

学校编码: 10384

分类号 _____ 密级 _____

学 号: 23120121152871

UDC _____

厦 门 大 学

硕 士 学 位 论 文

基于稀疏表示的自适应语音增强方法研究

Adaptive Speech Enhancement Method Based on Sparse
Representation

李 彬

指导教师姓名: 李 琳 副 教 授

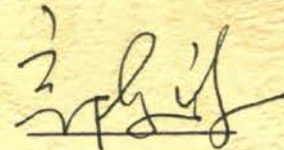
专 业 名 称: 电 路 与 系 统

论文提交日期: 2015 年 4 月

论文答辩时间: 2015 年 5 月

学位授予日期: 2015 年 月

答辩委员会主席:



评 阅 人:

2015 年 月

厦门大学博硕士学位论文摘要库

厦门大学学位论文原创性声明

本人呈交的学位论文是本人在导师指导下,独立完成的研究成果。本人在论文写作中参考其它个人或集体已经发表的研究成果,均在文中以适当方式明确标明,并符合法律规范和《厦门大学研究生学术活动规范(试行)》。

另外,该学位论文为()课题(组)的研究成果,获得()课题(组)经费或实验室的资助,在()实验室完成。(请在以上括号内填写课题或课题组负责人或实验室名称,未有此项声明内容的,可以不作特别声明。)

声明人(签名):



2015年 5月 24日

厦门大学博硕士学位论文摘要库

厦门大学学位论文著作权使用声明

本人同意厦门大学根据《中华人民共和国学位条例暂行实施办法》等规定保留和使用此学位论文，并向主管部门或其指定机构送交学位论文（包括纸质版和电子版），允许学位论文进入厦门大学图书馆及其数据库被查阅、借阅。本人同意厦门大学将学位论文加入全国博士、硕士学位论文共建单位数据库进行检索，将学位论文的标题和摘要汇编出版，采用影印、缩印或者其它方式合理复制学位论文。

本学位论文属于：

1. 经厦门大学保密委员会审查核定的保密学位论文，于 年 月 日解密，解密后适用上述授权。

2. 不保密，适用上述授权。

（请在以上相应括号内打“√”或填上相应内容。保密学位论文应是已经厦门大学保密委员会审定过的学位论文，未经厦门大学保密委员会审定的学位论文均为公开学位论文。此声明栏不填写的，默认为公开学位论文，均适用上述授权。）

声明人（签名）：



2015年 5月 24日

厦门大学博硕士学位论文摘要库

摘要

现代语音通信系统中，大部分的语音信号往往受到背景噪声的干扰，在一定程度上降低了语音信息的可辨性和人耳的听觉舒适性，不利于后继的语音信号处理工作的开展。为了改善语音质量，语音增强技术应运而生，希望在尽可能不引入新噪声的前提下，抑制背景噪声对语音信息的负面影响，提高带噪语音的可辨性。

鉴于语音信号的稀疏先验性，本文基于稀疏表示框架实现对带噪语音的增强处理。通过对字典训练算法和目标优化函数的分析，分别研究了基于 K-SVD (K-Singular Value Decomposition, K-奇异值分解) 自适应稀疏字典的语音增强方法、基于 CNMF (Convolution Nonnegative Matrix Factorization, 卷积非负矩阵分解) 的语音增强方法、基于 BPFA (Beta Process Factor Analysis) 融合模型的自适应语音增强方法和基于鲁棒贝叶斯自适应字典稀疏表示 (Robust Bayesian Dictionary Learning using Spike-slab Prior, RBDL) 模型的语音增强方法。为验证语音增强效果，本文采用信噪比 SNR 和主观评测 PESQ 分数作为评价标准。在 ZOIEUS 语音库上，分别针对携带不同信噪比高斯噪声或有色噪声的带噪语音，采用上述四种不同的语音增强方法 (KSVD、CNMF、BPFA 和 RBDL) 进行降噪处理，由实验统计出的主客观指标结果显示，相较于传统的语音增强方法 (谱减法 and 维纳滤波法)，本文基于稀疏表示实现的四种语音增强方法都能取得较好的降噪性能，其中 BPFA 算法和 RBDL 模型在计算速度上有很大优势，在干净语音及噪声类别未知的情况下能够获得很好的 PESQ 及 SNR 值。

本文的创新点如下：

(1) 基于 CNMF 字典融合策略，提出一种统一参考字典的语音增强算法，利用与待处理带噪语音无关的干净语音训练出统一参考字典，融合各种噪声字典，以解决干净语音未知情况下带噪语音增强的问题。与直接使用相关干净语音训练得到的字典进行 CNMF 语音增强处理相比较，统一参考字典的 CNMF 语音增强方法获得的 SNR 和 PESQ 主客观评价结果有一定程度的降低，但对有色噪声的抑制能力较强，具有较好的实用价值。

(2) 基于非参数贝叶斯概率模型，对字典原子增加 Bernoulli 先验分布及贝塔分布的约束，采用 BPFA 融合模型实现语音增强。不需要人工干预设置模型参数，基于 BPFA 的语音增强方法可根据输入语音信号自适应收敛得到最佳解决策略，相较于其他基于稀疏表示的语音增强方法，具有较快的计算效率和较高的降噪性能，对高斯噪声有很好的增强效果。

(3) 基于 RBDL 模型的语音增强算法，在 BPFA 框架基础上加入残余噪声分量，实验结果表明在高斯噪声处理能力上 RBDL 模型比 BPFA 算法更具有鲁棒性和自适应性，达到语音增强的目的。

关键词：稀疏表示；语音增强；自适应；CNMF；非参数贝叶斯；BPFA；RBDL

ABSTRACT

In Modern voice communication system, most of the speech signal is often affected by background noise, which reduces the differentiability of speech information of the comfort of auditory in some degree, and influences the subsequent speech signal processing. In order to improve the quality of speech, speech enhancement technology arises at the historic moment, in which we hope inhibit the negative impact of background noise on speech information and improve the differentiability of noisy speech under the prediction of introducing none noise.

This paper adopts sparse representation method to realize to the speech enhancement processing of noisy speech based on speech signal sparse apriority. Through the analysis of the dictionary training algorithm and target optimization function, respectively to study the speech enhancement method based on adaptive sparse dictionary K-SVD (K- Singular Value Decomposition), CNMF (Convolution Nonnegative Matrix decomposition) speech enhancement method, the BPFA (Beta Process Factor Analysis) fusion model and RBDL (Robust Bayesian Dictionary Learning using Spike – slab the Prior) model speech enhancement method. To verify the effect of speech enhancement, this paper uses the signal-to-noise ratio SNR and subjective evaluation PESQ scores as the evaluation standard. In ZOIEUS library, to carry different SNR colored Gaussian noise or colored noise, using the above three different speech enhancement method (K-SVD、 NMF、 BPFA and RBDL) for noise reduction processing, subjective and objective indicators by statistical experiment results show that compared with traditional speech enhancement methods (spectral subtraction and wiener filtering method), three kinds of speech enhancement method based on sparse representation can achieve better performance of noise reduction, the BPFA algorithm and RBDL model have great advantages on computing speed, can get high PESQ and SNR scores under the condition of clean speech is unknown.

In this paper, the innovation points are as follows:

(1) This paper puts forward a new algorithm using unified reference dictionary based on CNMF dictionary fusion strategy, gets reference dictionary by training long speech, explore the speech enhancement of unknown clean speech, experiments show that although the SNR and PESQ is lower than clean speech known, but CNMF can effectively restrain colored noise, and has better practical value.

(2) The speech enhancement using sparse prior BPFA fusion model solves the problem of low efficiency in traditional algorithms based on Bernoulli distribution and Beta distribution, and it has a very strong adaptive performance, experiment results show that contrast with K-SVD algorithm, the algorithm of BPFA produces a big promotion in the time efficiency, brings significantly increase with the PESQ values in the speech enhancement of low SNR.

(3) Speech enhancement algorithm based on RBDL model adds residual noise component on framework of BPFA, and the experiment results show that RBDL model is more robust and adaptive than BPFA algorithm on the Gaussian noise processing, so it can achieve the goal of speech enhancement.

Keywords: Sparse Presentation; Speech Enhancement; Adaptive; CNMF; Nonparametric Bayesian; BPFA; RBDL

目录

第一章 绪论	1
1.1 研究背景及其意义	1
1.2 语音增强技术发展	1
1.2.1 语音增强技术发展历程.....	2
1.2.2 语音增强算法分类.....	3
1.3 语音增强的基本理论知识	8
1.3.1 语音信号产生的数字模型.....	8
1.3.2 语音特性和人耳感知特性.....	9
1.3.3 噪声类别.....	10
1.4 语音增强算法的性能评价	11
1.4.1 语音质量的主观评价.....	11
1.4.2 语音质量的客观评价.....	12
1.5 本文主要研究内容与章节安排	14
第二章 稀疏表示的语音增强方法	15
2.1 冗余字典的建立方法	15
2.1.1 固定基原子的选择.....	16
2.1.2 基于数据驱动的字典原子法.....	17
2.2 目标函数优化方法	19
2.2.1 全局优化算法.....	19
2.2.2 贪婪算法.....	20
2.3 K-SVD 学习的语音增强方法.....	22
2.3.1 实验方法.....	22
2.3.2 实验结果.....	23
2.4 本章小结	26
第三章 卷积非负矩阵分解的语音增强算法.....	27
3.1 NMF 算法	27

3.1.1 基础的非负矩阵分解.....	27
3.1.2 改进的非负矩阵分解算法.....	28
3.2 非负矩阵分解的目标函数求解	29
3.2.1 基于欧氏距离的迭代规则.....	30
3.2.2 基于 KL 散度下的迭代规则	31
3.3 卷积非负矩阵分解的语音增强算法	32
3.3.1 CNMF 算法	32
3.3.2 盲语音的 CNMF 语音增强	34
3.4 实验结果及分析	35
3.4.1 W 矩阵中第三维 tt 的取值对结果的影响.....	35
3.4.2 干净语音已知时的 CNMF 语音增强	35
3.4.3 干净语音未知时的 CNMF 语音增强	36
3.5 本章小结	40
第四章 基于非参数贝叶斯自适应字典学习的语音增强.....	41
4.1 非参数贝叶斯方法	41
4.1.1 非参数贝叶斯方法的模型研究.....	42
4.1.2 非参数贝叶斯方法的应用研究.....	43
4.2 基于 BPFA 算法的语音增强研究.....	43
4.2.1 贝叶斯理论.....	43
4.2.2 BPFA 算法.....	46
4.3 基于 RBDL 模型的语音增强研究	48
4.3.1 基于 spike-slab 先验的字典学习	48
4.3.2 RBDL 模型	49
4.4 语音信号重建	50
4.4.1 理想信号重构.....	50
4.4.2 残余噪声项.....	52
4.4.3 高斯噪声项.....	53
4.5 实验比较与分析	53
4.5.1 BPFA 的实验结果及分析.....	53

4.5.2 RBDL 与 BPFA 相比.....	58
4.5.3 四种算法进行比较.....	59
4.6 本章小结	62
第五章 总结与展望	63
5.1 工作总结	63
5.2 工作展望	64
参考文献	65
硕士期间发表的论文	71
致谢.....	73

厦门大学博硕士学位论文摘要库

CONTENTS

Chapter I Introduction	1
1.1 Research Background and Significance.....	1
1.2 Speech Enhancement Technology Development.....	1
1.2.1 Development History of Speech Enhancement	2
1.2.2 Summary of Speech Enhancement Algorithm.....	3
1.3 Basic Knowledge of Speech Enhancement	8
1.3.1 Digital Model of Speech Signal.....	8
1.3.2 Speech Characteristics and Awareness of Human Ear	9
1.3.3 Noise Characteristics	10
1.4 Performance Evaluation of Speech Enhancement.....	11
1.4.1 Subjective Evaluation	11
1.4.2 Objective Evaluation.....	12
1.5 Main Contents and Chapter Arrangement	14
Chapter II Sparse Representation in Speech Enhancement.....	15
2.1 Redundant Dictionary Building Method	15
2.1.1 Sparse Representation in KLT Region	16
2.1.2 Dictionary Atom Method based on Data Driving.....	17
2.2 Objective Function Optimization Method.....	19
2.2.1 Global Optimization Algorithm.....	19
2.2.2 Greedy Algorithm	20
2.3 Speech Enhancement based on K-SVD Study.....	22
2.3.1 Experiment Method	22
2.3.2 Experiment Results	23
2.4 Chapter Summary	26
Chapter III CNMF in Speech Enhancement	27
3.1 Nonnegative Matrix Factorization	27

3.1.1 Fundamental NMF	27
3.1.2 Improved NMF	28
3.2 Redundant Dictionary Building Method	29
3.2 Objective Function of NMF	29
3.2.1 Iteration rules based on Euclidean distance	30
3.2.2 Iteration rules based on KL divergence	31
3.3 Convolution Nonnegative Matrix Factorization in Speech Enhancement	32
3.3.1 Convolution Nonnegative Matrix Factorization	32
3.3.2 Blind Speech Enhancement by CNMF	34
3.4 Experimental Results and Analysis.....	35
3.4.1 Influence on Result of Different t in W	35
3.4.2 Speech Enhancement by CNMF when Speech Known	35
3.4.3 Speech Enhancement by CNMF when Speech Unknown	36
3.5 Chapter Summary	40
Chapter IV Algorithm Based on Nonparametric Bayesian Learning Adaptive Dictionary	41
4.1 Nonparametric Bayesian Method.....	41
4.1.1 Model Research of Nonparametric Bayesian Method	42
4.1.2 Application Research of Nonparametric Bayesian Method.....	43
4.2 Speech Enhancement based on BPFA Algorithm.....	43
4.2.1 Bayesian Theory	43
4.2.2 BPFA(Beta Process Factor Analysis) Algorithm	46
4.3 Speech Enhancement based on RBDL.....	48
4.3.1 Dictionary Learning based on Spike-slab Prior	48
4.3.2 RBDL Model	49
4.4 Reconstruction of Speech Signal	50
4.4.1 Reconstruction of Ideal Signal.....	50
4.4.2 Outliers Noise Item	52
4.4.3 Gaussian Noise Item	53

Degree papers are in the “[Xiamen University Electronic Theses and Dissertations Database](#)”.

Fulltexts are available in the following ways:

1. If your library is a CALIS member libraries, please log on <http://etd.calis.edu.cn/> and submit requests online, or consult the interlibrary loan department in your library.
2. For users of non-CALIS member libraries, please mail to etd@xmu.edu.cn for delivery details.