

# 抑制坦克强背景噪声的改进谱减法研究

刘兴涛,王忠,张维

LIU Xing-tao, WANG Zhong, ZHANG Wei

四川大学 电气信息学院,成都 610065

School of Electrical Engineering and Information, Sichuan University, Chengdu 610065, China

E-mail: liuxingtao110@tom.com

LIU Xing-tao, WANG Zhong, ZHANG Wei. Improved spectral subtraction algorithm to restrain strong background noise in tank cockpit. *Computer Engineering and Applications*, 2010, 46(4): 134-135.

**Abstract:** Spectral subtraction algorithm, with little computation, real-time processing easily, and good enhancement effects, is a very effective method to deal with broad-band noise. According to the traditional and many kinds of improved spectral subtraction speech enhancement algorithms, an improved spectral subtraction algorithm is presented. This enhancement algorithm smooths noisy speech in the time-domain and statistically weighted noisy speech in the frequency-domain based on the characteristics of speech and noise respectively. Subjective and objective tests are made on this algorithm. The results show this algorithm is very effective to reduce the background noise and music noise, moreover ensure the intelligibility and naturalness of speech.

**Key words:** speech enhancement; spectral subtraction; tank noise; music noise

**摘要:** 谱减法是处理宽带噪声较为传统和有效的方法,它运算量较小,容易实时处理,增强效果也较好。根据经典谱减法及其各种改进形式的基本原理,提出一种新的改进谱减法语音增强算法。根据语音和噪声各自的特性,对带噪语音进行时域平滑和频谱统计加权处理。对该算法进行客观和主观测试表明:相对于传统的谱减法,该算法能更好地抑制背景噪声和音乐噪声,同时也较好地保持了语音的可懂度和自然度。

**关键词:** 语音增强;谱减法;坦克噪声;音乐噪声

DOI: 10.3778/j.issn.1002-8331.2010.04.043 文章编号: 1002-8331(2010)04-0134-02 文献标识码: A 中图分类号: TP391

语音通信过程中不可避免地会受到来自周围环境的干扰,接收到的语音不再是原始的纯净语音,而是受噪声干扰的带噪语音信号,很强的背景噪声例如机械、枪炮等均会影响语音通信质量。特别是装甲车中,它的通信不仅受到很强机械噪声的干扰,有时还受到枪炮声的随机干扰,在科学技术飞速发展的今天,如果处理得不好,可能会带来很多不必要的伤亡,甚至是毁灭性的。因此,需要对带噪语音进行语音增强,改善语音质量,消除背景噪声,提高语音的清晰度和可懂度,使人乐于接受。

目前语音信号的降噪方法大致有四类:噪声对消法、谐波增强法、基于参数估计的语音再合成法和基于语音短时谱估计的增强算法<sup>[1]</sup>。其中谱减法是目前最常用的语音增强技术,特点是运算量较小,易于实现,其思想是假设噪声是平稳的或是变化缓慢的加性噪声,并且假设语音信号与噪声是相互独立的条件下,从带噪语音的功率谱中减去噪声功率谱,从而得到较为纯净的语音频谱。但是该方法经典形式的不足之处是会产生具有一定节奏性起伏、听上去类似音乐的“音乐噪声”<sup>[2]</sup>。针对这一问题人们提出了各种改进形式,这些形式虽未将“音乐噪

声”彻底滤除,但的确在不同程度上使其有所削弱,令听者乐于接受。

该文对谱减法经典形式及多种改进形式在理论上加以分析,并在此基础上,提出了一种改进方案,此种方法不仅对抑制残留的音乐噪声有很好的效果,而且对基于坦克强背景噪声也有很好的效果。

## 1 基本谱减法原理

语音是非平稳随机过程,但在 10~20 ms 的分帧内可以近似看作是平稳的,因此语音的短时谱具有相对稳定性<sup>[3]</sup>。假设纯净语音信号  $s(n)$  被加性噪声  $d(n)$  所污染,形成带噪语音  $y(n)$ ,即

$$y(n) = s(n) + d(n) \quad (1)$$

式(1)中  $n$  为时间采样点。短时谱处理时要对信号分帧处理,分帧时时域信号首先分成相叠的帧信号,帧长为  $N$ ,帧移为  $N/2$ ,设观察到的一帧带噪语音信号为:

$$y_i(n) = s_i(n) + d_i(n), i=0, 1, \dots; n=0, 1, \dots, N-1 \quad (2)$$

其中,  $s_i(n)$  为一帧纯净语音,  $d_i(n)$  为一帧平稳高斯噪声,且有

基金项目:国家自然科学基金(the National Natural Science Foundation of China under Grant No.49901013);中国博士后基金。

作者简介:刘兴涛(1983-),男,硕士,研究方向:信号与信息处理;王忠(1964-),男,博士后,副教授,硕士生导师,主要研究方向为无线电导航理论及技术,无线与移动通信关键技术,网络通信理论与技术等。

收稿日期:2008-08-07 修回日期:2008-10-31

$s_i(n)$ 与 $d_i(n)$ 相互独立,为避免分帧时的截断效应,应在变换到频域之前对 $y_i(n)$ 加窗处理<sup>[4]</sup>。用 $Y_i(\omega)$ 、 $S_i(\omega)$ 、 $D_i(\omega)$ 分别表示 $y_i(n)$ 、 $s_i(n)$ 、 $d_i(n)$ 的傅里叶变换,对式(2)等号两端进行傅里叶变换,可得

$$Y_i(\omega) = S_i(\omega) + D_i(\omega) \quad (3)$$

对式(3)求功率谱<sup>[5]</sup>,有

$$|Y_i(\omega)|^2 = |S_i(\omega)|^2 + |D_i(\omega)|^2 + S_i(\omega)D_i^*(\omega) + S_i^*(\omega)D_i(\omega) \quad (4)$$

由于 $s(n)$ 和 $d(n)$ 独立,它们的互谱统计独立,故原始语音的估计值为:

$$|\hat{S}_i(\omega)|^2 = |Y_i(\omega)|^2 - |\hat{D}_i(\omega)|^2 \quad (5)$$

式(5)中 $|\hat{S}_i(\omega)|$ 和 $|\hat{D}_i(\omega)|$ 分别是对 $|S_i(\omega)|$ 和 $|D_i(\omega)|$ 的估计。

式(4)和式(5)都是按帧计算的,如果式(5)出现负的情况,最简单的处理是直接令其为0,即

$$|\hat{S}_i(\omega)|^2 = \begin{cases} |Y_i(\omega)|^2 - |\hat{D}_i(\omega)|^2, & \text{如果 } |Y_i(\omega)|^2 > |\hat{D}_i(\omega)|^2 \\ 0 & \text{否则} \end{cases} \quad (6)$$

由于人耳对语音信号的相位不敏感, $|\hat{D}_i(\omega)|^2$ 可以在无音段估计得到。带噪语音的相位直接与 $|\hat{S}_i(\omega)|^2$ 相乘恢复出增强后的语音,即

$$\hat{s}(n) = \text{IFFT}[|\hat{S}_i(\omega)| \cdot \exp(j \arg Y(\omega))] \quad (7)$$

式(7)中IFFT是傅里叶反变换, $\arg(Y_i(\omega))$ 为带噪语音的相位谱。

## 2 改进的谱减法

### 2.1 时域中语音平滑

在坦克环境中,语音通信不仅受到强机械噪声的影响,同时还受到例如枪声、炸弹的爆炸声等强随机噪声的影响,而这种噪声的很大特点就是具有脉冲特性,所以必须在时域中对其进行消除,这更有益于后继的降噪处理。

因为语音信号的短时平稳性,相邻的几帧语音信号都保持很强的相关性,而噪声具有随机性,所以可以用连续3帧对应点上的平均值代替当前帧所对应点上的值,即

$$\hat{y}_i(n) = \frac{y_{i-1}(n) + y_i(n) + y_{i+1}(n)}{3}, i=0, 1, \dots; n=0, 1, \dots, N-1 \quad (8)$$

### 2.2 频谱中语音的处理

式(6)显示,噪声谱的估计对整个算法效果的优劣至关重要。一般的做法是 $|D_i(\omega)|^2$ 在带噪语音的无声段用多帧统计平均值作为噪声谱的估计,同时引入有声/无声检测(采用一种既简单又准确的谱熵检测<sup>[6]</sup>),在被判为噪声帧时对估计的噪声谱进行更新<sup>[7]</sup>,即

$$|\hat{D}_i(\omega)|^2 = (1-\beta)|\hat{Y}_i(\omega)|^2 - \beta|\hat{D}_{i-1}(\omega)|^2, i=1, 2, \dots; 0 < \beta < 1 \quad (9)$$

$|\hat{Y}_i(\omega)|^2$ 为式(8)傅里叶变换。如果被判为语音帧则不作更改,保持原来估计的噪声谱。

根据语音信号失真及残留噪声对语音的影响,设定一个增益函数 $G_i(\omega)$ ,其表达式为:

$$G_i(\omega) = \frac{|\hat{Y}_i(\omega)| - a|\hat{D}_i(\omega)|}{|\hat{Y}_i(\omega)|}, i=1, 2, \dots; a \geq 0 \quad (10)$$

$$T = \frac{b\hat{D}_i(\omega)}{\hat{Y}_i(\omega)}, i=1, 2, \dots; b > 0 \quad (11)$$

其中, $T$ 为门限值, $a$ 为过减因子, $b$ 为增益补偿因子。

$$G_i(\omega) = \begin{cases} G_i(\omega) & \text{当 } G_i(\omega) > T, i=1, 2, \dots \\ ((G_i(\omega))^\lambda) & \text{否则 } \lambda > 0 \end{cases} \quad (12)$$

式中, $\lambda$ 为平滑因子,经过实验表明,此因子选择得当能很好地抑制噪声。则所得到的纯净语音信号为:

$$\hat{S}_i(\omega) = G_i(\omega) \cdot \hat{Y}_i(\omega), i=1, 2, \dots \quad (13)$$

### 2.3 “音乐噪声”的消除

经过以上的处理后,背景噪声将会得到极大的消除,语音的质量也会得到很大的改善,但是经过试听发现还存在一定的“音乐噪声”,所以有必要对其进一步的处理,达到语音质量再次提高。根据语音信号的短时平稳性和噪声的随机性,对语音进行进一步的处理,即

$$\hat{S}'_i = \sum_{j=-M}^{j=M} P_j |\hat{S}_{i-j}|, i=1, 2, \dots; -M \leq j \leq M \quad (14)$$

式中, $i$ 为帧序号, $j$ 为叠加帧序号, $P_j$ 为权值,且 $P_{-M} + P_{-M+1} + \dots + P_{M-1} + P_M = 1$ 。为了减少语音的失真和有效地抑制音乐噪声, $M$ 一般取2~3, $P_j$ 在靠近 $i$ 取值稍大一点较好。将式(14)带入式(7)就能得到最终的处理结果。

## 3 仿真实验

实验采用真实的录制于坦克驾驶舱的带噪语音“众人皆醉我独醒”,8kHz采样,16bit量化。对带噪语音采用汉明窗进行分帧,每帧512个采样点,帧间叠加128个采样点。式(9)中的参数 $\beta=0.98$ ;式(10)中的参数 $a=5$ ;式(11)中的参数 $b=0.01$ ;式(12)中的参数 $\lambda=2$ ;式(14)中的参数 $M=2, P=[0.08 \ 0.25 \ 0.34 \ 0.25 \ 0.08]$ 。

图1、图2和图3分别显示了原始带噪语音、基本谱减法和该文改进算法的时域波形。

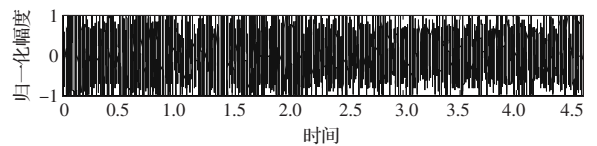


图1 原始带噪语音

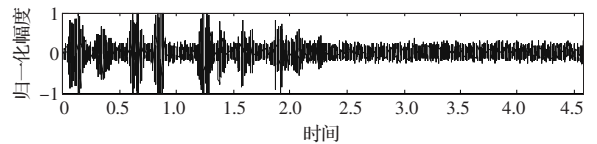


图2 基本谱减法处理结果

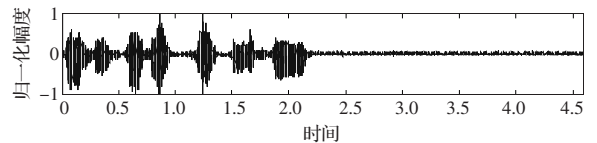


图3 该文改进方法处理结果

从图1~图3中比较可以看出,改进谱减算法比基本谱减法在信噪比的提高方面是相当明显的。同时经主观试听也表明:采用基本谱减法对含噪语音进行增强后虽然噪声已明显减少,但产生了有节奏的音乐噪声;而采用改进谱减算法不仅噪声有明显减少,并且音乐噪声也得到了极大的抑制,语音的清晰度和可懂度也较好。

## 4 结束语

由于环境复杂多变,语音增强方法虽说也在不断地更新,

(下转 138 页)