

语音信号 增强技术及其应用

徐 岩 王春丽 著



科学出版社

语音信号增强技术及其应用

徐 岩 王春丽 著

科学出版社

北 京

内 容 简 介

本书详细介绍了基于短时谱估计、自适应滤波、小波变换、子空间、盲源分离、噪声掩蔽、分数阶傅里叶、分形及神经网络等语音增强算法;通过语音质量评价系统,对语音质量进行了评价,并对语音增强算法及其效果进行了仿真比较;利用 DSP 技术及 OMAP 平台,将语音增强算法应用于无线语音通信系统,实现了较好的语音增强效果,论证了语音增强算法的有效性和可行性。

本书为从事语音增强技术处理的研究者提供了全面而又专业的参考,既可作为本领域研究生和高年级本科生的教学参考书,也可作为数字信号处理、通信以及电子信息相关领域的工程技术人员提供参考,适用于具有一定语音信号处理知识基础的读者。

图书在版编目(CIP)数据

语音信号增强技术及其应用/徐岩,王春丽著. —北京:科学出版社,2013
ISBN 978-7-03-039062-2

I. ①语… II. ①徐…②王… III. ①语音增强—研究 IV. ①TN912.35

中国版本图书馆 CIP 数据核字(2013)第 260730 号

责任编辑:汤 枫 裴 育 / 责任校对:刘亚琦
责任印制:张 倩 / 封面设计:蓝正设计

科 学 出 版 社 出 版

北京东黄城根北街 16 号

邮政编码:100717

<http://www.sciencep.com>

文 林 印 刷 厂 印 刷

科学出版社发行 各地新华书店经销

*

2014 年 1 月第 一 版 开本:B5(720×1000)

2014 年 1 月第一次印刷 印张:21 3/4

字数:423 000

定价:98.00 元

(如有印装质量问题,我社负责调换)

前 言

语音信号增强,简称语音增强,它不仅仅是一个数字信号处理技术方面的问题,还涉及人的听觉感知和语音学,是一项很复杂的技术。语音增强技术是从噪声背景中提取有用的语音信号,抑制与降低噪声干扰的技术,所研究的内容是改进语音质量,尽可能地消除背景噪声、提高信噪比,同时提高语音的自然度和可懂度。由于噪声来源众多,且随应用场合特性各不相同,从而增加了语音增强算法的复杂性,因此不可能设计出一种通用算法来解决所有的噪声问题。针对不同的噪声源,需要采取不同的语音增强算法,因此对各种语音增强技术进行研究是非常有必要的。

目前国内语音信号处理领域的参考书,大多侧重于介绍语音信号处理的基础理论,缺乏专业、系统性介绍语音增强技术的参考书。本书在作者长期从事语音增强研究工作的基础上,分析各种经典语音增强算法原理,根据不同的背景噪声改进部分算法,并结合目前多种新型语音处理技术来实现语音增强,同时通过语音质量评价、软硬件仿真,以及实际系统应用,论证高质量语音增强系统的可行性和有效性。

在参考国内外大量文献的基础上,本书结合作者多年来的专业知识沉淀、项目经验积累,本着系统介绍、重点突出的原则,体现理论性和实用性互补的思想,主要介绍内容包括:各种语音增强算法及其改进算法的原理、语音增强质量评价系统、语音增强软件仿真、语音增强系统在通信中的应用与实现等。本书不仅内容翔实、主次分明,而且与多种通信新技术,如混沌、分形、神经网络等相结合,高效、实时地解决了当代语音通信系统中的信号增强问题。书中图文并茂,算法模块简洁明了,文字解释通俗易懂。

本书系统地介绍语音增强技术及其应用,分为9章,依次从理论基础、算法结构、仿真及应用实现四个层面进行全面介绍。首先,对语音增强技术、语音信号分析处理技术进行描述;其次,介绍几种常用的语音增强技术;然后,介绍语音增强质量评价算法;最后,通过语音增强算法仿真、软硬件实现方法,对语音增强系统进行整体测试与性能评估。

本书的主要内容是作者与课题组同事及研究生长期科研工作的成果总结。在本书的撰写过程中,杨桂琴老师给予了热情的帮助,并提出了修改意见;研究生李洋洋负责资料收集、绘图、公式编辑等工作,对本书作出了贡献;研究生孟静、刘馨、唐建云、谭方、曹玉萍、闫亮也给予了热心的帮助。在此向他们表示衷心的感谢!

尽管作者在撰写过程中尽了最大努力,但受水平、学识和经验所限,书中难免会存在不妥和疏漏之处,衷心希望广大读者予以批评指正。

作 者

xuyan@mail.lzjtu.cn

2013 年 10 月

目 录

前言

第 1 章 语音增强技术概述	1
1.1 语音增强研究背景	1
1.2 语音信号与语音增强	2
1.2.1 语音信号特征	2
1.2.2 语音信号信息量	3
1.2.3 噪声特征及其分类	4
1.2.4 人耳感知特性	7
1.2.5 语音增强的信号模型	8
1.3 语音增强技术的发展	10
1.4 语音增强方法分类	11
1.5 语音增强效果评价	13
1.6 语音增强技术应用	14
1.7 本书主要内容	16
参考文献	17
第 2 章 语音信号分析处理技术	19
2.1 语音增强预处理技术	19
2.1.1 语音信号预滤波和数字化	20
2.1.2 语音信号预加重	22
2.1.3 语音信号加窗处理	23
2.2 语音增强时域分析处理技术	28
2.2.1 短时能量及短时平均幅度分析	29
2.2.2 短时平均过零率分析	29
2.2.3 短时自相关分析	30
2.3 语音增强频域分析处理技术	32
2.3.1 短时傅里叶变换分析	32
2.3.2 短时傅里叶逆变换分析	35
2.4 语音增强同态分析处理技术	38
2.4.1 同态处理	38
2.4.2 复倒谱及倒谱	39

2.4.3 复倒谱分析	40
2.5 语音增强线性预测分析处理技术	41
2.5.1 线性预测分析	42
2.5.2 线性预测方程组	43
2.5.3 线性预测等价参数	45
2.6 基于非线性理论的语音分析处理技术	46
2.6.1 基于混沌理论的语音分析处理技术	46
2.6.2 基于分形理论的语音分析处理技术	48
2.6.3 基于神经网络的语音分析处理技术	50
2.7 语音增强噪声估计技术	53
2.7.1 基于平稳环境下的噪声估计	53
2.7.2 基于非平稳环境下的噪声估计	55
2.8 本章小结	57
参考文献	57
第3章 语音增强短时谱估计算法	59
3.1 谱相减算法	59
3.1.1 幅度谱减法	59
3.1.2 改进的幅度谱减法	61
3.1.3 功率谱减法	63
3.1.4 改进的功率谱减法	65
3.2 维纳滤波算法	67
3.2.1 维纳滤波法时域实现	67
3.2.2 维纳滤波法频域实现	69
3.2.3 改进的维纳滤波法	71
3.2.4 卡尔曼滤波法	73
3.3 最小均方误差算法	76
3.3.1 基本型最小均方误差法	76
3.3.2 对数谱最小均方误差法	78
3.4 本章小结	80
参考文献	81
第4章 语音增强自适应滤波算法	83
4.1 自适应滤波	83
4.1.1 自适应滤波算法	83
4.1.2 自适应滤波器的性能指标	84
4.1.3 最佳滤波准则	85

4.2 最速下降自适应滤波·····	86
4.2.1 最速下降算法·····	86
4.2.2 最速下降自适应滤波器的性能指标·····	88
4.3 最小均方自适应滤波·····	91
4.3.1 最小均方算法·····	91
4.3.2 归一化最小均方算法·····	95
4.3.3 最小均方滤波器的性能指标·····	95
4.4 最小二乘自适应滤波·····	98
4.4.1 最小二乘自适应滤波算法·····	98
4.4.2 递归最小二乘自适应滤波算法·····	99
4.4.3 最小二乘滤波器的性能指标·····	100
4.5 自适应滤波算法的改进·····	100
4.5.1 自适应滤波算法的时域改进·····	100
4.5.2 自适应滤波算法的频域改进·····	104
4.6 本章小结·····	110
参考文献·····	111
第5章 语音增强小波变换算法·····	114
5.1 小波变换分析·····	115
5.1.1 连续小波变换·····	115
5.1.2 离散小波变换·····	117
5.1.3 多分辨率分析与 Mallat 算法·····	119
5.1.4 最优小波基·····	123
5.2 小波域语音信号增强·····	125
5.2.1 小波域信号增强·····	125
5.2.2 常用小波函数·····	127
5.2.3 语音增强中小波函数选取·····	131
5.3 小波阈值去噪法·····	132
5.3.1 小波阈值去噪算法原理·····	132
5.3.2 改进的阈值函数去噪法·····	137
5.4 小波模极大值去噪法·····	141
5.4.1 信号与噪声在小波变换各尺度上的不同传播特性·····	141
5.4.2 小波模极大值去噪算法原理·····	143
5.5 小波掩蔽去噪法·····	144
5.5.1 小波掩蔽去噪算法原理·····	144
5.5.2 改进型掩蔽去噪法·····	145

5.6 各种小波去噪法比较	148
5.7 本章小结	149
参考文献	150
第6章 语音增强其他优选算法	152
6.1 基于信号子空间的语音增强算法	152
6.1.1 信号子空间单通道语音增强算法	152
6.1.2 信号子空间多通道语音增强算法	154
6.2 基于盲源分离的语音增强算法	155
6.2.1 信号盲源分离	155
6.2.2 语音增强中的盲源分离	158
6.3 基于听觉掩蔽效应的语音增强算法	162
6.3.1 噪声掩蔽阈值	163
6.3.2 语音增强中的掩蔽效应	165
6.4 基于分数阶傅里叶变换的语音增强算法	167
6.4.1 分数阶傅里叶变换算法	167
6.4.2 基于分数阶域的谱减法语音增强	172
6.4.3 离散分数余弦变换自适应滤波算法	175
6.5 基于分形理论的语音增强算法	179
6.5.1 分形理论	179
6.5.2 语音增强中的分形理论	180
6.6 基于神经网络的语音增强算法	183
6.6.1 神经网络	183
6.6.2 语音增强中反向传播神经网络	184
6.6.3 语音增强中小波神经网络自适应滤波	186
6.7 本章小结	188
参考文献	189
第7章 语音增强质量评价	192
7.1 语音质量评价	192
7.1.1 听觉系统	192
7.1.2 语音质量	194
7.1.3 语音质量评价方法	195
7.2 语音质量主观评价	195
7.3 语音质量客观评价	197
7.3.1 客观评价系统	197
7.3.2 客观评价测度	198

7.3.3 客观评价算法	207
7.4 语音质量评价算法	212
7.4.1 语音质量评价算法的实现	213
7.4.2 基于听觉模型的客观评价算法	214
7.4.3 感知语音质量评价算法	220
7.4.4 主客观评价方法的相关度	228
7.5 本章小结	229
参考文献	229
第8章 语音增强算法仿真	231
8.1 语音信号处理与仿真软件	231
8.1.1 语音编辑	232
8.1.2 语谱图生成	233
8.1.3 语音增强仿真工具	235
8.1.4 语音增强仿真准备	236
8.2 语音增强算法仿真	237
8.2.1 高斯白噪声仿真实验	237
8.2.2 粉红噪声仿真实验	243
8.2.3 工厂噪声仿真实验	247
8.2.4 算法仿真性能分析	253
8.3 熵函数最优小波基选取仿真	254
8.3.1 Shannon 熵最优小波基选取仿真实验	255
8.3.2 SURE 熵最优小波基选取仿真实验	264
8.3.3 threshold 熵最优小波基选取仿真实验	269
8.3.4 算法仿真性能分析	272
8.4 小波阈值计算仿真	275
8.4.1 阈值函数的选取	275
8.4.2 阈值函数中调节因子及阈值选取	275
8.4.3 算法仿真及结果分析	277
8.5 语音增强质量评价算法仿真	280
8.5.1 分段信噪比仿真	280
8.5.2 语音感知质量评价算法仿真	284
8.5.3 算法仿真性能分析	287
8.6 本章小结	290
参考文献	290

第9章 语音增强系统设计与应用	292
9.1 基于 TMS320C6416 的语音增强系统硬件设计与实现	293
9.1.1 DSP 处理技术	293
9.1.2 基于 TMS320C6416 的语音增强系统硬件结构设计	295
9.1.3 基于 TMS320C6416 的语音增强系统工作原理.....	303
9.2 基于 TMS320C6416 的语音增强系统软件设计与实现	307
9.2.1 语音增强系统软件设计	307
9.2.2 基于 TMS320C6416 的语音增强系统软件实现.....	308
9.2.3 基于 TMS320C6416 的 FFT 算法软件实现	313
9.3 基于 TMS320C6416 的语音增强系统性能测试	317
9.4 基于 OMAP3 平台的语音通信增强系统设计	318
9.4.1 OMAP 概述	318
9.4.2 OMAP3 体系结构	320
9.4.3 OMAP3 软件开发平台的构建.....	322
9.4.4 基于 OMAP3 的无线语音通信系统设计	326
9.5 本章小结	336
参考文献.....	336

第 1 章 语音增强技术概述

语音是语言的声学表现,它不仅是人类交流信息最自然、最有效、最方便的手段,而且也是人类进行思维的一种依托。随着通信技术的发展,语音作为一种典型的非平稳随机信号,已经成为人们日常生活、工作中不可缺少的一部分。在语音通信过程中,不可避免地会受到来自周围环境噪声和设备内部噪声等的各种干扰,由于噪声的存在会使语音处理系统的性能恶化,因此语音增强是解决噪声污染的一种有效方法。语音增强的目的就是 from 被污染的语音信号中,提取尽可能纯净的语音信号,改善语音质量,使听者不觉得疲劳,同时提高语音的可懂度。

1.1 语音增强研究背景

语音增强早在 20 世纪 60 年代就受到了人们的重视,在随后的四十多年里,很多学者对这一课题进行了研究。1978 年, Lim 和 Oppenheim 提出了语音增强的维纳滤波方法;1979 年, Boll 提出了谱相减方法来抑制噪声;1980 年, Maulay 和 Malpass 提出了软判决噪声抑制方法;1984 年, Ephraim 和 Malah 提出了基于最小均方误差短时谱幅度估计的语音增强方法。20 世纪 80 年代以后,随着高速 DSP 的发展,语音增强逐渐走向实用,同时新的语音增强方法又相继涌现,常见的语音增强方法主要有:基于最小均方(least mean square, LMS)自适应滤波的噪声抵消语音增强法、基于短时谱(short time spectrum, STS)估计的语音增强法、基于语音生成模型的语音增强法、基于梳状滤波器的谐波增强法、基于阵列话筒的语音增强法、基于听觉模型的语音增强法、基于人工神经网络的语音增强法、基于信号子空间分解的语音增强法、基于小波变换的语音增强法等。语音增强是一门涉及面很广的交叉学科,它不但与语音信号数字处理理论有关,而且涉及模式识别、数理统计、语音学等。此外,语音增强所面临的噪声形式也可能众多,因此要有效地增强语音,必须对语音和噪声特性有充分的了解。

语音信号处理技术可分为四个主要研究领域,分别是语音编码和压缩技术、语音识别技术、语音合成技术及语音增强技术。而语音增强技术可以广泛应用于语音通信领域、语音识别和语音编码系统中。尤其是近年来,语音识别技术获得突破性进展,作为人机交互的一种很自然、便捷的方式,越来越受到人们的重视。特别是在 2008 年北京奥运会的成功举办中,移动式语音识别和翻译系统发挥了重要作用。但是目前的语音识别系统大多工作在安静的环境下,背景噪声的引入会严重

影响识别系统的性能。在低速率语音编码系统中,以较低的码速率传输语音信号能够节省大量传输带宽,而噪声的存在严重影响模型参数的提取,使得重建语音的质量急剧恶化,甚至变得完全不可懂。但通过语音增强,输入声码器信号的抗噪能力显著提高。在上述情况下,加入语音增强系统,可抑制背景噪声,提高语音通信质量;也可将语音增强系统作为预处理器,提高语音处理系统的抗干扰能力,稳定系统性能。

语音增强是从噪声背景中提取有用的语音信号,抑制、降低噪声的干扰。语音增强的目的主要是改进语音质量,尽可能地消除背景噪声,提高信噪比(signal to noise ratio, SNR),同时提高语音自然可懂度和说话人的可辨度。噪声来源取决于实际的应用环境,因而噪声特性可以说变化无穷,所以在实际应用时,要根据具体的噪声情况和特定环境,选用不同的语音增强方法,才能达到最好的语音增强效果。

随着数字信号理论的成熟,语音增强技术已发展成为语音信号处理的一个重要分支。近年来,随着计算机和 DSP 技术的发展和成熟,语音增强的实时实现成为可能,可广泛应用于无线电话会议、手机、娱乐系统、多媒体应用、智能家电、场景录音和军事窃听等领域。在实际应用中,这些系统在其前端加上语音增强系统作为预处理器,以提高其抗干扰能力,稳定其系统性能。因此语音增强技术可以广泛应用于各种语音信号处理领域中,本书所研究的语音增强技术理论、应用与实现有着重要的现实意义。

1.2 语音信号与语音增强

1.2.1 语音信号特征

由于语音的生成过程与发音器官的运动过程密切相关,而且人类发音系统在产生不同语音时的生理结构并不相同,因此使得产生的语音信号是一种时变的、非平稳的随机信号,例如,声道面积随着时间和距离而改变,气流速率随着声门处压力变化而变化等。但是由于人类发声器官变化速率具有一定的限度,而且远小于语音信号的变化速率,可以认为人的声带、声道等特征在一定的时间内基本不变。因此假定语音信号是短时平稳的,即语音信号的某些物理特性和频谱特性在 10~30ms 的时间段内是近似不变的,具有相对的稳定性,这样就可以运用分析平稳随机过程的方法来分析和处理语音信号,在语音增强中正是利用了语音信号短时谱的平稳性。

语音是由人的发音器官发出来的、具有一定语义的声音。语音中的元音是在发音过程中,气流通过口腔不受阻碍发出的音,每个元音的特点是由声道的形状和尺寸决定的。语音中的辅音是指在发音的时候,从肺里出来的气,经过口腔或者鼻

腔时受到阻碍而形成的音,根据声带是否振动又可分为清辅音与浊辅音两种:如果声带不振动,发出的辅音叫做清辅音,简称清音,在汉语音学中也叫噪声;声带振动发出的辅音叫做浊辅音,也叫浊音。在语音信号处理中基本上就分为清音和浊音两大类。清音和浊音在特性上有明显的区别,清音没有明显的时域和频域特性,看上去类似于白噪声,并具有较弱的振幅;而浊音在短时谱上有明显的特征,具有以下两个特点:

(1) 在时域上呈现出明显的周期性,这是因为浊音的激励源为周期脉冲气流。

(2) 频谱中有明显的几个凸起点,它们的出现频率与声道的谐振频率相对应。这些凸起点称为共振峰,其频率称为共振峰频率。共振峰按频率由低到高排列为第一共振峰、第二共振峰,依次类推。

在语音增强中可以利用浊音所具有的明显的周期性来区别和抑制非语音噪声,而清音由于类似于白噪声,与宽带平稳噪声很难区分。

语音信号可以用统计分析特性来描述。由于语音是非平稳的随机过程,所以长时间的时域统计特性在语音增强的研究中意义不大。语音信号短时谱幅度的统计特性是时变的,只有当分析帧长趋于无穷大时,根据中心极限定理,才能近似认为其具有高斯分布。实际应用时只能将其看做是在有限帧长下的近似描述。在宽带噪声污染的语音信号增强中,可将这种假设作为分析的前提。

1.2.2 语音信号信息量

在语音增强技术中,语音信号到底包含多少信息量,需要多少比特才能被无失真地表示出来,这是一个很复杂的问题,涉及对信号失真的评价。目前常用的有三种评价方法,其中两种是由 Flanagan 提出的,另一种是由 Johnston 提出的,它们是在建立在以下三种不同的失真评价基础之上的:

(1) 语音信号的信噪比;

(2) 接收语音信号时,信号由听觉外围处理之后,人们在主观上能够感觉到的失真;

(3) 人在接收语音信号时,不正确接收音素的数目和正确接收音素的数目之比。

在上述三种情况下,实际所获得的比特率首先选择能够接受的失真等级,然后计算该失真等级所需的理论比特率。在第三种测量音素失真的方法下,可以将接受的失真级设置为零。如果所有的音素都能正确传送,那是所期望的最好性能。假设相邻的音素之间不出现相关,则平均信息速率很容易计算。按照 Shannon 信息理论,每一个符号需要的平均比特数或信息量 I 如下所示:

$$I = - \sum_i p_i \lg p_i \quad (1-1)$$

式中, p_i 为每一个符号 i 出现的概率; I 为信息量。英语有 42 个音素(符号),汉语

的音素有 48 个,其中,辅音 22 个,单元音 13 个,复元音 13 个。正常情况下,说话速率大约是每秒钟 10 个音素。利用音素出现的相对概率表,能够计算出每一个符号的信息量大约是 5bit,得到的全部信息速率大约是 50bit/s。其中,自然的寂静也包含在这个比特速率内。而系统仅仅传送音素序列,缺少发音人声音的个性特征,也就是声带的形状和对声道的描述。另一方面,相邻音素之间的相关也被忽略了。基于这些因素的考虑,可以把这一估计作为语音信息所需的比特率下限,或者人们感知语音信号的最低要求。

另外,采用第一种方法,将语音信号的信噪比作为失真评价,在不考虑编码器结构的情况下,可以得到语音信号信息速率的上限。在具有电话带宽的信号中,估计最大信息速率时,必须考虑合理的噪声等级。

假设 P 是语音信号的平均功率, W 是语音信号的带宽, G 是附加的噪声信号功率,并假设附加的噪声信号是高斯白噪声, C 表示语音所需最大的信息速率,根据 Shannon 理论,对于包含附加噪声 G 的语音信号, C 可由下式计算:

$$C = W \lg \left(1 + \frac{P}{G} \right) \quad (1-2)$$

式中,如果语音信号的带宽 W 为 3.5kHz,信噪比 $\left(\text{SNR} = 10 \lg \frac{P}{G} \right)$ 为 30dB,则它所包含的最大信息速率为 35kbit/s。这是语音所需的信息速率上限。公式(1-2)中,对于语音信号所存在的短期相关和长期相关,都没有考虑。而信号中所存在的结构性相关,则意味着冗余度,它能够在传输之前被去除,从而降低信息速率。

第二种估计方法包含了人的感知和理解;这是因为声音信号由人的听觉器官处理以后,其信息速率降低了。声音信号的某些特点,会由于人听觉系统的掩蔽效应而不被注意到。例如,在一个特有频率上的低幅度纯音,可以被一个靠近该频率更响的纯音掩蔽掉。去除人们在感觉上不能区分的特点以后,再来考虑信号的信息速率是比较恰当的。将理解失真评价的阈值设置为零,即听不到失真,则需首先计算语音信号的傅里叶变换,然后按频带进行计算,要求的量化器步长应使量化噪声在掩蔽阈值以下。掩蔽阈值和频带宽度都是建立在听觉系统基础之上的,所得到的信息速率估计称为理解熵。对于连续语音,理解熵约为 10kbit/s,相当于执行透明的语音编码所需的平均速率。因此人的感知和理解在语音增强技术中有很重要的作用。

1.2.3 噪声特征及其分类

噪声是指一切干扰人们休息、学习和工作的声音,即人们不需要的声音。此外,杂乱的振幅和频率,断续或统计上无规律的声振动也称为噪声。噪声来源于实际的应用环境,因而其特性复杂。对噪声进行划分的标准很多,各种分类方法的分

析角度不同。

根据噪声对语音频谱的干扰方式不同,可以把噪声主要分为加性噪声和乘性噪声。

1) 加性噪声

加性噪声是指当噪声对语音的干扰表现为两者信号在时域进行相加时,该噪声被称为加性噪声,显然噪声和语音在频域也为相加关系。

实际环境中背景噪声可以看成加性噪声,如风扇的声音、汽车引擎声、周围人说话声等。加性噪声是对背景噪声一种比较贴切的表述。麦克风等声音采集设备在正常工作的范围内,可以近似看成是一个线性系统,即产生信号的幅度和声强成正比。从能量角度看背景噪声和语音的声强是相叠加关系,两者对麦克风共同作用所形成的带噪语音信号等于各信号之和。严格来说背景噪声和语音不可避免地存在非线性作用,但这不是带噪语音的主要成分。由于背景噪声的广泛存在性,针对这类噪声的研究已成为语音增强领域的一个重点,本书主要研究的也是加性噪声。

2) 乘性噪声

乘性噪声是指噪声和语音在频域是相乘的关系,在时域和语音则是卷积的关系,因此也称为卷积噪声。在实际应用中乘性噪声主要体现在在语音采集、麦克风传输中电话信道和无线信道的频率选择特性。乘性噪声可以通过某种变换如同态滤波,转变为加性噪声,从而可以用加性噪声的方法来处理乘性噪声。

根据噪声统计特性随时间变化的程度不同,可将噪声分为周期噪声、脉冲噪声、缓变噪声和平稳噪声。

1) 周期噪声

发动机产生的干扰、市电干扰都是周期噪声。它的特点是在频域上具有许多离散的线谱。这种周期性噪声可以用梳状滤波器加以滤除,用数字信号处理的方法来实现。实际环境中产生的周期性噪声并非简单的只含线性谱分量,而是由许多窄带谱组成。该类型噪声往往是时变的,并与语音信号频谱重叠,必须采用自适应滤波的方法才有可能自动识别和区分噪声。

2) 脉冲噪声

打火、放电都会产生脉冲噪声。脉冲噪声表现为时域波形中出现的窄脉冲。只要脉冲噪声不是太密,一般可用内插法来去除这种噪声。

3) 缓变噪声

缓变噪声是在实际场合经常遇到的噪声,这种噪声的统计特性会随时间缓慢变化。人群噪声是典型的缓变噪声。

4) 平稳噪声

平稳噪声是指噪声的统计特性不随时间发生变化。由于噪声源的复杂性,在日常生活中所遇到的噪声大多不是平稳的,但对平稳噪声的研究却是重要的基础。

按照噪声覆盖频率范围可将噪声分为全频带噪声(也称宽带噪声)和窄带噪声。

1) 全频带噪声

覆盖了信号全部频率带的噪声称为全频带噪声或宽带噪声。它的来源有很多种,如热噪声、气流(如风)、呼吸噪声、量化噪声以及各种随机噪声源。对于平稳的全频带噪声通常可以认为是高斯白噪声;对于不具有白色频谱的噪声,可以先进行白化处理,然后转化为白噪声。

2) 窄带噪声

只覆盖信号的部分频率带的噪声称为窄带噪声,又称带选噪声。“口哨”噪声就是一种带选噪声。

此外,还有不同学科领域研究的噪声类型,如振动噪声、环境噪声、声噪声、航空噪声、建筑噪声、电磁噪声、光学噪声等;以产生噪声的来源物体而命名的噪声类型,包括机器噪声、发动机噪声、风扇噪声、飞机噪声、天电噪声、太阳噪声、宇宙噪声、生物噪声、电子噪声等。

在语音增强系统中常用的噪声有高斯白噪声、粉红噪声和工厂噪声。三者都是加性噪声,不同的是工厂噪声是一种非平稳噪声,存在一段尖锐的类似脉冲噪声的噪声,而高斯白噪声和粉红噪声则是平稳噪声。粉红噪声是指用正比于频率的频带带宽测量时,频谱连续且均匀的信号。粉红噪声是通过对一般电子设备方便产生的白噪声进行滤波后获得的,这种滤波器称为粉红噪声滤波器。本书主要使用的是 noisex-92 噪声库中的白噪声、粉红噪声及工厂噪声三种噪声。

1) 白噪声

白噪声(white noise)是由高质量的模拟噪声发生器获得的,是指功率谱密度在整个频域内均匀分布的噪声。所有频率具有相同能量的随机噪声称为白噪声。

2) 粉红噪声

粉红噪声(pink noise)定义为在与频带中心频率成正比的带宽(如倍频程带宽)内具有相等功率的噪声或振动,是由高质量的模拟噪声发生器获得的。粉红噪声的频率分量功率主要分布在中低频段。从波形角度看,粉红噪声是分形的,在一定的范围内音频数据具有相同或类似的能量。从功率(能量)角度来看,粉红噪声的能量从低频向高频不断衰减,曲线为 $1/f$,通常在线性频率坐标下,其功率谱密度以 3dB 每倍频程的速率下降。

3) 工厂噪声

工厂噪声(factory floor noise)一般是指工业设备机器在运转时产生的噪声。在 noisex 噪声库中工厂噪声包括两种:工厂车间噪声 1 的 factory floor noise 1,这类噪声主要在板切割及电器设备焊接附近记录;工厂车间噪声 2 的 factory floor noise 2,这类噪声主要在汽车生产车间记录。