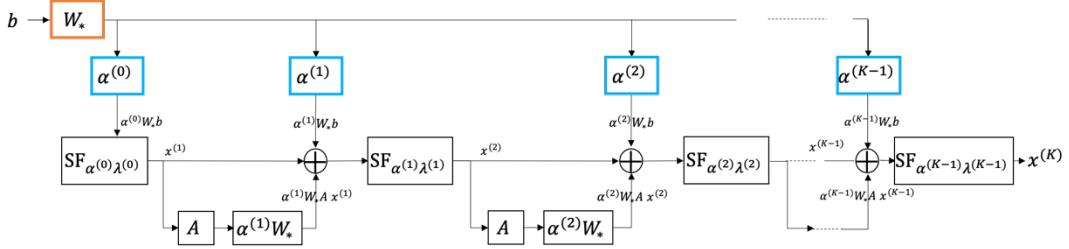
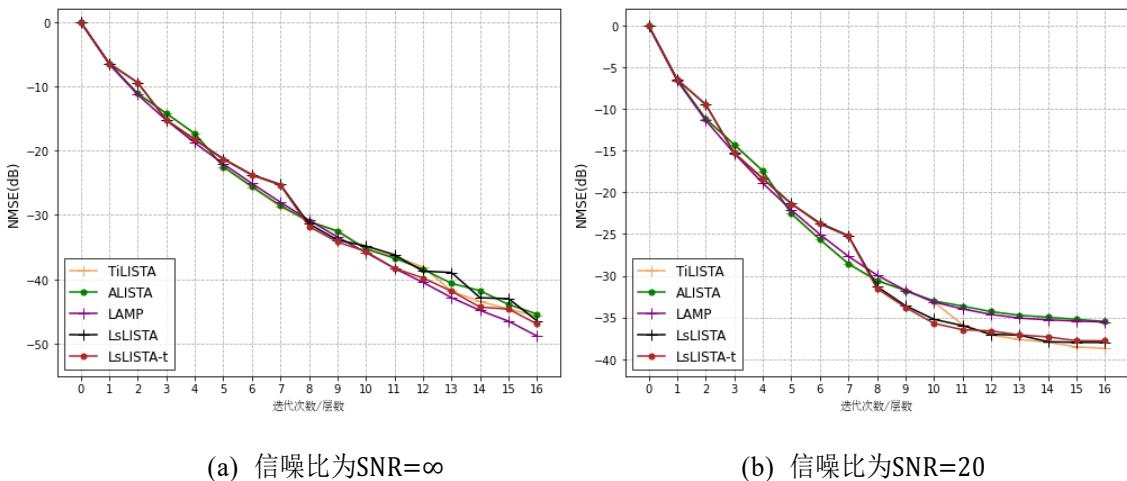
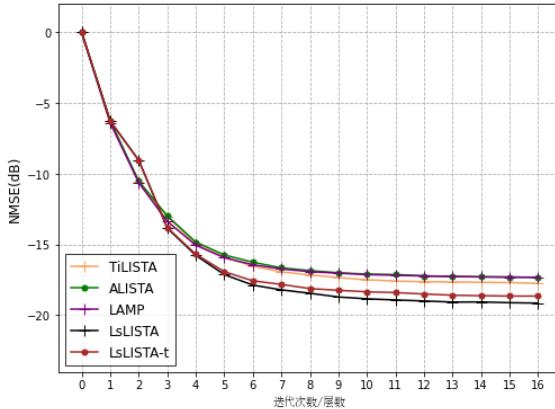


图 4-2 K 层 LsLISTA-t 算法的框图。蓝色方框是学习步长 α 的方案。橙色方框 W_* 是一个共享的权重。学到的步长也用于软收缩算子

4.2.4 仿真实验

本节验证权重改进后的算法，首先对比算法进行说明，第一、由于 LsLISTA 具有与 TiLISTA^[55]相同数量的可训练参数，ALISTA^[55]为其改进算法，因此将 TiLISTA 和 ALISTA 设置为主要对比算法；第二、本小节同样增加 LAMP 算法作为对照实验；第三、为了体现本文所提权重改进后的算法优点，LsLISTA 也为对比算法。实验结果如图 4-3 所示。

(a) 信噪比为 $\text{SNR}=\infty$ (b) 信噪比为 $\text{SNR}=20$



(c) 信噪比为SNR=10

图 4-3 LsLISTA-t 算法的验证

通过分析图 4-3 可知, 对比 TiLISTA、ALISTA、LAMP 和本章所提算法 LsLISTA-t, LsLISTA-t 性能略微要更好, 当信噪比较小时优势更加突出。对比 LsLISTA, LsLISTA-t 性能相差不多, 但注意到 LsLISTA-t 的训练参数更少, 模型更加简单, 本章所提改进算法相对更优。

4.3 绑定的权重与其人脸识别

4.3.1 模型求解

本章对可训练稀疏编码的权重结构进行了系统的讨论, 验证并在本文所提算法 LsLISTA-t。这种权重设计能够大大简化基于可训练稀疏编码的人脸识别的模型复杂度。对于采用 tied 权重的方法进行求解人脸的稀疏表达可以表示为:

$$x^{(k+1)} = \text{SF} \left(x^{(k)} + \alpha^{(k)} \tilde{W}_* (\tilde{b} - R A x^{(k)}), \lambda^{(k)} \alpha^{(k)} \right). \quad (4-4)$$

具体的结构示意图如图 4-4 所示。可以看到采用绑定权重的学习步长的方法减少了可训练参数的数量, 对于大规模的应用比如本文的人脸识别应用来说会大幅度减少计算量以及存储, 模型结构也相对简单。而相比于一般的基于机器学习的方法来说, 其一般都具有比较深层次的网络模型, 其中大多包含了大量的矩阵权重需要进行学习, 而经过本章优化后的人脸识别模型随着模型迭代次数的增长也始终只具有一个矩阵权重需要进行学习, 训练参数大大减少, 模型复杂度也大大降低。

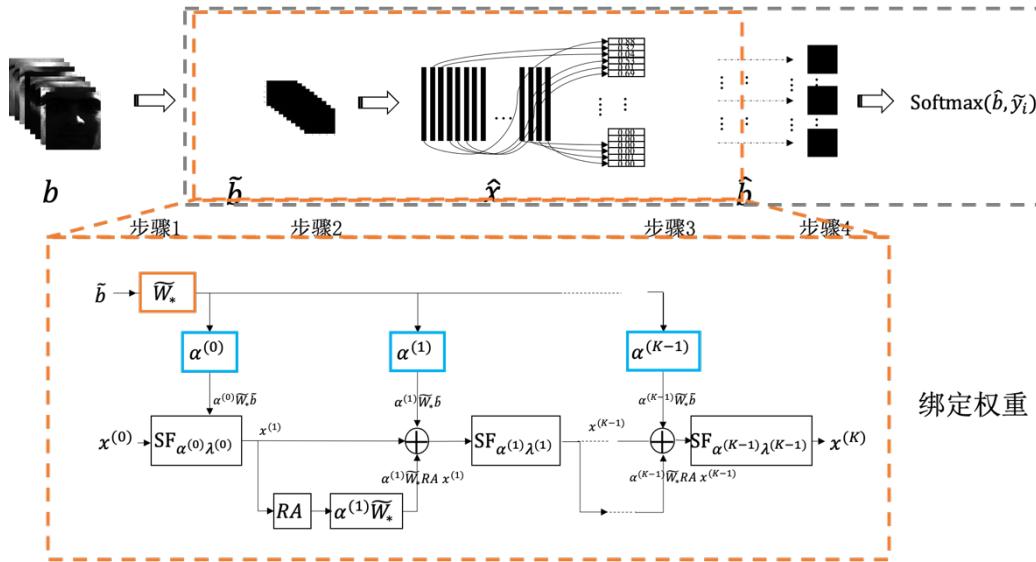


图 4-4 基于 LsLISTA-t 的稀疏编码人脸识别模型框架图。上半部分为流程结构图，其中灰色虚线方框内为采用可迭代算法展开为深度神经网络后的网络结构；下半部分橙色方框内为 LsLISTA-t 的稀疏编码方法

4.3.2 结果分析

求解人脸稀疏表达的 ℓ_1 最小化问题，本章重点验证 tied 权重结构下的稀疏表达人脸识别模型的有效性，因此，主要将采用 tied 模型前后的算法 LsLISTA 以及 LsLISTA-t 进行对比。

4.3.2.1 光照条件下的人脸识别

针对光照条件下的人脸识别，总共分为两个部分来对本小节所提的算法进行验证。第一部分采用 Extend Yale B 数据集对算法进行验证，第二部分采用 AR 人脸数据集对算法进行验证。

Extend Yale B 数据集上的实验结果：本小节针对特征空间为 120 维和 56 维的输入特征进行实验，迭代次数为 10 次，结果如图 4-5 所示。

通过分析图 4-5 可知，对比改进权值前后的算法 LsLISTA 和 LsLISTA-t，权值改进后识别率得到了较大的提升。对比传统稀疏编码算法 100 次迭代与学习的稀疏编码算法 10 次迭代后的最佳识别率（如表 4-2 所示），120 维时，本章所提算法 LsLISTA-t 识别率最优，达到 90.90%，而 56 维时，经过 100 次迭代后的 FISTA 识别效果更优一点，但本章算法表现不差且迭代次数仅需 10 次。对比第三章中所提算法 TsLISTA，本章所提算法更优，所提改进权值的方案得到有效验证。

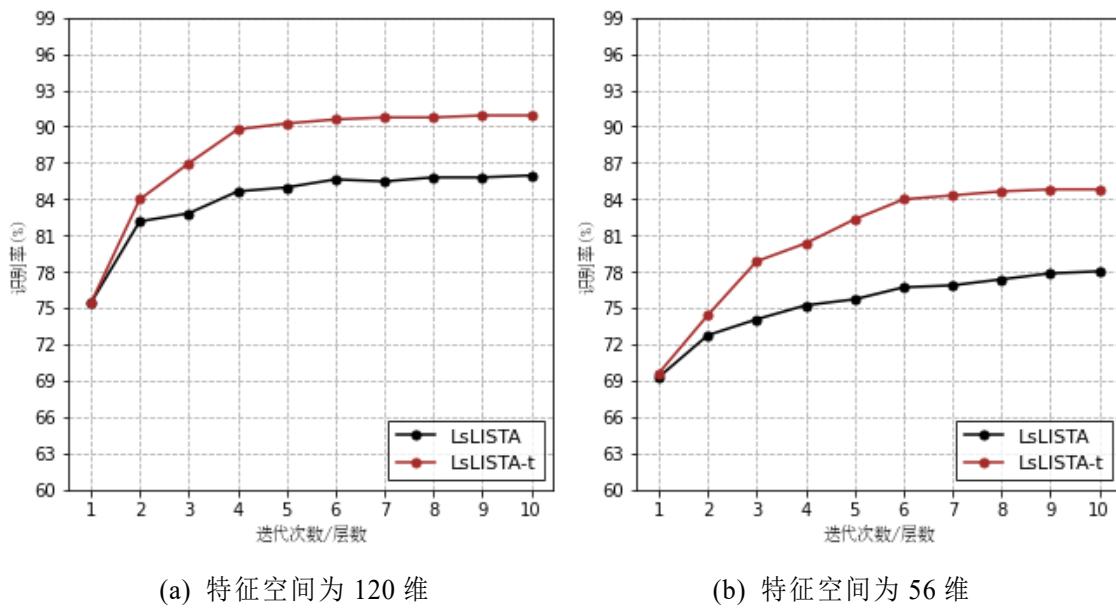


图 4-5 LsLISTA-t 在 Extend Yale B 数据集上的验证

表 4-2 光照条件下算法在 Extend Yale B 数据集上的识别结果。其中 FISTA 的识别率为 100 次迭代后的结果，其他算法为 10 次迭代后的结果

特征空间	识别率 (%)			
	FISTA	TsLISTA	LsLISTA	LsLISTA-t
120	89.75	89.42	85.95	90.90
56	85.78	81.32	78.01	84.79

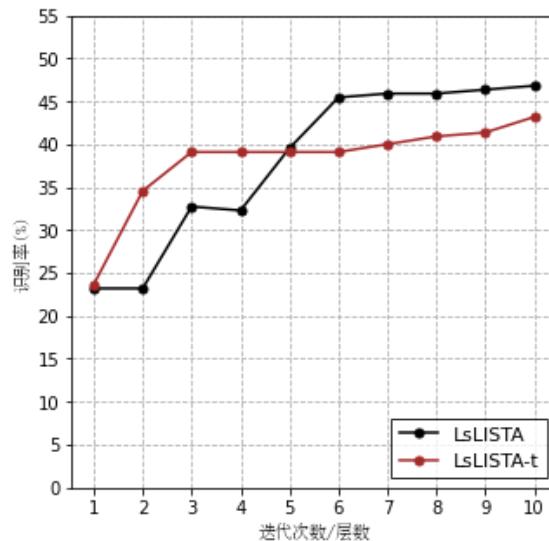


图 4-6 LsLISTA-t 在 AR 人脸数据集上关于光照环境的验证

AR 人脸数据集上的实验结果：本小节针对特征空间为 120 维的输入特征进行实验，从图 4-6 可知，权重改进后的算法在字典基数减少的情况下依旧具有一定优势，可以实现迭代次数更少的情况下，实现更高的识别率。

4.3.2.2 遮挡及其他条件下的人脸识别

针对遮挡条件下的人脸识别，本文采用 AR 人脸数据集进行验证。本小节针对特征空间为 120 维的输入特征进行实验，结果如图 4-7 所示。

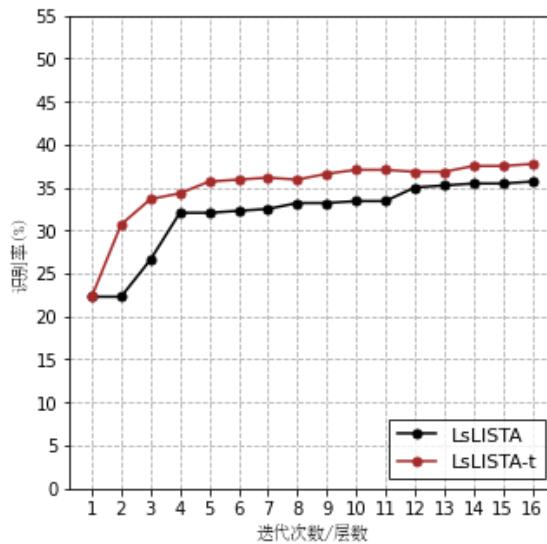


图 4-7 LsLISTA-t 在 AR 人脸数据集上关于遮挡及其他环境的验证

通过分析图 4-7 可知，权重改进的算法在简化模型的同时还提高了识别效率，在遮挡实验中也得到了预期结果。而从表 4-3 可以得到，改进后的算法较传统迭代算法更优，但相比于调整步长的算法 TsLISTA 仍然要更弱，本文将在下一章中继续基于模型简化后的算法 LsLISTA-t 进行改进，在保持简单模型的前提下得到更优结果。

表 4-3 遮挡及其他环境下算法在 AR 数据集上的识别结果。其中 FISTA 的识别率为 100 次迭代后的结果，其他算法为 10 次迭代后的结果

特征空间	识别率 (%)			
	FISTA	TsLISTA	LsLISTA	LsLISTA-t
120	34.55	42.27	35.68	37.73

4.4 本章小结

第三章中设计的模型能够较好的达到人脸识别效果但存在模型较为复杂的缺点，本章主要讨论如何简化基于可训练稀疏编码的人脸识别模型。一方面对现有的权重结构进行介绍与总结；另一方面尝试通过一种交替共享权重结构来验证说明既存 tied 权重结构的合理性，从而说明本文算法权重结构设计的合理性。对比一般的基于机器学习思想的方法，本章所提模型大大减少了训练参数及其所带来的计算以及存储负担。最后采用此种权重结构来重建本文所提的人脸识别模型，并在仿真实验以及两种数据集多种复杂环境下检验重构的简化模型的可行性，实验结果表明该模型具备一定的实用性。

第五章 动能补偿的稀疏表达人脸识别

5.1 引言

第三章和第四章从步长和权重结构角度研究得到了高效的可训练稀疏编码方法来实现快速人脸稀疏表达，主要是从迭代软阈值算法自身固有的算法结构进行考虑。研究^[27,48,49]表明，软阈值迭代算法在每次迭代后得到的稀疏解成分并没有得到充分的求解，通过对稀疏解进行“超调”以及扩大元素来实现更加快速的收敛性能。FISTA 算法中添加的一种“Nesterov”加速度的方法对稀疏解进行作用，但从第三章的结果可知其性能远远不及学习的迭代算法；研究^[49]提供了多种“门”策略来实现补偿，但这些策略都不同程度的使得算法变得更加复杂，为了实现更快收敛，牺牲了时间以及其他资源的耗损。

本章从研究^[44,49]中所提出的补偿稀疏编码估计能力方面考虑，首先验证结合 FISTA 的“动能”设计将 FISTA 组合成为一个可训练的迭代算法。其次，实现稀疏解补偿，文章假设在理想数据中存在一种最优序列的“动能”补偿，提出一种学习“最优”动能的可训练稀疏编码算法。如何设计实现更快收敛速度同时更加轻便的算法是本章稀疏表达人脸识别算法设计的重心。

5.2 稀疏编码与动能补偿

Wu 等人^[49]推测 (L) ISTA 的逐元素更新“落后于”最佳稀疏解，因此建议需要超调量才能达到最佳。该研究通过使用“门”方案对稀疏解作用，添加了额外的方程以扩大所学习的 ISTA 的稀疏估计的元素部分。其在算法中引入了“增益门”和“过冲门”，以实现更好的性能。对于求解 $k+1$ 次迭代后的稀疏解，“增益门”作用在软阈值算子收缩之前的 $x^{(k)}$ 上，而“过冲门”则作用在软阈值算子收缩之后的稀疏解 $\tilde{x}^{(k+1)}$ 和前一个稀疏解 $x^{(k)}$ 上。其“增益门”可以表示为：

$$q^{(k+1)} = \text{SF} \left(W_e^{(k)} b + S^{(k)} \left(\text{gate}_k(x^{(k)}, b | \Lambda_g^{(k)}) \odot x^{(k)} \right), \theta^{(k)} \right), \quad (5-1)$$

其中 $\text{gate}_k(\cdot | \cdot): \mathbb{R}^n, \mathbb{R}^m \rightarrow \mathbb{R}^n$ 输出 n 维向量， $\Lambda_g^{(k)}$ 包含所有的可训练的参数。
 $\text{gate}_k(\cdot | \cdot)$ 被表示为：

$$\text{gate}_k(x^{(k)}, b | \Lambda_g^{(k)}) = 1 + g_k(x^{(k)}, b | \Lambda_g^{(k)}) = 1 + \mu_k \text{th}^{(k-1)} f_k(x^{(k)} | v_k), \quad (5-2)$$

$\mu_k \in \mathbb{R}$ 是一个训练参数, $\text{th}^{(k-1)}$ 是第 $k - 1$ 层的阈值参数, 三种“增益门”函数 $f_k(x^{(k)} | v_k)$ 分别如下:

$$f_k(x^{(k)} | v_k) = \text{ReLU}(1 - \text{ReLU}(v_k | x^{(k)}|)), \quad (5-3)$$

$$f_k(x^{(k)} | v_k) = 1/(v_k | x^{(k)}| + \epsilon), \quad (5-4)$$

$$f_k(x^{(k)} | v_k) = \exp(-v_k | x^{(k)}|), \quad (5-5)$$

其中, $v_k \in \mathbb{R}$ 是一个训练参数, ϵ 是正标量避免出现零为分母的情况, 等式 (5-3) 为分段函数, 等式 (5-4) 为基于反比例的函数以及指数函数对应等式 (5-5)。分段线性函数和指数函数对所有层都有效, 基于反比例的函数仅在较深的层中起作用。因此, 该研究提出组合分段线性函数和基于反比例的函数, 在前 10 层中使用前者, 在 10 层后使用后者。本文认为, 以手动指定不同迭代层次使用不同功能函数的方式使该算法变得复杂, 并且这三种添加的门函数需要更多的资源来进行计算。其“过冲门”策略更进一步使“门”策略成为时间和输入变化的功能:

$$\tilde{x}^{(k+1)} = \text{SF}\left(W_e^{(k)} b + S^{(k)} x^{(k)}, \theta^{(k)}\right), \quad (5-6)$$

$$x^{(k+1)} = \text{og}_k(x^{(k)}, b | \Lambda_o^{(k)}) \odot \tilde{x}^{(k+1)} + \left(1 - \text{og}_k(x^{(k)}, b | \Lambda_o^{(k)})\right) \odot x^{(k)}, \quad (5-7)$$

其中 $\text{og}_k(\cdot | \cdot) : \{\mathbb{R}^n, \mathbb{R}^m\} \rightarrow \mathbb{R}^n$ 输出 n 维向量, $\Lambda_o^{(k)}$ 包含所有的可训练的参数。 $\text{og}_k(\cdot | \cdot)$ 被表示为:

$$\text{og}_k(x^{(k)}, b | \Lambda_o^{(k)}) = 1 + a_0 \sigma(W_o x^{(k)} + S_o b) |\sum_i b_i|, \quad (5-8)$$

$$\text{og}_k(x^{(k)}, b | \Lambda_o^{(k)}) = 1 + \frac{a_0}{|\tilde{x}^{(k+1)} - x^{(k)}| + \epsilon}, \quad (5-9)$$

方程 $\sigma(\cdot)$ 是一个 sigmod 函数, 当方程式为 (5-8) 时, $\Lambda_o^{(k)} = \{a_o, W_o, S_o\}$, 当方程式为 (5-9) 时, $\Lambda_o^{(k)} = \{a_o\}$ 。

本文认为添加额外的方程可能对在一定程度上会增加时间以及计算消耗。此外，这种方式将会使可训练参数的数量在某些设置下增长太多，如所有的“增益门”函数以及“过冲门”中的等式（5-8）的设计。为了适当增益稀疏估计的组成部分，本文倾向于直接扩展 FISTA。本小节首先遵循 LISTA 的想法，使 FISTA 成为一种学习算法，其次本文介绍了针对展开的 FISTA 的学习“最优动能”的方案并本文总结现有学习迭代算法的可训练参数，最后进行实验验证。

5.2.1 可训练迭代算法和 Nesterov 加速度

FISTA 比 ISTA 能更加快速逼近稀疏解，同时学习的迭代算法显示，其性能要远远优于传统迭代算法。如 Wu 等人的研究^[49]表明，稀疏解没有得到充分学习需要对其稀疏元素进行扩大，本文提出是否可以通过简单的添加“动能”来实现充分学习。通过组合深度学习的策略与添加“动能”方法，本小节尝试将 FISTA 的迭代部分展示为深度神经网络。通过将 FISTA 算法与学习的策略结合起来，此处将可训练的 FISTA 分解为两部分，分别为学习的 ISTA 和 FISTA 中描述的“Nesterov”加速度。本文以 LISTA 为例，添加“Nesterov”加速度来进行说明：

$$z^{(1)} = x^{(0)}, t_{(1)} = 1, \quad (5-10)$$

$$x^{(k)} = \text{SF}(W_e b + Sx^{(k)}, \theta), \quad (5-11)$$

$$t_{(k+1)} = \frac{1}{2} \left(1 + \sqrt{1 + 4t_{(k)}^2} \right), \quad (5-12)$$

$$z^{(k+1)} = x^{(k)} + \frac{t_{(k)} - 1}{t_{(k+1)}} (x^{(k)} - x^{(k-1)}). \quad (5-13)$$

同样，可以将等式（5-11）替换成其他学习型的迭代算法，如 LISTA-CP、ALISTA 或本文第三、四章中所提算法等。

5.2.1.1 算法验证

本小节将“Nesterov”加速度分别添加至基础算法 LISTA、LISTA-CP、ALISTA 以及本文第三章所提算法 LsLISTA-t 中，将基础算法设置为对比算法。实验设置遵循第三章中的仿真实验设置，此处仅设置无噪音环境（SNR=∞）和噪音环境（SNR=20）两种情况。

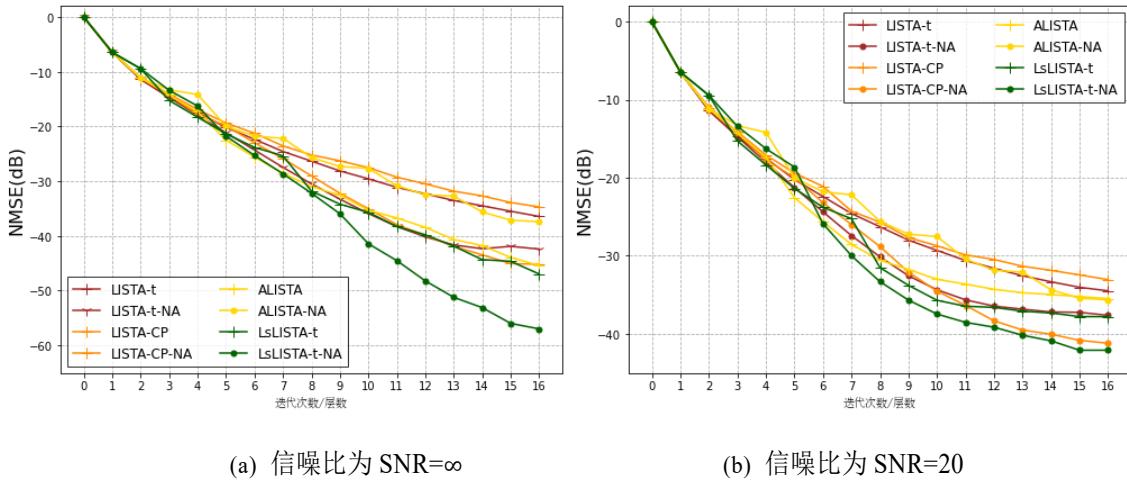


图 5-1 可训练迭代算法和“Nesterov”加速度的验证

通过分析图 5-1 可知，本文对比“-NA”方案与其对应的基础算法，LISTA-t-NA，LISTA-CP-NA 和 LsLISTA-t-NA 比原始方案实现了更优结果，其中 LsLISTA 表现最佳，这验证了上文中的假设，即组合学习的展开迭代算法和“Nesterov”加速度是有效的。但是本文注意到，ALISTA-NA 的性能相比于 ALISTA 反而变得更差，本文推测，对于通过计算得到矩阵 W_e 的 ALISTA 来说，添加的“Nesterov”加速度“太弱”，其中“Nesterov”加速度中的特殊线性表达系数（如等式（5-12）和（5-13）所示）太小无法加速稀疏解，这导致 ALISTA-NA 需要更多的迭代次数或层数才能达到更好的性能。据此，本文提出疑问，即对于图 5-1 中的其他算法来说，“动能”步长是否足够？在训练后是否存在依赖训练数据的“最优加速度”，即等式（5-12）和（5-13）中的特殊线性组合能否被转化为一个可以学习的最优值？如果存在的话，则组合 ALISTA 和这个学习到的“动能”本文可以更加明显的快速逼近稀疏解，得到更优实验结果。

5.2.2 可训练迭代算法与最佳动能补偿模型

FISTA 中添加的“Nesterov 加速度”是通过将特殊的线性变换乘以软收缩操作的最后两个输出之间的差而得到的。特殊的线性变换（等式（5-12）和（5-13））是介于 0 和 1 之间的单调递增序列（图 5-3 中蓝色线段所示）。为了使原始 FISTA 快速收敛，通常经过几十次迭代后，需要将标量 $t_{(k)}$ 重置为 0。上小节本文已经展示了具有“Nesterov 加速度”的学习型迭代算法显示出更好的性能，但这些算法仍在使用“手工方式”来计算添加的功能，本章追求一种轻量级的学习型迭代算法，该算法无需手写程序，具有相对较少数量的可训练参数和简单的方程表达同时能够实现快速的稀疏表达。

回归到本文上小节（5.2.1.1）所提出的“在训练后是否存在依赖训练数据的“最优加速度”，即等式（5-12）和（5-13）中的特殊线性组合能否被转化为一个可以被学习到最优的正尺度？”的猜想，本文通过假设存在依赖于数据的最优值，该最优值有助于扩大训练后的稀疏估计量的组成部分，引入一种将特殊线性变换转换为标量并学习这个标量进而实现学习“动能”效果的策略。本文结合使用第四章所提算法 LsLISTA-t 和学习的标量的策略，将其数学表示为：

$$z^{(1)} = x^{(0)}, \quad (5-16)$$

$$x^{(k)} = \text{SF}(z^{(k)} + \alpha^{(k)} W_*(b - Az^{(k)}), \lambda^{(k)} \alpha^{(k)}), \quad (5-17)$$

$$z^{(k+1)} = x^{(k)} + \gamma(x^{(k)} - x^{(k-1)}), \quad (5-18)$$

其中， $\gamma \in \mathbb{R}$ 是一个从 $(t_{(k)} - 1)/t_{(k+1)}$ 转化而来的标量，可训练参数为 $\Theta = \{W_*, \gamma, \{\alpha^{(k)}, \lambda^{(k)}\}_{k=0}^{K-1}\}$ 。本文将此方案命名为 LM-FISTA。 γ 的初始值为 0， γ 是一个共享标量， $k > 1$ 时开始存在。共享标量 γ 可以帮助本文观察从 1 层模型到 K 层模型的“最佳”加速度趋势。本文推测随着层数或迭代次数的增加，学习的 γ 将会达到介于 0 和 1 之间的稳定“最佳”点。动能补偿同时可以添加到其他基础算法中，比如 LISTA-CP 和 ALISTA 等，本章所提的算法具有一定的扩展性。其算法框架图如图 5-2 所示。

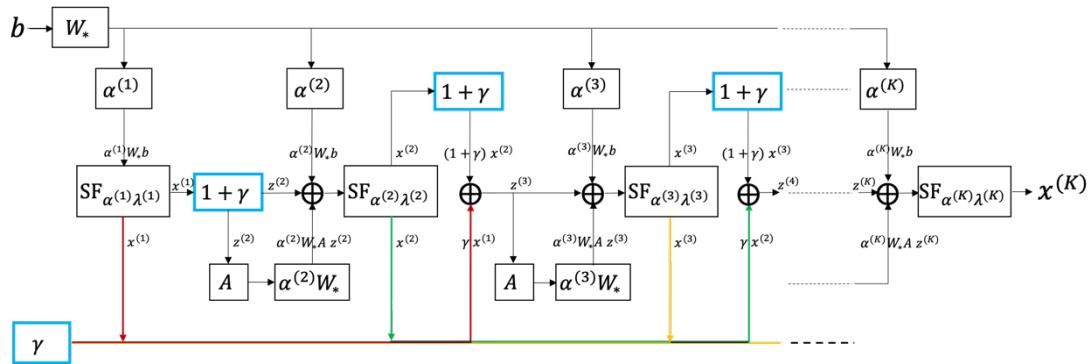


图 5-2 K 层 LM-FISTA 算法的框图。蓝色方框内为学习的标量 γ 的使用， $z^{(1)} = x^{(0)} = \mathbf{0}$ ，红色的指示线为稀疏解 $x^{(1)}$ 乘以标量 γ 后作为加法运算的部分输入，绿色以及黄色的指示线表示同样的步骤，不同颜色的指示线表示该步骤按顺序执行

5.2.2.1 学习标量 γ 的验证

本文首先组合学习的“动能”策略和 LISTA-CP 和 ALISTA，将其命名为 LM-LISTA-CP 和 LM-ALISTA，验证带有“LM”标记的算法中的共享学习标量 γ 经

过 1 层模型至 K 层模型训练学习后的值。其次，本文注意到 LISTA-CP 和 ALISTA 研究中都添加了一种“ss”策略，本文同样采用此种策略。最后，本文添加最初的特殊线性组合 $(t_{(k)} - 1)/t_{(k+1)}$ 来进行对比。

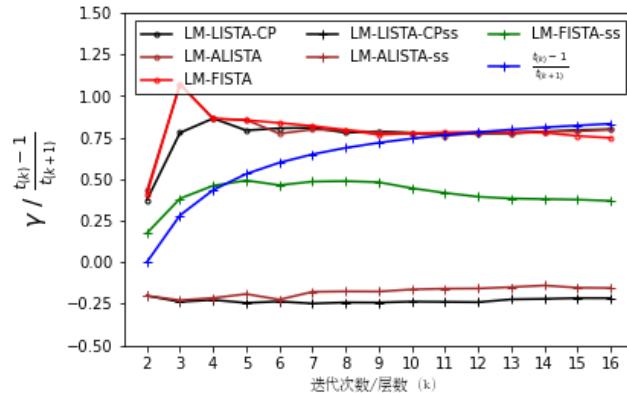


图 5-3 可训练迭代算法和“Nesterov”的验证

根据图 5-3 所示的验证，有“LM-”前标记没有“-ss”后标记的算法（如：LM-FISTA 等）表明，随着迭代次数 k 的增加， γ 趋于稳定并低于 1，据此本文可以得出结论，存在依赖数据最优的标量 γ 可以使得学习“动能”最优。 $(t_{(k)} - 1)/t_{(k+1)}$ 的曲线表明，在经过十层以上的迭代后，其能获得“好”的标量，但获取速度太慢，会导致前期动能补偿不足。然而，从有“LM-”前标记并且“-ss”后标记的算法（如：LM-FISTA-ss 等）可以看出，仅有本文所提算法 LM-FISTA-ss 可以学习到“优质动能”，即训练后的标量 γ 趋于稳定并低于 1。LM-LISTA-CPss 和 LM-ALISTA-ss 的 γ 值趋于稳定但低于 0，低于 0 意味着得到的稀疏解的组成成分已经足够大，甚至还需要再减少才能实现最优，本文推测 LISTA-CPss 和 ALISTA-ss 通过“支持选择”策略已经为稀疏解得到最优的“选择”成分，无法通过学习动能算法进一步提升性能，可以看到“支持选择”策略能够得到良好的性能，但是同时本文注意到这种策略有两个超参数需要被手动调节，本章节所提算法 LM-FISTA 仅仅添加了一个标量参数进行学习，成本较低且性能有明显提升，同时不需要进行手动调节超参数，更加轻便且 γ 的值得到验证，至此，本文所做猜想被直接验证。

5.2.3 算法参数比较

如前所述，本文追求一种轻量级的学习型迭代算法，可训练参数的数量对于评估深度学习模型至关重要。本文在表 5-1 中列出了现存的基于 ISTA 的可训练变体算法的可训练参数。

表 5-1 类 LISTA 算法的不同权重结构的训练参数比较

LISTA-t ^[64]	LISTA-CPss ^[44]	ALISTA-ss ^[45]	TsLISTA	LsLISTA-t	LM-FISTA
$O(mn + n^2 + K)$	$O(Kmn + K)$	$O(K)$	$O(Kmn + K)$	$O(mn + K)$	$O(mn + K)$

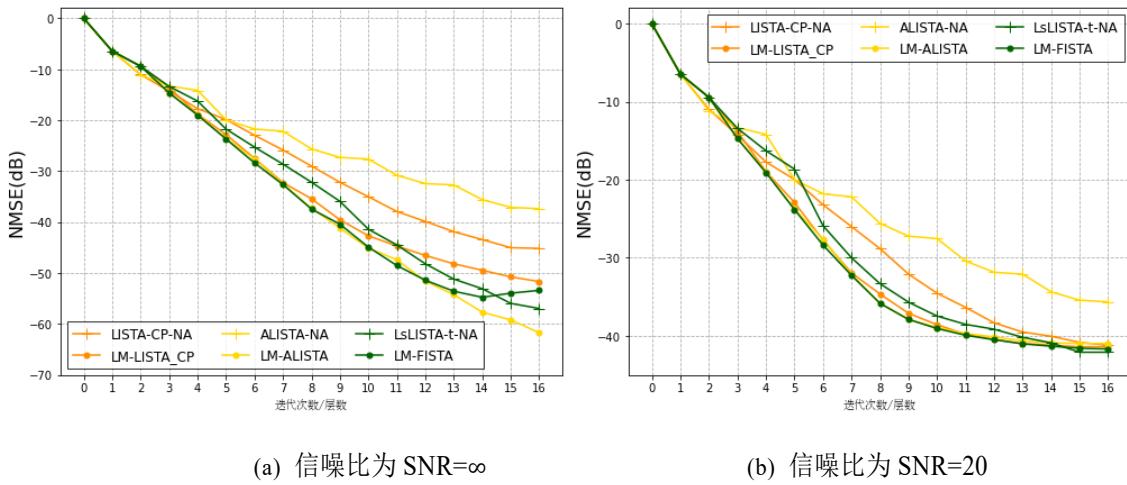
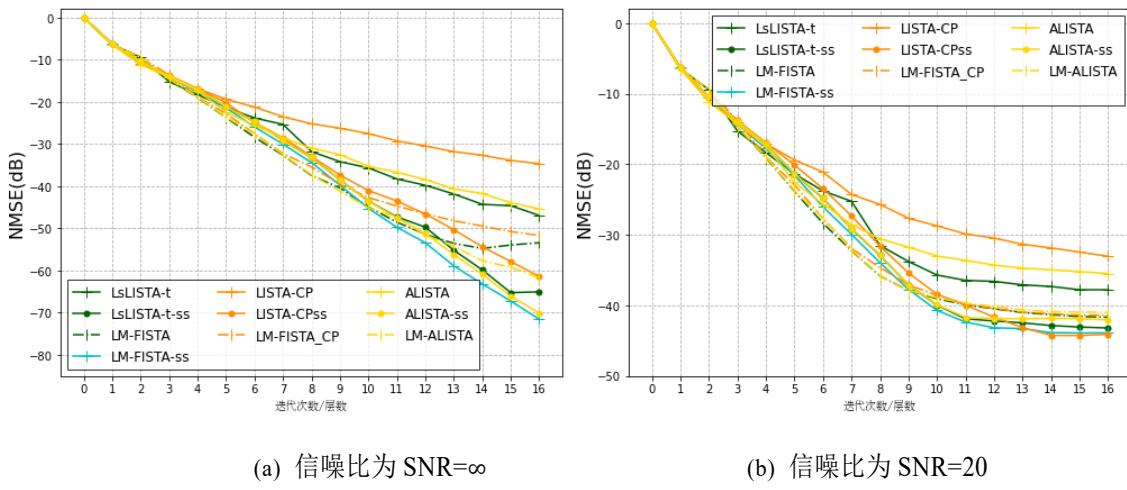
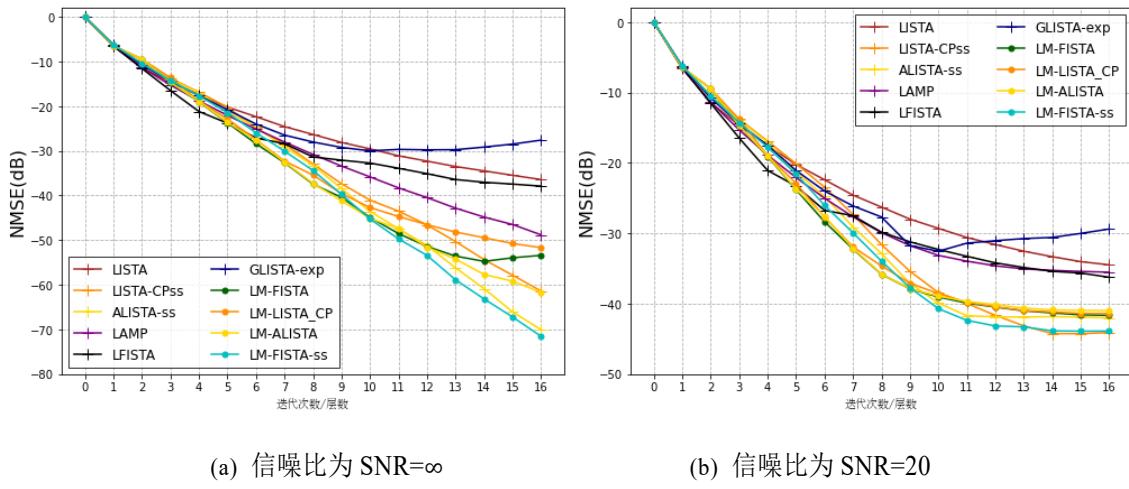
通过表 5-1 可知， ALISTA-ss 具有最少数量的可训练参数， LISTA-CP 和 TsLISTA 具有相同的数量； LsLISTA-t 的数量较 LISTA-CPss 更少。注意到， LFISTA 的训练参数为 $O(Kmn + Kn^2 + Kn^2 + K)$ ， GLISTA 的训练参数则由不同的“门”策略决定，最少为 $O(Kmn + Kn^2 + K)$ 。而 LM-ALISTA 为 $O(K)$ ，基于本章所提策略学习动能的算法仅仅需要多添加一个标量训练参数。

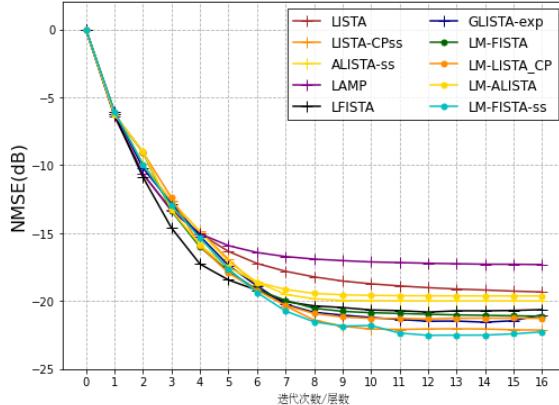
5.2.4 仿真实验

本小节验证基于动能补偿策略的算法。本文将实验分成三组，第一组、将直接添加“Nesterov 加速度”的算法，LISTA-CP-NA、ALISTA-NA 和 LsLISTA-t-NA 与动能补偿后的算法 LM-LISTA-CP、LM-ALISTA 与 LM-FISTA 进行对照实验，验证动能补偿的策略要比直接添加“Nesterov 加速度”要简洁和性能更优，其实验结果如图 5-4 所示；第二组、本小节将动能补偿策略与”ss”策略^[44]进行对比，同时注意到“ss”与动能策略在一定条件下并不互斥，所以本文也将其应用到本章所提算法 LM-FISTA 中，本文用带有“-ss”尾标记的算法表示其添加了“ss”策略，其实验结果如图 5-5 所示；第三组、最后本小节将本文所提算法与当前最为出色的一些工作进行完整的对比，其实验结果如图 5-6 所示。

第一组实验验证结果分析（图 5-4）：对比“-NA”标记的算法与本文所提动能补偿的“LM-”标记的算法，“LM-”标记的算法性能更优，且计算量相对更少，本文所提算法 LM-FISTA 结果最优。

第二组实验验证结果分析（图 5-5）：对比“-ss”标记的算法与其基本算法，本文所提算法 LsLISTA-t 更优；对比所有“-ss”标记的算法，本文算法 LsLISTA-t 相比于 ALISTA-ss 更优，相比于 LISTA-CPss 前 12 层具有明显优势，而 LM-FISTA-ss 从 10 层开始也明显更优；对比“LM-”标记的算法与“-ss”标记的算法，“LM-”标记的算法在前 9 层明显更优，后期性能相对较差，但总体来说具有可比性，“LM-”的算法在计算上面更加简单，计算量相对更少。

图 5-4 第一组验证实验, 信噪比分别为 SNR 为 ∞ 和 20图 5-5 第二组验证实验, 信噪比分别为 SNR 为 ∞ 和 20



(c) 信噪比为 SNR=10

图 5-6 第三组验证实验，信噪比分别为 SNR 为 ∞ 、20 和 10

第三组实验验证结果分析（图 5-6）：对比当前所有算法，可以看到在多种不同噪声设置下随着迭代次数的增加，本章所提算法 LM-FISTA-ss 性能最优，本章所提算法被直接验证。本章所提算法在当前可训练的稀疏编码算法中具有最优性能，能够实现快速的稀疏表达，符合文章所提算法的预期效果。

5.3 动能补偿的稀疏表达人脸识别

5.3.1 模型求解

本章在前文改进的前提下，研究如何在少数迭代次数下更快更优获得人脸的稀疏表达。本章的前半部分通过仿真实验验证了一种学习动能补偿的策略，这种动能补偿策略相比于传统的手动计算的动能补偿来说更加轻便、仅仅只需添加一个学习参数便能够达到比较好的效果。对于采用动能补偿的方法进行求解人脸的稀疏表达可以表示为：

$$z^{(1)} = x^{(0)}, v^{(k)} = \tilde{b} - RAz^{(k)}, \quad (5-16)$$

$$x^{(k)} = SF(z^{(k)} + \alpha^{(k)} \tilde{W}_* v^{(k)}, \lambda^{(k)} \alpha^{(k)}), \quad (5-17)$$

$$z^{(k+1)} = x^{(k)} + \gamma(x^{(k)} - x^{(k-1)}). \quad (5-18)$$

具体的结构示意图如图 5-7 所示。可以看到采用绑定权重的学习步长的方法减少

了可训练参数的数量，对于大规模的应用比如本文的人脸识别应用来说会大幅度减少计算量以及存储，模型结构相对一般的动能补偿的算法更加简单。

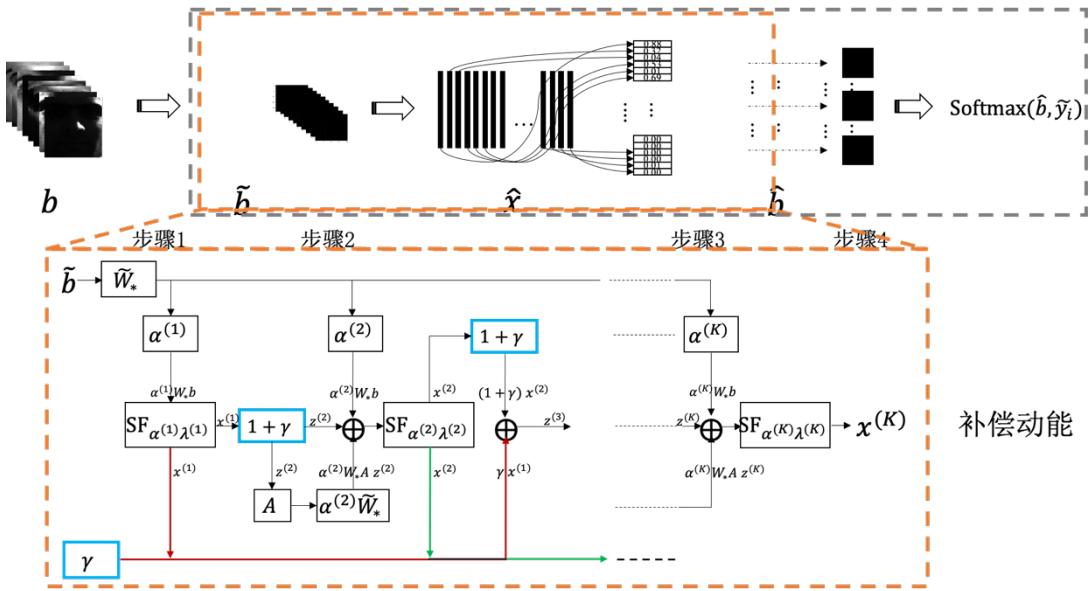


图 5-7 基于 LM-FISTA 的稀疏编码人脸识别模型框架图。上半部分为流程结构图，其中灰色虚线方框内为采用可迭代算法展开为深度神经网络后的网络结构；下半部分橙色方框内为补偿动能的稀疏编码方法

5.3.2 结果分析

5.3.2.1 光照条件下的人脸识别

针对光照条件下的人脸识别，本章主要在 Extend Yale B 数据集上进行验证，本文扩展对随机特征空间的调研，在原有 120 维和 56 维设定上增加 504 维进行进一步分析。求解人脸稀疏表达的 ℓ_1 最小化问题，由于前文已知 FISTA 算法要比 ISTA 更优，所以本章的传统迭代算法的对照算法采用 FISTA。此外对于当前基于学习的迭代算法，本文注意到虽然 ALISTA 算法在当前现有的研究工作中表现最为出色，但是它并不适用于人脸识别等分类任务，因此本文采用其替代算法 TiLISTA 算法以及经典基准算法 LISTA 进行比较，结果如图 5-8 所示。

通过分析图 5-8 可知，本文所提算法 LM-FISTA 在这三种不同特征空间（504、120 以及 56 维）的情况下，分别仅仅只需要一次迭代就可以达到 81.49%、74.88% 以及 69.59% 的识别率。同等条件下，基准算法 LISTA 分别需要 5、3 及 4 次迭代才能达到相同效果，而 TiLISTA 也需要两次左右才能得到与本文所提算法类似的效果，而非学习型的传统算法 FISTA 则需要 30 层以上才能达到比肩效果。同时，本文与研究^[25]进行横向对比，该研究表明，在 Extend Yale B 数据集上验证的识别率需要 40 次迭代以上才能达到较好的结果，而本文仅需 7 次迭代。

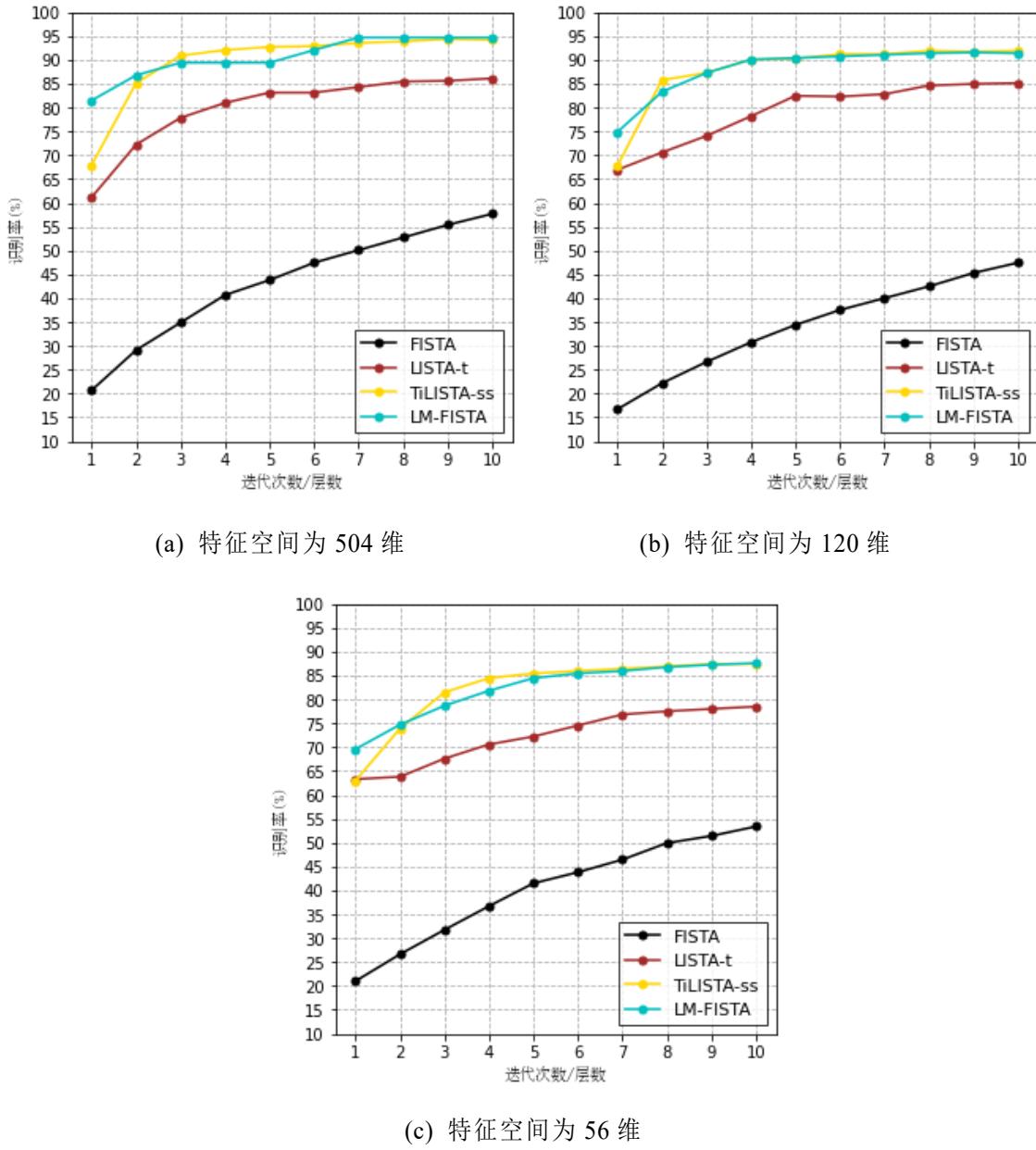


图 5-8 LM-FISTA 在 Extend Yale B 数据集上的验证

表 5-2 光照环境下算法在 Extend Yale B 数据集上的识别结果。其中 FISTA 的识别率为 100 次迭代后的结果，其他算法为 10 次迭代后的结果

特征空间	识别率 (%)			
	FISTA	LISTA-t	TiLISTA	LM-FISTA
504	91.07	86.12	94.38	94.71
120	89.75	85.12	91.90	91.57
56	85.78	78.51	87.43	87.60

通过分析表 5-2 可以得到, 从前十次迭代最好的识别效果看, FISTA 需要 100 次迭代才能到可比的结果, 本文所提算法在 504 维和 56 维的结果中都相对比其他结果更优, 120 维时也能达到和 TiLISTA 相差不远的结果。本文所提算法在 Extend Yale B 数据集上得到良好验证。

5.3.2.2 遮挡及其他条件下的人脸识别

针对遮挡条件下的人脸识别, 本文采用 AR 人脸数据集进行验证。求解人脸稀疏表达的 ℓ_1 最小化问题, 由于前文已知 FISTA 算法要比 ISTA 更优, 所以本章的传统迭代算法的对照算法采用 FISTA。此外对于当前基于学习的迭代算法, 本文采用经典算法 LISTA 以及当前研究中较为出色的 TiLISTA 算法进行比较。

AR 人脸数据集上的实验结果: 本小节针对特征空间为 120 维的输入特征进行实验, 迭代次数为 16 次, 结果如图 5-9 所示。

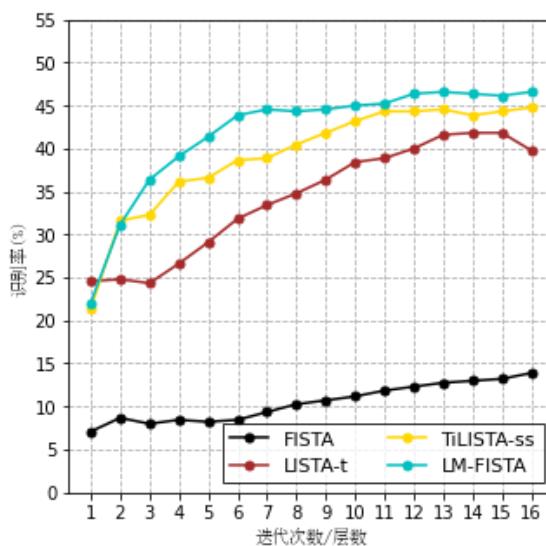


图 5-9 LM-FISTA 在 AR 人脸数据集上关于光照环境的验证

从图 5-9 和表 5-3 可知, 本章所提算法 LM-FISTA 较其他同类型算法在遮挡环境中依旧体现较大优势, 本章所提算法得到良好验证。

表 5-3 遮挡及其他环境下算法在 AR 人脸数据集上的识别结果。其中 FISTA 的识别率为 100 次迭代后的结果, 其他算法为 16 次迭代后的结果

特征空间	识别率 (%)			
	FISTA	LISTA-t	TiLISTA	LM-FISTA
120	34.55	41.82	44.77	46.59

5.4 本章小结

本章首先对基于动能补偿的稀疏编码算法进行调研，进而提出结合传统算法 FISTA 与展开迭代算法成深度神经网络的思想，通过实验结果得到当前稀疏编码算法可能存在动能不足，需要补偿的假设。其次，文章提出一种学习动能的方法来实现“最优”动能增益，更进一步加快人脸稀疏表达。本章将本章所提算法与传统稀疏编码算法、当前最优的可训练稀疏编码算法及以前的研究结果进行比较，设计了多组对比实验，实验结果表明本章所提算法较当前同类型最优算法在图像信号恢复、重建仿真实验中能够实现最精准的信号复原，在人脸识别实验中能够实现较快的人脸识别速率以及较优的人脸识别率，本章所提算法得到良好的验证。

总结与展望

1 论文主要工作总结

基于稀疏表达的人脸识别方法通过简单的随机矩阵实现特征提取，相比传统人脸识别方法以及当前基于机器学习的人脸识别方法在特征提取上有效减少了因输入图像维度较大而产生的计算以及资源消耗，但其通常需要上百次的迭代次数后才能得到较好的人脸稀疏表达。本文的目的是针对这种人脸识别方案进行改进，从而找到一种轻量级且具有成本效益的稀疏表达人脸识别方法，以加快稀疏表达人脸识别的识别速度以及减少资源消耗。本文首先通过步长调整和步长学习的思想再次确认了实时计算“最佳”步长可以有效加快稀疏编码；其次，实验证了 tied 权重结构的合理性并将此运用到本文所提基于学习步长思想的算法中有效的简化了模型；最后，实验性的证明了稀疏编码算法 FISTA 存在数据相关的可学习的“最优”动能，进而采用这种策略在人脸稀疏表达的求解中取得了出色的性能。文章在图像信号恢复仿真实验以及人脸数据集上的验证实验都显示出所提算法能够简化模型以及减少计算资源的消耗。更进一步，本文所提算法不仅可以加快人脸稀疏表达的求解过程，而且本文仿真实验同样表明其对图像信号重建具有较大优势。以下分点对研究内容进行总结：

(1) 针对当前稀疏表达人脸识别方法对于较低维的特征空间输入通常需要数十上百次的迭代才能取得有效结果的缺陷，本文重点考虑具有成本效益的框架模型，提出采用将传统稀疏编码迭代算法展开为深度神经网络的方式对 ℓ_1 最小化问题进行求解，以在给定的迭代次数下产生尽可能低的误差。文章扩展性的将基于稀疏表达的人脸识别转化为数据驱动的机器学习方式，使得本文构建的模型同时具备理论保障和实际效益，并有效证明了可训练的稀疏编码方法能有效加快人脸稀疏表达。

(2) 从获得“最优”合适步长的角度出发以加快求解人脸稀疏表达，本文首先提出训练 ISTA 并调整其步长。本文的策略通过使用反向传播使 ISTA 得到训练，因此它不仅适合训练数据集上的稀疏字典，而且可以提高收敛速度。其次，为了提高有效性并更加适应实时环境，本文提出用一种学习的方法代替了步长调整。改进步长的策略在图像信号重建仿真实验和真实人脸数据集上的结果都证明了其在低维特征空间输入的情况下仍能达到较高的速率，其表明“合适”步长对快速逼近稀疏解有重要作用。

(3) 针对现有基于稀疏表达的人脸模型在一定程度上具有较大的复杂性以及基于深度学习的模型权重很多很冗余的缺点,本文重点考虑构建轻量级的框架模型,在充分调研和验证可训练的稀疏编码的权重特性的前提下,文章提出采用一种 tied 权重设计思想来重新建模稀疏表达人脸识别模型。它能有效帮助简化模型以及实现计算以及资源减负。本文的模型得到了图像信号重建仿真实验和人脸识别实验的验证。

(4) 为了更进一步加快人脸稀疏表达,本文重新调研 FISTA 中引入的动量项,并假设该动能项对算法的加速存在“动能”不足,需要优化补偿的弱点。本文建议学习一个与迭代索引无关的动能项,并实验性验证该动能项最终趋于稳定并得到最优值,其证明了可以通过学习得到稀疏编码算法 FISTA 的“最优”动能实现快速稀疏表达。图像信号重建仿真实验和低维特征空间作为输入的人脸识别实验结果进一步验证了该策略能够快速以及更加精准的实现稀疏表达。

2 本文创新点

本文的研究主要在以下几个方法方面有所创新:

(1) 本文针对人脸稀疏表达过慢的缺点,提出采用可训练的稀疏编码方式进行人脸稀疏表达。据本文所知,本文是首先提出采用可训练的稀疏编码方式来实现人脸识别的方法。

(2) 为了提高人脸稀疏表达能力,本文提出采用改进步长的方案。本文提出在展开的 ISTA 中使用后向传播的特征分解,从而可以调整步长。通过学习步长的方法提高实时人脸识别效率。文章实现了仅仅只需几次迭代就能在低维特征作为输入的模型上获得较好的识别结果。

(3) 为了简化稀疏表达人脸识别模型并更进一步实现快速人脸稀疏表达,本文首先实验性的验证了既存的一种 tied 权重设计,并通过此种方式良好的简化了稀疏表达人脸识别模型。其次,通过对稀疏解决方案提出需要动能补偿假设,进而结合简化的模型学习“最优”加速度,以达到快速求解人脸稀疏表达的预期目标,同时实现了资源以及计算减负。

3 未来工作展望

虽然本文在基于稀疏编码建模人脸识别的研究上已经达到了基本的研究目标,但理论与实际两方面都还需要更加缜密的剖析与拓展:

(1) 首先，本文研究的中心在于采用可训练的稀疏编码的方式来建模基于稀疏表达的人脸模型，对多种复杂环境验证了其高效和有效性，但对于如何加强其在噪声以及遮挡环境下的鲁棒性、如何结合现有的字典学习方案更进一步提高人脸识别率的工作还存在不足，这将是本文未来的研究方向。其次，当前针对对抗样本所产生的人脸识别破解漏洞也将是本文未来的研究内容。

(2) 其次，本文采用可训练的稀疏编码算法替代传统算法进行人脸稀疏表达，通过多个实验验证了其快速的收敛性能，但缺乏相应的数学分析，这也将是本文的重点的未来研究方向之一。

(3) 最后，纵观当前可训练稀疏编码的研究成果，本文可知，许多科研工作者都关注于 LISTA 算法的变体算法的研究。其他一些稀疏编码迭代算法还没得到真正的延伸与拓展，究竟哪一种稀疏编码算法更适用于什么样的领域，性能能否更加优越，还需要更多的研究者来进行学习与探讨。推进可训练稀疏编码算法对其他场景的应用也将是本文未来的重要研究工作方向。

参考文献

- [1] Girshick R. Fast r-cnn[C]//Proceedings of the IEEE international conference on computer vision. 2015: 1440-1448.
- [2] Liu W, Anguelov D, Erhan D, et al. Ssd: Single shot multibox detector[C]//European conference on computer vision. Springer, Cham, 2016: 21-37.
- [3] Zhang L, Lin L, Liang X, et al. Is faster R-CNN doing well for pedestrian detection?[C]//European conference on computer vision. Springer, Cham, 2016: 443-457.
- [4] Shiqi Y, Jia W, Shengyin W, et al. Libfacedetection: a library for face detection in images[EB/OL].(2016-11-01)[2021-4-28].<https://github.com/ShiqiYu/libfacedetection>.
- [5] Seetaface. SeetaFaceEngine[EB/OL]. (2016)[2021-4-28].<https://github.com/seetaface/SeetaFaceEngine>.
- [6] Sun Y, Wang X, Tang X. Deep convolutional network cascade for facial point detection[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2013: 3476-3483.
- [7] Kowalski M, Naruniec J, Trzcinski T. Deep alignment network: A convolutional neural network for robust face alignment[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. 2017: 88-97.
- [8] Parkhi O M, Vedaldi A, Zisserman A. Deep face recognition[J]. Proceedings of the British Machine Vision, 2015:1(3)-6.
- [9] Schroff F, Kalenichenko D, Philbin J. Facenet: A unified embedding for face recognition and clustering[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2015: 815-823.
- [10] Sun Y, Wang X, Tang X. Deep learning face representation from predicting 10,000 classes[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2014: 1891-1898.
- [11] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[J]. arXiv preprint arXiv:1409.1556, 2014.
- [12] Szegedy C, Liu W, Jia Y, et al. Going deeper with convolutions[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2015: 1-9.
- [13] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 770-778.
- [14] Lu C, Tang X. Surpassing human-level face verification performance on LFW with gaussian face[C]//Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence. 2015: 3811-3819.

- [15] Huang G B, Mattar M, Berg T, et al. Labeled faces in the wild: A database for studying face recognition in unconstrained environments[R]. Workshop on faces in 'Real-Life' Images: detection, alignment, and recognition. 2008.
- [16] Grother P J, Ngan M L, Hanaoka K K. Ongoing face recognition vendor test (frvt) part 2: Identification[R]. National Institute of Standards and Technology, 2018.
- [17] Grother P J, Ngan M L, Hanaoka K K. Ongoing face recognition vendor test (frvt) part 1: Verification[R]. National Institute of Standards and Technology, 2018.
- [18] Crumpler W. How accurate are facial recognition systems--and why does it matter[R]. Center for Strategic and International Studies, 2020.
- [19] Facial recognition: top 7 trends (tech, vendors, markets, use cases & latest news) [EB/OL]. (2018-5-22)[2021-04-28].<https://www.thalesgroup.com/en/markets/digital-identity-and-security/government/biometrics/facial-recognition>.
- [20] RealSafe 升级对抗样本防火墙，保障人脸识别更安全. [EB/OL].(2020-2-8)[2021-4-28]
<https://realai.ai/company-trends/info/79>.
- [21] 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.
- [22] Yang M, Zhang L, Yang J, et al. Robust sparse coding for face recognition[C]//CVPR 2011. IEEE, 2011: 625-632.
- [23] 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.
- [24] Deng W, Hu J, Guo J. Extended SRC: Undersampled face recognition via intraclass variant dictionary[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(9): 1864-1870.
- [25] 蔡家柱. 基于稀疏表达的人脸识别算法研究与实现[D]. 成都: 电子科技大学, 2015.
- [26] Daubechies I, Defrise M, De Mol C. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint[J]. Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences, 2004, 57(11):1413-1457.
- [27] Beck A, Teboulle M. A fast iterative shrinkage-thresholding algorithm for linear inverse problems[J]. SIAM journal on imaging sciences, 2009, 2(1): 183-202.
- [28] Donoho D L, Maleki A, Montanari A. Message-passing algorithms for compressed sensing[J]. Proceedings of the National Academy of Sciences, 2009, 106(45): 18914-18919.
- [29] Tropp J A, Gilbert A C. Signal recovery from random measurements via orthogonal matching pursuit[J]. IEEE Transactions on information theory, 2007, 53(12): 4655-4666.
- [30] Needell D, Vershynin R. Uniform uncertainty principle and signal recovery via regularized orthogonal matching pursuit[J]. Foundations of computational mathematics, 2009, 9(3): 317-334.

- [31] Wang J, Kwon S, Shim B. Generalized orthogonal matching pursuit[J]. IEEE Transactions on signal processing, 2012, 60(12): 6202-6216.
- [32] Sprechmann P, Litman R, Ben Yakar T, et al. Supervised sparse analysis and synthesis operators[J]. Advances in Neural Information Processing Systems, 2013, 26: 908-916.
- [33] Sprechmann P, Bronstein A M, Sapiro G. Learning efficient sparse and low rank models[J]. IEEE transactions on pattern analysis and machine intelligence, 2015, 37(9): 1821-1833.
- [34] Borgerding M, Schniter P, Rangan S. AMP-inspired deep networks for sparse linear inverse problems[J]. IEEE Transactions on Signal Processing, 2017, 65(16): 4293-4308.
- [35] Yang M, Liu W, Luo W, et al. Analysis-synthesis dictionary learning for universality-particularity representation based classification[C]//Proceedings of the AAAI Conference on Artificial Intelligence. 2016.
- [36] Gregor K, LeCun Y. Learning fast approximations of sparse coding[C]//Proceedings of the 27th international conference on international conference on machine learning. 2010: 399-406.
- [37] Li Y, Osher S. Coordinate descent optimization for l_1 minimization with application to compressed sensing; a greedy algorithm[J]. Inverse Problems & Imaging, 2009, 3(3): 487.
- [38] Rubinstein R, Elad M. Dictionary learning for analysis-synthesis thresholding[J]. IEEE Transactions on Signal Processing, 2014, 62(22): 5962-5972.
- [39] Hershey J R, Roux J L, Weninger F. Deep unfolding: Model-based inspiration of novel deep architectures[J]. arXiv preprint arXiv:1409.2574, 2014.
- [40] Metzler C A, Mousavi A, Baraniuk R G. Learned D-AMP: Principled neural network based compressive image recovery[J]. arXiv preprint arXiv:1704.06625, 2017.
- [41] Donoho D L, Maleki A, Montanari A. Message-passing algorithms for compressed sensing[J]. Proceedings of the National Academy of Sciences, 2009, 106(45): 18914-18919.
- [42] Donoho D L, Maleki A, Montanari A. Message passing algorithms for compressed sensing: I. motivation and construction[C]//2010 IEEE information theory workshop on information theory (ITW 2010, Cairo). IEEE, 2010: 1-5.
- [43] Wang Z, Yang Y, Chang S, et al. Learning A Deep l_∞ Encoder for Hashing[J]. arXiv preprint arXiv:1604.01475, 2016.
- [44] Chen X, Liu J, Wang Z, et al. Theoretical linear convergence of unfolded ISTA and its practical weights and thresholds[J]. arXiv preprint arXiv:1808.10038, 2018.
- [45] Liu J, Chen X. ALISTA: Analytic weights are as good as learned weights in LISTA[C]//International Conference on Learning Representations (ICLR). 2019.
- [46] Moreau T, Bruna J. Understanding trainable sparse coding via matrix factorization[J]. arXiv preprint arXiv:1609.00285, 2016.
- [47] Kim D, Park D. Element-Wise Adaptive Thresholds for Learned Iterative Shrinkage Thresholding Algorithms[J]. IEEE Access, 2020, 8: 45874-45886.
- [48] Zhou J T, Di K, Du J, et al. Sc2net: Sparse lstms for sparse coding[C]//Proceedings of the AAAI Conference on Artificial Intelligence. 2018.

- [49] Wu K, Guo Y, Li Z, et al. Sparse coding with gated learned ista[C]//International Conference on Learning Representations. 2019.
- [50] Gers F A, Schraudolph N N, Schmidhuber J. Learning precise timing with LSTM recurrent networks[J]. Journal of machine learning research, 2002, 3(Aug): 115-143.
- [51] Xin B, Wang Y, Gao W, et al. Maximal sparsity with deep networks?[J]. arXiv preprint arXiv:1605.01636, 2016.
- [52] Giryes R, Eldar Y C, Bronstein A M, et al. Tradeoffs between convergence speed and reconstruction accuracy in inverse problems[J]. IEEE Transactions on Signal Processing, 2018, 66(7): 1676-1690.
- [53] Srerter H, Giryes R. Learned convolutional sparse coding[C]//2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2018: 2191-2195.
- [54] Tolooshams B, Dey S, Ba D. Scalable convolutional dictionary learning with constrained recurrent sparse auto-encoders[C]//2018 IEEE 28th International Workshop on Machine Learning for Signal Processing (MLSP). IEEE, 2018: 1-6.
- [55] Zhang Jian, Ghanem B. ISTA-Net: Interpretable optimization-inspired deep network for image compressive sensing[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2018: 1828-1837.
- [56] Yang Yan; Sun Jian; Li Huibin, et al. ADMM-CSNet: A deep learning approach for image compressive sensing[J]. IEEE transactions on pattern analysis and machine intelligence, 2018, 42(3): 521-538.
- [57] Hosseini S A H, Yaman B, Moeller S, et al. Dense recurrent neural networks for inverse problems: History-cognizant unrolling of optimization algorithms[J]. arXiv:1912.07197 [physics], Dec. 2019.
- [58] Wu D, Kim K, Dong B, et al. End-to-end lung nodule detection in computed tomography[C]//International workshop on machine learning in medical imaging. Springer, Cham, 2018: 37-45.
- [59] Kemelmacher-Shlizerman I, Seitz S M, Miller D, et al. The megaface benchmark: 1 million faces for recognition at scale[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 4873-4882.
- [60] Tibshirani R. Regression shrinkage and selection via the lasso[J]. Journal of the Royal Statistical Society: Series B (Methodological), 1996, 58(1): 267-288.
- [61] Kavukcuoglu K, Ranzato M A, LeCun Y. Fast inference in sparse coding algorithms with applications to object recognition[J]. arXiv preprint arXiv:1010.3467, 2010.
- [62] Jarrett K, Kavukcuoglu K, Ranzato M A, et al. What is the best multi-stage architecture for object recognition?[C]//2009 IEEE 12th international conference on computer vision. IEEE, 2009: 2146-2153.
- [63] Ablin P, Moreau T, Massias M, et al. Learning step sizes for unfolded sparse coding[C]//Advances in Neural Information Processing Systems. 2019: 13100-13110.

- [64] Borgerding M, Schniter P. Onsager-corrected deep learning for sparse linear inverse problems[C]//2016 IEEE Global Conference on Signal and Information Processing (GlobalSIP). IEEE, 2016: 227-231.
- [65] Georghiades A S, Belhumeur P N, Kriegman D J. From few to many: Illumination cone models for face recognition under variable lighting and pose[J]. IEEE transactions on pattern analysis and machine intelligence, 2001, 23(6): 643-660.
- [66] Martinez A M. The AR face database[R]. CVC Technical Report24, 1998.
- [67] Wang Z, Ling Q, Huang T. Learning deep ℓ_0 encoders[C]//Proceedings of the AAAI Conference on Artificial Intelligence. 2016.

致 谢

饮其流时思其源，成吾学时念吾师，本论文是在导师陈国栋的悉心指导下完成的。导师渊博的专业知识，严谨的治学态度，精益求精的工作作风，诲人不倦的高尚师德，严以律己、宽以待人的崇高风范，朴实无华、平易近人的人格魅力对我影响深远。在硕士课题的每一个阶段，陈老师都对我的课题悉心引导并提出宝贵意见，倾注了导师大量的心血。导师的支持和帮助让我有机会去到国外学习与交流，导师的授业解惑让我对未来有了更加清晰的人生规划，更加坚定了继续科研的信念，谨向导师表示崇高的敬意和衷心的感谢！

在此特别致谢日本长崎大学大学院工学研究科情报应用工学分野酒井智弥副教授在本人 2019-2020 年为期一年的交换留学过程中对本人学业、科研方向的指导，这些指导对我毕业论文后期的算法创新与研究方面有着一定的奠基作用。

本论文的顺利完成，离不开各位老师、同学和朋友的关心和帮助。在此感谢福州大学物理与信息工程学院李建微老师、张金星老师、张红老师、陈星辉老师等对我的关心、支持和帮助。他们在我三年的科研和学习中同样扮演了重要的角色，给予我诸多帮助，让我受益良多。

致谢友人 Matthieu Moreau、王之琪、杨文以及贺晶在学习和生活上对我的支持与关怀。致谢赴日留学期间，酒井研究室的成员以及其他友人对我学业的鼓励，同时让我的留学生活更加充实。

感谢我的室友陈惠斌、实验室同窗张神德、杨怀颖以及研究室学弟学妹等，他们在我三年的科研生活中给予了颇多鼓励与关怀，使我三年的研究生活倍感温暖。

求学多年，最需要感谢是我的家人，感谢他们的无时不刻的精神支持与物质供予，焉得谖草，言树之背，养育之恩，无以回报，他们永远健康快乐是我最大的心愿。

衷心感谢在百忙之中抽出时间审阅本论文的专家教授，感谢答辩委员会的各位老师和专家们对我的论文提出的宝贵建议，为我今后的学习和研究开拓了思路。

个人简历

基本信息：

周筝，女，1995年03月出生，籍贯湖南湘乡。

教育与工作经历：

2012年09—2016年06月 就读于成都大学信息科学与技术学院数字媒体技术专业，获学士学位。

2017年03—2017年10月 就职于盘石软件有限公司（上海），工作岗位：C++软件开发工程师。

2019年09—2020年09月 赴境外交换留学，留学于日本长崎大学大学院工学研究科情报应用工学分野。

2018年09至今 就读于福州大学物理与信息工程学院数字媒体技术专业。

在学期间的研究成果及发表的学术论文

在读期间已发表和录用的论文：

第一作者（1篇）

- 【1】 Preserving Optimal Step Size of unfolded ISTA. 第 23 回画像の認識理解シンポジウム (The 23rd Meeting on Image Recognition and Understanding, 第 23 届图像识别与理解学会). <https://sites.google.com/view/miru2020/>. 2020. (日本学会发表) .

第二作者（1篇）

- 【1】 スパース解法の計算グラフに基づく識別器の深層学習(基于稀疏编码计算图的分类器的深度学习方法) .第 23 回画像の認識理解シンポジウム (The 23rd Meeting on Image Recognition and Understanding, 第 23 届图像识别与理解学会) .<https://sites.google.com/view/miru2020/> . 2020. (日本学会发表) .

参与的科研项目及成果：

- 【1】 深層学習による口腔細胞診および診断に役立つ画像特徴の考察(基于深度学习的口腔细胞学诊断相关的图像特征的考察) .第 38 回日本口腔腫瘍学会総会・学術大会 (第 38 回日本口腔肿瘤学会总会・学术大学) . <https://www.c-linkage.co.jp/jsoo38/> . 2019. (日本学会发表) .
- 【2】 CT 画像における頸部リンパ節の自動検出と口腔がん転移識別(CT 图像上宫颈淋巴结的自动检测和口腔癌转移的识别) . 第 38 回日本口腔腫瘍学会総会・学術大会 (第 38 回日本口腔肿瘤学会总会・学术大学) . <https://www.c-linkage.co.jp/jsoo38/> . 2020. (日本学会发表) .
- 【3】 一种可测水域温度的户外游泳变色救生浮球, 专利号: ZL 2019 2 1677770.4

在校期间所获荣誉：

- 【1】 获得 2018 级学术型硕士研究生中期优秀学业奖学金三等奖;
- 【2】 获得 2019-2020 学年福州大学优秀学生海外访学基金;
- 【3】 获得 2019-2020 学年福州大学优秀共青团干部;
- 【4】 获得 2018-2019 学年福州大学优秀共青团干部;