

通信技术

基于 Matlab 仿真的语音信号增强算法研究

杨国荣 靳丽君

(西安铁路职业技术学院, 西安 710014)

摘要 语音增强技术是语音信号处理的一个重要分支,也是语音识别系统的重要组成部分。简单介绍了现有的各种语音增强算法,并对目前应用最为普遍的谱相减法作了系统的描述。对传统型、增强型、改进型的谱相减法进行了探讨验证,提出了一种基于 C 语言环境的算法程序,详细地介绍了该程序并画出了它的算法流程图,而且进行了算法仿真,通过分析它在 -5 dB 0 dB 5 dB 的信噪比下的输入输出波形,证明了该程序可以较好地实现语音增强。

关键词 语音增强 谱相减法 Martin 算法 N-S 流程图

中图分类号 TN912.34; **文献标志码** A

随着语音技术研究的深入和实际应用的增多,各种语音处理系统都面临着进一步提高性能的问题。语音增强是其中的关键技术之一。从 20 世纪 60 年代开始,语音增强的研究就一直没有停止。20 世纪 70 年代由于数字信号处理理论的成熟,语音增强曾经形成了一个研究热潮,取得了一些基础性成果。20 世纪 80 年代以后,VLSI 技术的发展为语音增强的实时实现提供了可能。目前,除了基于信号处理理论研究外,针对人的听觉系统的生理特性研究、语言学中上下文联想智能的研究等,都在进一步推动着语音增强的发展^[1]。

近年来随着数字技术的发展,出现了许多处理方法来解决这一问题。常用的方法有谱相减法、维纳滤波法、多带通滤波器法、自适应滤波器法等。本文对目前应用最为普遍的谱相减法作了系统的描述。

1 谱相减法

谱相减法总体上运算量较小,容易实时实现,增强效果也较好,是目前处理宽带噪声的最通用技术方法。即从带噪语音估值中减去噪声频谱估计,而得到纯净语音的频谱。由于人耳对语音频谱分量的相位不敏感,因而这种方法主要针对短时幅度谱^[2]。原理如下:

假定语音为平稳信号,而噪声和语音为加性信号且彼此不相关。此时带噪语音信号可表示为

$$y(t) = s(t) + n(t)。$$

式中 $s(t)$ 为纯净语音信号 $n(t)$ 为噪声信号。而用 $Y(\omega)$ $S(\omega)$ 和 $N(\omega)$ 表示 $y(t)$ $s(t)$ 和 $n(t)$ 的 Fourier 变换,则有下列关系存在

$$Y(\omega) = S(\omega) + N(\omega)。$$

对功率谱则有

$$|Y(\omega)|^2 = |S(\omega)|^2 + |N(\omega)|^2。$$

因为假定噪声为不相干的,所以不会出现语音与噪声的乘积项。只要从 $|Y(\omega)|^2$ 中减去 $|N(\omega)|^2$ 便可恢复 $|S(\omega)|^2$ 。因为噪声是局部平稳的,可以认为语音前的噪声与语音期间的噪声功率谱相同,因而可以利用语音前的“寂静帧”来估计噪声^[3]。

2010 年 6 月 30 日收到 7 月 20 日修改

第一作者简介:杨国荣(1976—),女,河北衡水人,西安铁路职业技术学院电子信息系教师,讲师,研究生,研究方向:通信与信息系统方向研究。E-mail: ygr01@tom.com。

而语音是不平稳的,而且只处理一小段加窗信号,上式可写为

$$|Y(\omega)|^2 = |S(\omega)|^2 + |N(\omega)|^2 + S_w^*(\omega) \times N_w^*(\omega) + S_w(\omega) N_w(\omega)。$$

式中,下标 w 表示加窗信号,* 表示复共轭。可以根据观测数据估计 $|Y(\omega)|^2$,其余各项必须近似统计均值。由于 $s(t)$ 和 $n(t)$ 独立,则互相关谱的统计均值为 0,所以原始语音的估值为

$$|S_w(\omega)|^2 = |Y(\omega)|^2 - [|N_w(\omega)|^2]。$$

式中, \wedge 表示估计值, $[|N_w(\omega)|^2]$ 为无语音时 $|N(\omega)|^2$ 的统计均值。因为涉及到估值,实际中此值可能为负,而功率谱不能为负,故可令负值为 0、改变其符号或平滑滤波。

为了 IFFT 再现语音,还需要 $S(\omega)$ 的相位,在此用 $P_h[S(\omega)]$ 表示。并用带噪语音的相位来近似。因而

$$S_w(\omega) = |S_w(\omega)| \exp(jP_h[Y_w(\omega)])。$$

则恢复的语音是估值的 IFFT。谱减法的原理框图如图 1。

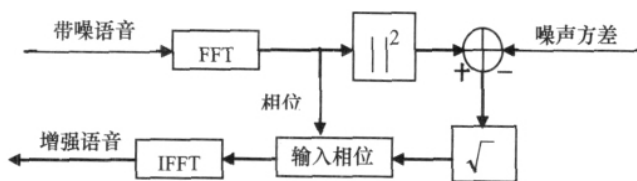


图 1 谱减法的原理框图

图 1 中, $\sqrt{\quad}$ 的处理用以将功率转化为幅度。

2 谱相减法的 C 语言程序简介

实验中,用 C 语言做出谱相减法的程序。其程序的结构如图 2 示。

其中, `main()` 函数为 `ssub.c`。他提供了一个人机对话的界面。在运行程序时,会出现一个 DOS 界面窗口,提示输入:相应的参数(用“-?”表示)、待处理信号的文件名和处理后信号所存文件的文件名。这样它就会选择相应的算法去输入待处理信号文件中的数据,并存入处理后信号所存的文件中。本文选择的是 Martin 算法,所以要输入参数”

-m”选择 Martin 算法。 `Ssub` 把输入文件 `9_!.dat` 读入,通过 Martin 算法,把处理结果存入输出文件 `out.bin` 中^[4]。

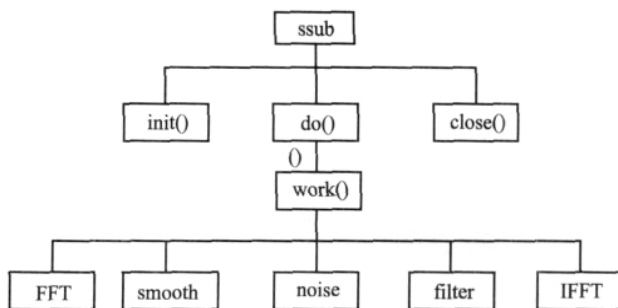


图 2 谱相减法 C 语言程序结构图

在 Martin 算法开始时需要初始化, `martin_init()` 来完成初始化。它对算法中所要用到的参数、帧长、标志位赋初值。并对累加缓冲、能量谱数组、常数 π 初始化,并构造宽度为 256 的 Hanning 窗,为执行 Martin 算法做准备^[5]。

在 Martin 算法中,我们选用 Hanning 窗作为窗函数,因为它可以有效的降低吉布斯效应,减小旁瓣分散去的能量。实验过程中,窗的长度不同,帧与帧之间的重叠程度不同,消噪的效果也不一样。我们用长分别是 256 点、128 点和 64 点的语音帧进行测试,发现不同帧长对噪声段的消噪效果影响不大;但语音段不同,窗口长度越长,听觉失真效果越小。原因是帧的长度取得比较短时,信号的频域分量变化比较快;而长窗时相邻两帧的差别较小。因此,我们取较长的 256 点作为一帧的长度。由试验可知,当 3/4 重叠时消噪效果最好。因此我们采用 256 点为一帧,窗函数 75% 重叠的滑动^[6]。

`martin_do()` 来完成算法的运行,它由预处理、调用 `martin_work()` 和输出结果三部分组成。开始时,用 `getfloat()` 从输入文件中读取一帧的数据读到 `inbuf` 数组中,计数 f 赋初值为 0。调用 `martin_work()` $f+1$ 。而后,再向后移 64 点再读取一帧数据到 `inbuf` 中,调用 `martin_work()`。如此重复操作。直到读完输入文件中的数据。最后把结果存入输出文件中。而 `martin_work()` 所作的是把 Martin 算法的过程一一理顺。`martin_work()` 又要调用

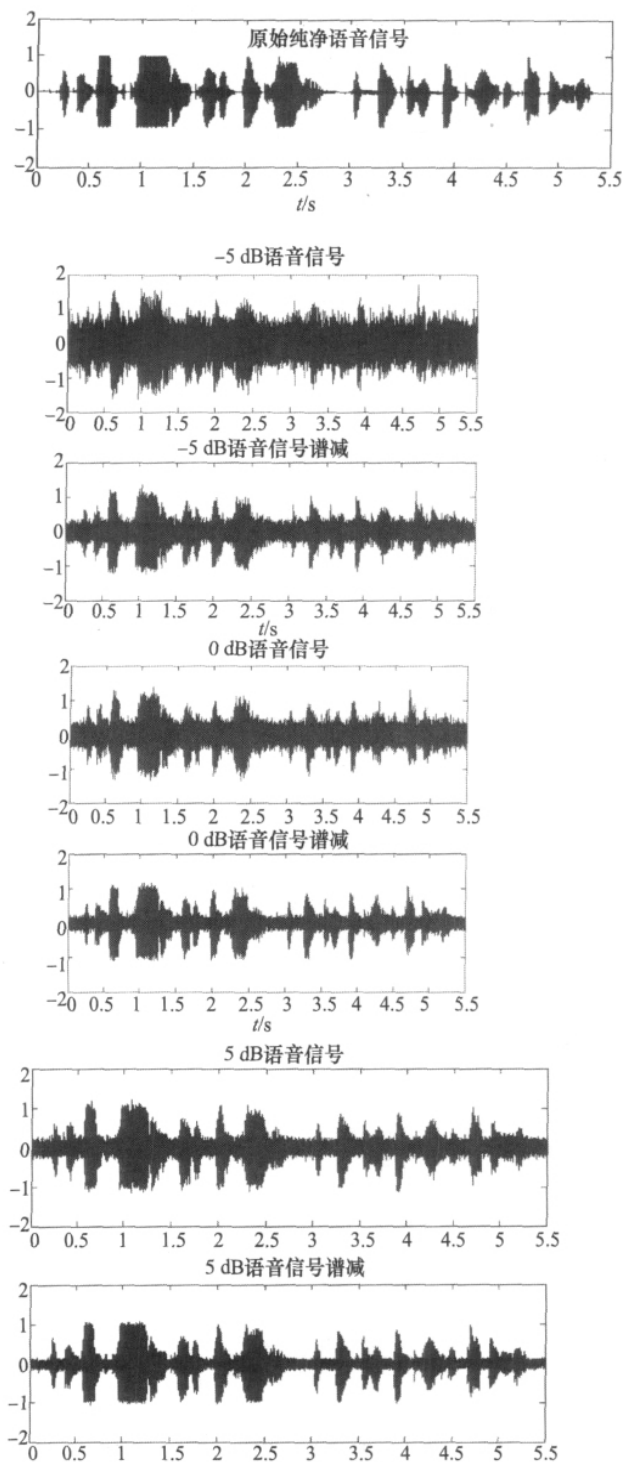


图3 原始纯净语音信号、-5 dB、0 dB、5 dB

的信噪比下的带噪信号及谱减后的信号的时域波形图

一些子程序来完成。这些子程序分别是: `fft()`、`smooth()`、`noise()`、`filter()` 及 `ifft()`。其中 `fft()`

和 `ifft()` 是一个程序,由于输入参数的顺序不同,进行的运算也就不同。最后 `martin_close()` 来关闭此算法^[7]。

3 谱相减法的数字仿真

我们用 Matlab 作为工具进行仿真。在 Matlab 中读入带处理文件,并把它存入 `in.dat` 中,然后用 C 语言程序运行,结果存在 `out.bin` 中,最后在 Matlab 中把 `out.bin` 读出^[8]。

试验中,我们选用 -5 dB、0 dB、5 dB 的信噪比语音信号进行试验,画出了原始纯净语音信号、-5 dB、0 dB、5 dB 的信噪比下的带噪信号及谱减后的信号的时域波形图。如图 3 所示,图中横坐标是时间(t),单位是秒(s),纵坐标是信号幅度大小。

经过以上反复试验我们可以得出,谱减法有一定的去噪功能,且在低信噪比下效果明显,信噪比能提高 6 dB 以上,但主观试听留有较为明显的音乐噪声。

4 结论

由于在对语音信号进行数字处理研究时一般都是假设在理想情况下,因此当很多技术从实验室走向实际应用时,环境噪声所带来的问题越来越严重,特别是作为语音处理技术中最有效手段之一的线性预测,很容易受到噪声的影响。因此研究各种技术将“干净”语音从语音和噪声的混合体中提取出来是十分必要的^[9]。本文通过谱相减法从带噪语音中尽可能提取纯净的原始语音,很大程度上消除了背景噪声,改进了语音质量。但由于噪声信号都是随机产生的,完全消噪几乎不可能,因此语音增强技术的研究将是一个长期探讨的课题。

参考文献

- 1 蔡文龙. 低信噪比下的语音增强技术研究. 秦皇岛: 燕山大学, 2005; 71—86
- 2 胡 钢, 沈文轩. 一种基于噪声动态检测的语音端点检测算法. 鞍山科技大学学报, 2004; (2): 47—50

3 张雄伟 ,陈 亮 ,杨吉斌. 现代语音处理技术及应用. 北京: 机械工业出版社 2003

4 蒋海霞 ,成立新 ,陈显治. 一种改进的谱减语音增强方法. 解放军理工大学学报 2001; (1) : 41—44

5 刘有恒. 信号检测与估计. 北京: 人民邮电出版社 ,1989

6 李宏伟 ,段艳丽 ,郭 英. 基于帧建重叠谱减法的语音增强算法及实现. 空军工程大学学报 2001; (5) : 59—62

7 刘 波. 语音转换的关键技术研究. 长沙: 国防科学技术大学 2005

8 于哲舟 ,杨佳东 ,蒲东兵 ,等. 多门限声纹识别方法. 吉林大学学报(信息科学版) 2005; (2) : 93—96

9 金学骥. 语音增强算法的研究与实现. 杭州: 浙江大学 2005

The Research of Speech Signal Enhancement

Algorithm Based on Matlab Simulation

YANG Guo-rong , JIN Li-jun

(Xi' an Railway Vocational and Technical Institute ,Xi' an 710014 ,P. R. China)

[Abstract] Speech enhancement technology is an important branch of speech signal processing ,but also the important component of speech recognition system. The existing algorithm of speech enhancement is introduced briefly , and applied the current application of the most popular spectral subtraction of the system. Discussed and verified for traditional ,enhanced and improved the spectral subtraction ,proposed the algorithm procedures based on C language environment ,the detailed algorithm is introduced in this program and drew out the flow chart ,through the analysis of the input and output waveform in - 5 dB ,0 dB ,5 dB , proved the program can achieve speech enhancement.

[Key words] speech enhancement spectral subtraction Martin algorithm N-S flow chart