

非空气传导语音的噪声特性分析及处理

陈玲, 王喆, 姚佩阳
(空军工程大学 电讯工程学院, 陕西 西安 710077)

摘要:喉部送话器具有很强的抗环境噪声干扰的能力,但使用时内部会产生一定的噪声。本文利用同步录音的麦克风语音信号和喉部送话器语音信号的相关性,用自适应线性滤波法提取出了喉部送话器的噪声信号,对其特性进行了分析,并提出了解决方案。

关键词:非空气传导语音;噪声特性;谱减法;语音增强

中图分类号:TN911.7 **文献标识码:**A **文章编号:**1009-3516(2003)02-0052-03

一般的语音识别系统采用空气传导的受话器如麦克风作为语音传感器。当在强背景噪声的环境下使用时,识别率将明显下降。为提高语音识别系统的抗噪性能,可以使用喉部送话器等非空气传导受话器。喉部送话器使用时紧贴使用者的喉部,使用者说话时声带振动,引起喉部送话器中的碳膜发生形变,使其电阻发生变化,从而使其两端的电压发生变化,于是振动信号转化为电信号,即语音信号。空气中传导的声波无法使喉部送话器中的碳膜发生形变,所以喉部送话器感受不到空气传导的声音,因而具有很强的抗环境噪声干扰的能力。但喉部送话器在使用时,其内部会产生一定的噪声,致使其传导的语音不够清晰,影响了语音识别率。针对这一问题,本文对喉部送话器的噪声特性进行了分析,并提出了处理方法。

1 喉部送话器语音的噪声特性分析

1.1 喉部送话器语音的噪声提取

首先在安静的实验室环境下,将麦克风和喉部送话器选择 8 kHz 采样率和 16 bit 量化精度进行同步录音。图 1 是麦克风及喉部送话器录入的“你好”语音信号的时域波形。

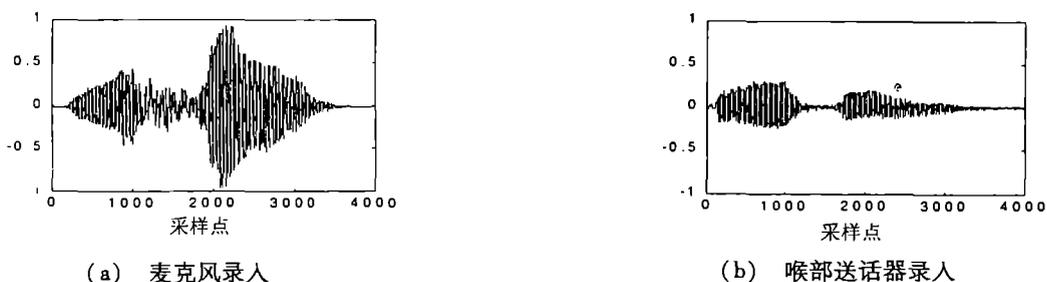


图1 “你好”语音信号时域波形

从图中明显可以看出,喉部送话器语音的波形幅度有一定的衰减。从录音回放效果来看,喉部送话器语音的声音沉闷,音强较小,且噪声较大。

实验中采用自适应滤波器来提取噪声信号。其原理源于自适应噪声抵消器。自适应噪声抵消器用一个参考噪声作为辅助输入,当辅助输入噪声与语音中的噪声完全相关时,自适应噪声抵消法能完全排除噪声的随机性,彻底地抵消语音中的噪声成分^[1]。由于麦克风语音信号和喉部送话器语音信号是同步录音的,所

收稿日期:2002-10-08

基金项目:军队科研基金资助(KJD03302A)

作者简介:陈玲(1972-),女,陕西西安人,硕士生,主要从事语音识别研究。

以这两个信号必然是相关的。但是,喉部送话器语音在中高频上有较大的衰减,其语音的频率范围大大降低,约为(0~1.5) kHz,而麦克风语音的频率范围是(0~5.5) kHz^[2],所以,两种语音还是有较大差别的。为了更好地提取噪声,在实验中我们先将麦克风语音信号通过一个截止频率为 2.0 kHz 的低通滤波器,然后连接到自适应噪声抵消器的参考输入端,将喉部送话器语音放在信号源端,这样就可以比较彻底地滤除语音,提取噪声。其工作原理见图 2。这里,设喉部送话器语音 $x(n) = s(n) + n(n)$, 其中 $s(n)$ 为纯净语音, $n(n)$ 为噪声,设低通滤波后的语音信号为 $s_1(n)$ 。由以上分析可知, $s(n)$ 与 $s_1(n)$ 是完全相关的。其工作原理实质上为以均方误差 $E[e^2(n)]$ 最小为准则,对喉部送话器输出的语音信号 $s(n)$ 进行最优估计,即 $y(n) = s(n)$, 然后从含噪语音中减去 $s(n)$, 提取 $n(n)$ 。提取到的噪声如图 3 所示,其中图 3(a) 是语音段噪声的短时域波形,图 3(b) 是其短时频域幅度谱。实验中又对无声段信号进行了测量,图 3(c) 是无声段噪声短时频域幅度谱。

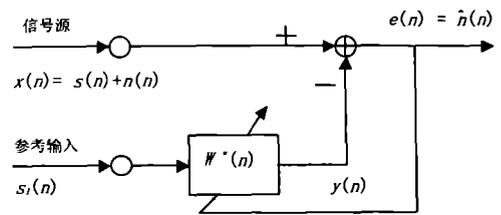
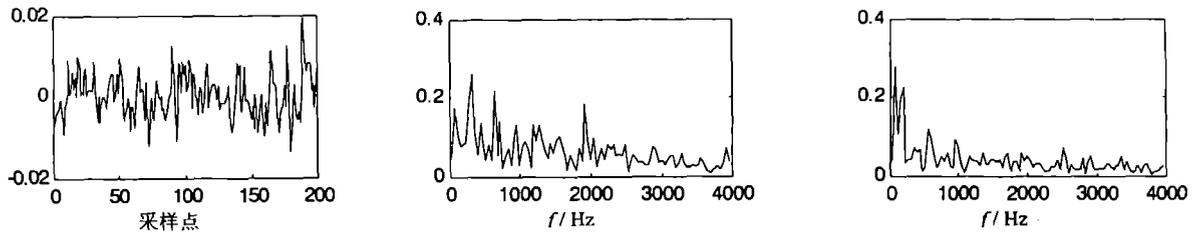


图 2 利用自适应噪声抵消器提取喉部送话器噪声



(a) 语音段噪声短时时域波形 (b) 语音段噪声短时频域幅度谱 (c) 无声段噪声短时频域幅度谱

1.2 喉部送话器语音的噪声特性分析

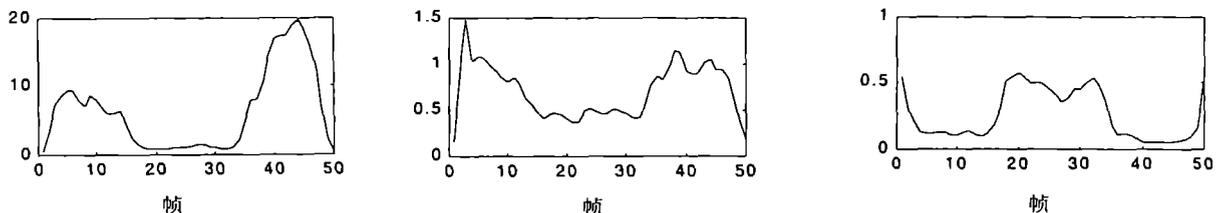
从图 3 可以看出,喉部送话器有声段噪声的频谱除个别频段有尖峰外,其余部分相对比较平坦,且噪声频谱低频处较高频处幅值略大,大致为斜率较小的下降三角噪声。而无声段的噪声为 $1/f$ 噪声,也就是红噪声。

对喉部送话器的噪声特性进行认真分析后,发现喉部送话器语音的噪声大小与该帧语音的总能量的大小有关系。噪声能量基本上随着语音能量的变化而变化,但又不纯粹是正比关系。经过大量的观察实验,噪声能量 ω_N 可近似为

$$\omega_N = 0.02E_s + E_n \tag{1}$$

其中 E_s 是喉部送话器带噪语音一帧的能量, E_n 实际上是一个常量,它是无声段噪声的平均帧能量。

图 4 是一个由喉部送话器送出的完整的汉字的语音能量、噪声能量及二者比例的分布关系,横坐标单位是帧。从图中可以看出:在语音能量高的地方,噪声所占的比例小;在语音能量低的地方,噪声所占的比例反而大。分析其原因,是因为喉部送话器的噪声由两部分组成,一部分是由语音引起的,这部分噪声可称为噪声 1,它与语音能量成正比;另一部分噪声和语音没有关系,是喉部送话器的固有噪声,这部分噪声称为噪声 2,其值可用无声段噪声的平均帧能量来近似估计。当语音能量很大时,总噪声的大小主要由噪声 1 决定,噪声 2 可以忽略不计;但当语音能量较小时,就必须考虑噪声 2 的影响了。这也就是喉部送话器语音的噪声能量与语音能量之间并不纯粹是正比关系的原因。



(a) 有声段语音能量分布 (b) 有声段噪声能量分布 (c) 有声段噪声能量与语音能量的比值

图 4 一个完整汉字的语音能量、噪声能量分布及二者比例关系

2 喉部送话器语音的去噪

由以上分析可知,喉部送话器语音的噪声特性近似于斜率较小的下降三角噪声。而噪声能量可以通过式(1)进行估算,所以可以采用谱减法进行去噪,相位由有声段带噪语音频谱的相位来取代。实验中先在 MATLAB 数学软件中产生白噪声^[3],然后用一条下降直线来描述其噪声谱。经过大量实验,确定这条直线的斜率为 $-m/k$,在有声段 $m = \sqrt{0.02E_s + E_n}$,而在无声段 $m = \sqrt{E_n}$ 。这里的 E_s 和 E_n 都是在频域上计算出的有声段语音的帧能量和无声段噪声的平均帧能量; $k = N/2$, N 是每帧语音信号傅立叶变换所取的点数。图 5 是噪声频域幅度谱之一。

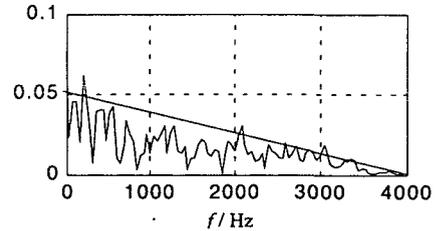


图 5 噪声频域幅度谱之一

在实验中,用这种方法对图 6(a) 所示的带噪语音进行去噪后,其信噪比提高了约 14 dB,图 6(b) 是去噪后语音的时域波形。从图上看,噪声得到了很好的消除。人耳主观上听起来,去噪后的语音噪声减小,清晰度提高了。

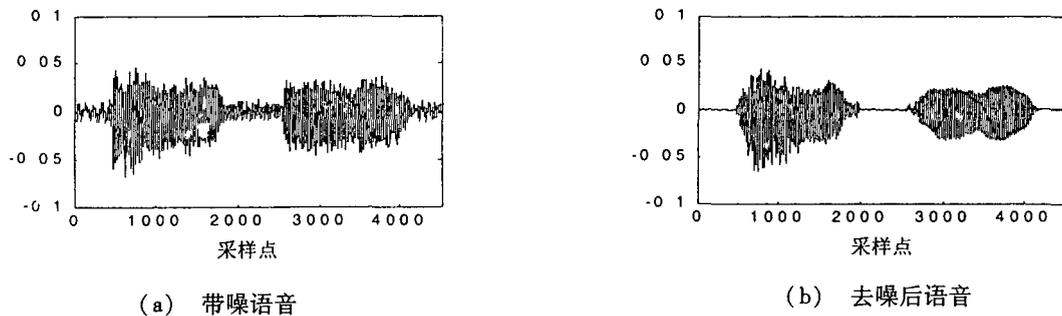


图 6 带噪语音去噪前后时域波形

参考文献:

- [1] 朱民雄,闻 新. 计算机语音技术[M]. 北京:航空航天大学出版社,2001.
- [2] 张歆奕,何 强. 非空气传导语音的分析及识别[J]. 信号处理,1999,(2):63-67.
- [3] 楼顺天,李博菡. 基于 MATLAB 的系统分析与设计[M]. 西安:西安电子科技大学出版社,2000.
- [4] 徐旭华,王 刚,郭 英,等. 一种基于噪声生成模型的语言消噪算法[J]. 空军工程大学学报(自然科学版),2001,2(1):45-48.

(编辑:门向生)

The Analysis and Processing of the Characteristics of the Non-air Spread Speech Noise

CHEN Ling, WANG Zhe, YAO Pei-yang

(The Telecommunication Engineering Institute, Air Force Engineering University, Xi'an 710077, China)

Abstract: The laryngophone, in spite of its strong ability of anti-environment, will produce a certain noise itself while being used. In this paper, by making use of the relevance of the microphone speech signal with the laryngophone speech signal recorded synchronously, the noise signal of the laryngophone is abstracted by the auto-adaptive linear filter, and based on this, its characteristics are analyzed and the solution is put forward.

Key words: non-air spread speech; noise characteristics; subtraction in frequency domain; speech enhancement