

低信噪比下的语音增强处理

李 国 锋

(南京市公安科学技术研究所, 南京 210001)

1994 年 5 月 9 日收到

摘要 本文介绍了一种基于功率谱减的方法来增强带有白噪声的语音信号。过量功率谱减是语音增强的一个有效的方法,其处理后产生的纯音噪声采用中心限幅的方法可以很好地得到抑制。

本文中对 $-8\text{dB} \sim +3\text{dB}$ 的汉语语句处理,信噪比可提高 $6\text{dB} \sim 10\text{dB}$ 。

关键词 低信噪比,语音增强,功率谱

Enhancement of low snr speech

Li Guofeng

(Nanjing Public Security Science and Technology Institute, Nanjing 210001)

Abstract This paper describes a method for enhancing speech degraded by white noise, based on the power spectrum subtraction. The- Subtraction of the noise power spectrum over-estimate is an effective method. A new noise appearing in the processed signal, referred to as “music noise”, can be effectively eliminated by using “center-clipping process”.

In experiment, sentences of Chinese with about $-8\text{dB} \sim +3\text{dB}$ signal-to-noise ratio was processed and improved up to $6\text{dB} \sim 10\text{dB}$.

Key words Low signal-to-noise ratio, Speech enhancement, The power spectrum

1 引言:

在语音通讯中,随时随地都会受到各种各样的噪声干扰,噪声环境中的通讯成为一个亟待解决的问题;在语音分析、合成等研究中,线性预测在实验室里的研究中是语音处理技术的有效手段,但当语音处理由实验室走向实际应用时,线性预测手段却是最容易受噪声影响的,如果把线性预测看作频谱匹配过程,则在大量噪声使频谱畸变时,预测器就设法与畸变的频

谱匹配,而不是与基本语音匹配,因此环境噪声的存在所带来的问题就愈来愈重要了。

谱减法作为处理宽带噪声的传统方法对于抑制背景噪声有一定的作用,特别是在无参考信号源的单话筒录音系统下是一个有效的方法,但是传统的谱减法处理后的语音往往保留了大量的噪声,所以在实际中不能很好地应用。采用过量功率谱减的方法^[3]可使处理效果产生明显的提高,同时采用中心限幅、噪声功率估计等的方法,使处理结果达到实际应用的要求。

2 基本方法:

假定语音为平稳信号,而噪声为加性和不相关的,这样可以把带噪声语音信号表示如下:

y(t) = s(t) + n(t) (1)

其中 s(t) 是语音信号, n(t) 是噪声, y(t) 是带噪声语音信号。

用 Y(w)、S(w)、N(w) 分别表示相应的傅里叶变换,则式(1)有下列的变换关系:

Y(w) = S(w) + N(w) (2)

则有下列功率谱关系:

|Y(w)|^2 = |S(w)|^2 + |N(w)|^2 (3)

(因为假定噪声与语音信号不相关,因此二者的乘积项为零可省略。)

由式(3)可得如下表示:

P_s(w) = P_y(w) - P_n(w) (4)

其中, P_s(w)、P_y(w)、P_n(w) 是相应的功率谱, P_s(w) 的值可由语音之间的寂静段估值得来。

通过实验表明,利用式(4)处理后,将有近 1/3 的噪声保留在语音信号中,而且还会产生纯音噪声。因此式(4)可修改为下式:

P_s(w) = { P_y(w) - a * P_n(w) 当 P_y > a * P_n; b * P_n(w) 其它 (5)

(其中 a ≥ 1, b < 1)

理论上可以推算^[2], 高斯噪声累积概率密度 F(N) = 1 - exp[-N^2/σ_N^2]。当 N^2 = σ_N^2, 即式(5)中 a = 1 时, F(N) = 63%, 表明可减去 63% 的噪声功率;当 N = 2σ_N, 即 a = 4 时, F(N) = 98.2%, 这样仅剩 1.8% 的噪声;另外, 语音的能量往往集中在某一些频段, 相对幅度较噪声高。因此可以采用过量噪声功率谱的方法来提高降噪效果。实验中, 在辅音段取 a = 1.5—3, 元音段取 a = 2.5—4。

3 实现过程:

采用过量功率谱减法实现降低噪声, 其处理框图如图 1。

实现过程中几个实际问题的处理如下:

3.1 噪声功率谱的估计: 在通常的谱减

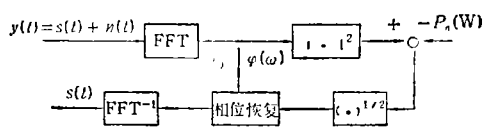


图 1 基本原理框图

法中, 噪声的功率估计往往采用语音段开始的前几帧或语音间寂静段来求功率谱。用开始的几帧求功率谱作为以后处理的对照会因时变的不同产生较大的误差; 而语音间的寂静段与语音段的分界在信噪比较低的情况下不易分清, 因此也容易产生不正确的结果。

本文采用带噪声语音的处理前后能量的变化情况来判断是语音段还是非语音段^[4]。公式如下:

D_k = (1/N * sum_{i=1}^N y_k^2(i)) / (1/N * sum_{i=1}^N s_k^2(i)) (6)

其中: s_k(i) 是第 k 帧处理后的值, y_k(i) 为处理前的值。

经实验验证, 对于仅有背景噪声的非语音段, 处理前后帧内平均能量衰减很大, 即 D 值较大; 而对于语音段则 D 值较小。这样, 就可判断出当前处理帧是否寂静段, 若为寂静段, 则该帧处理前的能量值 P_n(w) 作为下一帧的相减比照就比较精确。

3.2 利用谱减法对语音处理后, 语音信号中保留有大量的纯音噪声, 这种噪声对语音的影响亦很大。在 Berouti. M. 等的方法^[3]中, 为掩蔽纯音噪声而保留少量宽带噪声, 实现方法是令式(5)中的 b = 0.01—0.5。通过实验表明, 保留宽带噪声的方法不能达到满意的去噪效果。本文采用中心限幅的方法, 即通过其高幅部分而削除其低幅部分, 可以较好地抑制纯音噪声, 对语音可懂度也没有明显损伤。

3.3 在式(3)中我们假定语音信号为稳定信号, 实际上语音信号是时变的, 因此只能用一小段加窗的语音信号作为稳定信号来处理, 则式(3)改为下式:

|S_w(w)|^2 = |Y_w(w)|^2 - |N_w(w)|^2 (7)

其中,式中的 w 下标表示加窗信号。

由于海明窗^[1]的一半叠加的结果比较平稳,所以为了避免处理过程中,正FFT乘时窗函数,逆FFT后除时窗函数的麻烦,本文在处理过程中采用海明窗,每次叠加1/2帧,使帧与帧之间平滑相接。

3.4 在频域处理过程中仅处理了功率谱而忽略了相位,逆变换中采用原来相位谱来恢复时域信号。由余崇智等文^[2]中的实验结果表明,虽然原相位谱与实际语音本身的相位有所不同,但是其对语音的懂度影响不大,因此在处理中可用保留的相位谱来恢复语音信号。

3.5 带有色噪声的语音信号的处理:有色噪声是有周期性的噪声^[3],如电扇声、电钻声等,它在频谱中表现为许多有规律的离散的窄谱峰,而且不同各帧中分布规律相似,对其处理可采用平均谱对照相减的方法来完成。而平均谱的取得,用语音段开始前的若干帧的频谱相加平均而来。处理后信噪比可提高7dB—12dB。

4 实验结果:

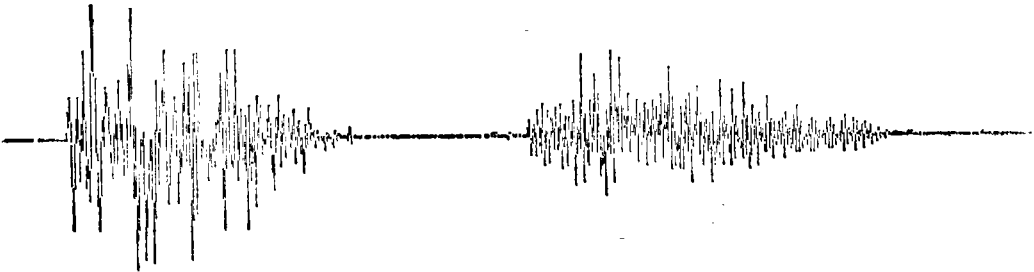
我们在实验中使用一段语音信号“到黑龙江”作为纯净原始语音,并以两种噪声作为处理对象。

一种是宽带噪声:利用计算机叠加随机数产生一定信噪比的加性白噪声的语音信号,在实验过程中调整不同的 α 值和中心限幅值,其处理结果可使信噪比提高6dB—10dB,参见图2。图中只显示了前两个的结果。

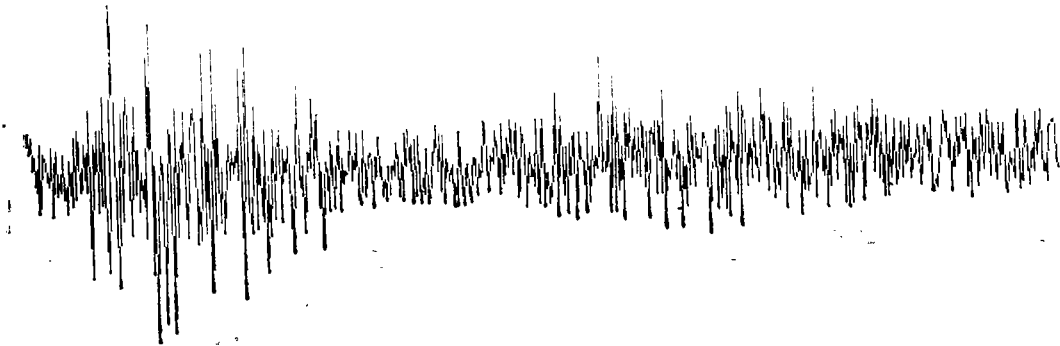
一种是有色噪声:利用带有很强电扇声的“到黑龙江”语音信号作为处理对象,可以使信噪比提高7dB—12dB。

5 结束语:

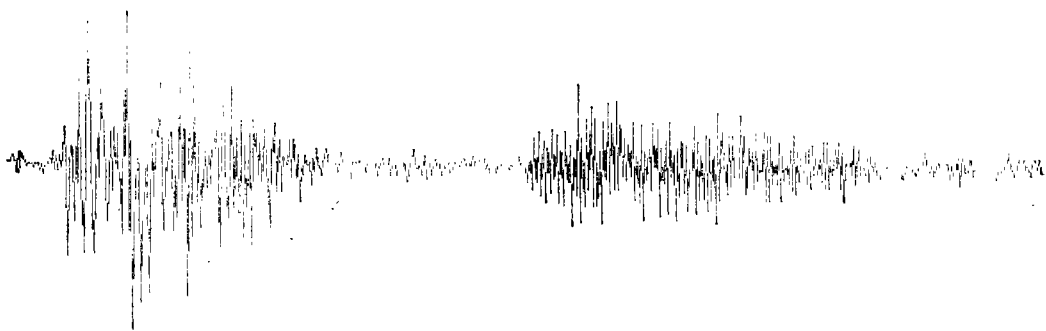
本文中采用了噪声功率谱估计的方法并使其很好地应用于实验中证明了其在实际处理中的可行性;过量功率谱相减的方法是语音增强的一种有效的方法,在实验中可以使信噪比提高7dB左右。本文的方法我们曾经用于处理实际的带环境噪声的录音带,有比较明显的效



(a)



(b)



(c)

图 2 (a)原始语音信号(“到黑”2 字) (b) 带有白噪声的语音信号(“到黑”2 字)
(c) 增强处理后的语音信号(“到黑”2 字)

果。

参 考 文 献

[1] (美) T. W. 帕森斯著,文成义,常国琴,王化周等译.
语音处理,第一版.北京:国防工业出版社,1990.257—
268.

[2] 余崇智,华一满,胡春年等. 声学学报,1983, 8(1): 14—21.
[3] Berouti, M Schwartz R. and Makhoul J. ICASSP, 1979: 208—211. (注: 无卷期号)
[4] 沈亚强,程仲文. 第二届全国人机语音通讯学术会议论文集,1992: 399—404.
[5] 陈永彬编. 数字信号处理,第一版. 南京: 南京工学院出版社,1987: 97—102.

北京长城全周影院的音质设计

郑 敏 华

崔 东 元

(中国科学院声学研究所, 北京 100080) (北京工业大学, 北京 100022)

1994 年 6 月 14 日收到

摘要 一种具有大影像、360 度水平视野、接近于全方位的立体声效果,以高保真的声画环境将观众包围,使观众产生强烈临场感的全周影院在国际上也为数不多,而在我国建造更是一种新的尝试。本文通过介绍北京长城全周影院的音质设计和声学性能的测试说明,虽然这种新形式电影系统与常规电影有明显区别,具有很多特殊要求,但通过建筑、声学、光学设计者的共同努力及密切配合,最终获得了视听性能满意的结果,使影院达到了国际先进水平,也为全周影院的再次建造和进一步提高提供了科学依据。

关键词 全周影院,混响时间,背景噪声

Acoustic design of Beijing “Great Wall” circle vision theatre

Zheng Minhua

(Institute of Acoustics, Academia Sinica, Beijing 100080)

Cui Dongyuan

(Beijing Polytechnic University, Beijing 100022)

Abstract A circle vision theatre is a theatre which can provide audience with a