

文章编号:1000-2073(2002)03-0058-04

基于减谱法的语音增强和噪声消除的研究^{*}

陶 智¹,葛 良²

(1. 苏州大学 理学院物理系,江苏,苏州 215006;

2. 苏州大学 信息学院通信与电子工程系,江苏,苏州 215006)

摘 要:介绍了减谱法进行语音增强的一种方法. 分别对语音和噪声信号进行傅立叶变换,求得它们的频谱,相减得到的是去噪后语音的频谱,再进行傅立叶反变换,即可得到增强语言信号,从而有效地抑制了噪声.

关键词:减谱法;傅立叶变换;语音增强

中图分类号: TN92

文献标识码: A

0 引言

语音通信过程中不可避免地要受到来自周围环境、传播媒介引入的噪声,通信设备内部电噪声以及其它噪声的影响. 在多数实际情况下,语音信号总是含有噪声成分. 语音信号的质量,根据噪声的大小及形式而下降,从部分听不清到全部听不懂,因此需要对带噪语音信号进行语音增强^[1].

语音增强是语音信号处理与识别系统的重要组成部分. 事实上从带噪语音的信号中提取完全纯净的语音是不可能的,因此语音增强的主要作用是改进语音质量,消除背景噪声,提高语音的自然度,使人乐于接受,这是一种客观度量;二是可提高语音可懂度,这是一种主观度量.

本文采用带噪语音的频谱减去采集到的噪声幅度谱从而达到语音增强. 这种方法运算量较少,容易实时实现.

^{*} 收稿日期:2002 - 01 - 20

作者简介:陶智(1970 年 12 月),男,江苏泰州人,实验师,在职硕士,主要从事语音信号的研究.

1 原理

1.1 理论分析

设观察到的一帧带噪信号为 $y(n) = s(n) + d(n)$, $0 \leq n \leq N-1$, 其中 $s(n)$ 为纯静音, $d(n)$ 为平稳加性高斯噪声. 便于分析, 将 $y(n)$ 在一组基 $\{ \phi_k(n) \}$ 上展开, 即 $\{ \phi_k(n) \}$ 满足:

$$\phi_k(n) \phi_k(m) = \delta_{n,m} \quad (1)$$

同时 $y(n)$ 的展开式为

$$y(n) = \sum_{k=0}^{N-1} Y_k \phi_k(n), \quad Y_k = \sum_{n=0}^{N-1} y(n) \phi_k(n) \quad (2)$$

若 $y(n)$ 的相关长度小于 N 帧长, 那么一簇近似的特征函数是

$$\phi_k(n) = \frac{1}{\sqrt{2}} \exp(j \frac{2\pi nk}{N}) \quad (3)$$

由(2)式容易看出, $y(n)$ 的展开过程相当于离散傅立叶变换, Y_k 对元的频谱分量.

1.2 原理

设 $y(n)$ 的傅立叶变换为

$$Y_k = |Y_k| \exp(j \phi_k) \quad (4)$$

$s(n)$ 的傅立叶变换为

$$S_k = |S_k| \exp(j \phi_k) \quad (5)$$

$d(n)$ 的傅立叶变换为 N_k , 由 $y(n) = s(n) + d(n)$, 则有 $Y_k = S_k + N_k$. 由此可得

$$|Y_k|^2 = |S_k|^2 + |N_k|^2 + S_k N_k^* + S_k^* N_k \quad (6)$$

由于 $s(n)$ 与 $d(n)$ 独立, 所以 S_k 与 N_k 独立. 而 N_k 为零均值的高斯分布, 所以有

$$E[|Y_k|^2] = E[|S_k|^2] + E[|N_k|^2] \quad (7)$$

由此可得原始语音的估计值

$$| \hat{S}_k | = [|Y_k|^2 - E(|N_k|^2)]^{\frac{1}{2}} = [|Y_k|^2 - \sigma_n^2]^{\frac{1}{2}} \quad (8)$$

这就是谱相减法的基本原理, 如图1所示^[2]:

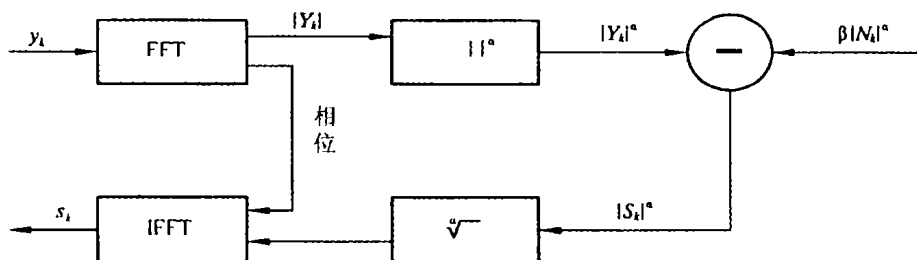


图1 谱相减法原理框图

在实际的增强过程中,更多地使用的是谱相减法的改进形式:

$$|S_k| = \{ [|Y_k| - \sigma_n(k)]^{\frac{1}{\alpha}}, \quad \} \quad (9)$$

这里 α 为大于零的一个小数. 这种改进形式与普通谱相减法的差别在于引入 α 与 $\sigma_n(k)$ 两个参数. 噪声估计值乘以系数 α 是基于这样的一种考虑: 由于在减谱法过程中, 是以无声期间的统计平均的噪声方差代替当前分析帧各频率点的噪声频谱分量. 而噪声频谱具有高斯分布, 即其幅度随机变化的范围很宽. 因此相减时, 若该帧某频率点噪声分量较大, 因此会有很大一部分保留, 在频谱上呈现随机出现的尖峰, 在听觉上形成了有节奏性起伏的类似音乐的残留噪声. 为了解决上述问题, 在语音能量较高的区域, 可以令 $\alpha > 1$, 即有意识地多减去一点, 这样可以更好地相对突出语音频谱. 当然失真的可能也会增大. 所以我们在选取的时候, 一定要根据实验的结果来确定, 当然我们也可以根据语音当时的大小来动态地确定大小. 调节参数也可以获得类似的效果^[3].

2 实验结果及讨论

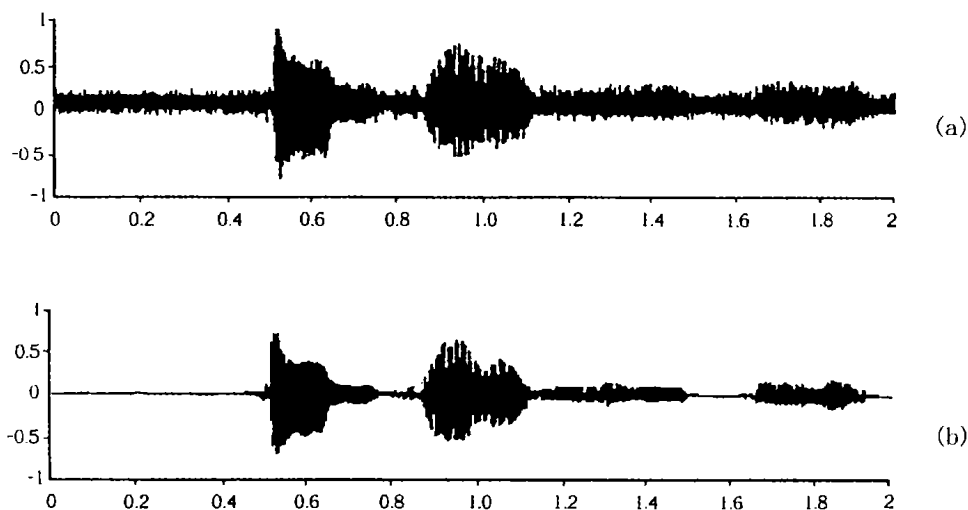


图2 一段受扰语音去噪结果

图2(a)是含噪语音,图2(b)是用减谱法进行语音增强后的图.比较两幅图,我们可以清楚地看出噪声被有效地去除了.我们对两段长度为30s的含噪语音进行了去噪,信噪比的改善情况如表1所示.

我们在这里主要对不同的 α 进行了测试,测试结果发现如果原始信号的信噪比较高,去噪效果随 α 变化不大,但是如果原始的语音信噪比较低的话,则 α 的选择对提高信噪比的影响较大.结果实验测定,我们发现当 α 取2的时候效果一般比较好.

3 结 论

用减谱法进行语音增强,噪声能得到很明显的抑制,但无法避免地引入了音乐噪声,但我们在实验中,用了语音信号的平滑、尾部处理、参数改变等方法,使得音乐噪声的干扰得到了很好的抑制.

表 1 参数 、 改变后的语音信号信噪比改善情况

			原始语音 / dB	去噪语音 / dB
Voice1. wav	2	1	- 1. 06	9. 08
Voice1. wav	2	2	- 1. 06	9. 17
Voice1. wav	2	3	- 1. 06	9. 16
Voice1. wav	2	4	- 1. 06	8. 32
Voice1. wav	2	5	- 1. 06	8. 10
Voice1. wav	2	6	- 1. 06	8. 10
Voice1. wav	2	7	- 1. 06	7. 67
Voice1. wav	2	8	- 1. 06	7. 52
Voice1. wav	2	9	- 1. 06	7. 16

参考文献 :

[1] 易克初 ,田斌 ,付强. 语音信号处理[M]. 北京 :国防工业出版社 ,2000.

[2] WEI Wei ,CHEN Yan-pu. Speech enhancement by spectral component selection[J]. Proceedings of ICSP, 2000 ,674 - 678.

[3] 陶智 ,葛良. 基于小波变换的语音增强与噪声消除的研究[J]. 苏州大学学报 ,2001 ,17(4) ,74 - 77.

Speech enhancement and noise reduction based on the spectral subtraction

TAO Zhi¹ ,GE Liang²

(1. Dept. of Phys. ,School of Sci. ,Suzhou Univ. ,Suzhou 215006 ,China ;
2. Dept. of Electron ,School of Information ,Suzhou Univ. Suzhou 215006 ,China)

Abstract :A general problem of removing additive background noise from the noisy speech in the spectral subtraction is described. Fourier transform is used in noisy speech and noise respective to get their frequency spectrum. The frequency spectrum of pure speech is got through subtration. After inverse Fourier transform of subtration spectrum , the pure speech is got.

Key words:spectral subtraction ;Fourier transform ; speech enhancement

(责任编辑 :周建兰)