

非平稳环境下基于人耳听觉掩蔽特性的语音增强*

刘海滨¹ 吴镇扬¹ 赵力¹ 曾毓敏²

(1.东南大学无线电系, 南京 210096; 2.南京师范大学物理系, 南京 210097)

摘要: 传统的语音增强算法往往仅对平稳噪声或缓慢变化的噪声有效, 且残留的音乐噪声较大。对此, 本文研究了一种非平稳环境下基于听觉掩蔽效应的语音增强算法。该算法对传统谱减法的功率谱估计算法进行改进, 根据最小均方误差原则和语音信号的听觉掩蔽阈值调整功率谱估计的参数, 并引入了基于最小值统计特性的噪声估计算法, 使估计的噪声更好地跟踪噪声的变化。实验结果表明: 该算法对平稳和非平稳的噪声都得到较好的增强效果, 且较好地抑制了音乐噪声。

关键词: 语音增强; 听觉掩蔽阈值; 非平稳性; 音乐噪声

Speech Enhancement based on Human Auditory Masking Properties under Non-stationary Environments

Liu Haibin¹ Wu Zhenyang¹ Zhao Li¹ Zeng Yumin²

(1.Dept. of Radio Engineering, Southeast University, Nanjing, 210096, China)

(2.Dept. of Physics, Nanjing Normal University, Nanjing, 210097, China)

Abstract: This paper addresses the problem of single channel speech enhancement under stationary and non-stationary environments, which based on the masking properties of human auditory system. This algorithm can overcome the deficiency of the conventional speech enhancement algorithms, which were only efficient for stationary environments and have large level of musical residual noise. During the estimation of power spectrum of the speech, the parameters of the estimator can be modified by the MMSE and the masking threshold of the speech, by this way, we can find the best trade off among the amount of noise reduction, the speech distortion and the level of musical residual noise. For the best tracking the variation of the environment, the method of minimum statistics was introduced for noise power spectrum estimation. Objective and subjective evaluation of the proposed algorithm is performed with several noise types in the Noisex-92 database with different time frequency distributions. The evaluations confirm that the enhanced speech by proposed algorithm is more pleasant to a human listener for every noise conditions.

Key words: speech enhancement; auditory masking threshold; non-stationary; musical noise

1 引言

在语音通信技术中, 语音增强技术得到了极为广泛的应用。国内外的许多学者对语音增强算法进行了研究^[1-7], 在平稳的声学环境及信噪比较大的情况下得到了较好的语音增强效果, 但在非平稳的声学环境下, 低信噪比的含噪语音信号的增强仍然是一项具有挑战性的工作。由于噪声特性各

异, 语音增强算法的方法也各不相同, 目前常用的基于短时幅度谱估计的方法有减谱法、维纳滤波法和最小均方误差估计(MMSE)的方法等, 其中减谱法和维纳滤波法尽管导出的出发点不同, 但形式上并没有太大差别, 可以认为是统一的。减谱法和维纳滤波法的缺点是完全放弃了对语音频谱的分布假设, 而 MMSE 方法则有效地解决了这一缺点, 但由于需要事先估计或假设语音频谱地分布, 因而其计算量较

收稿日期: 2002年12月9日; 修回日期: 2003年2月28日

*国家自然科学基金(合同号: 60272044)资助项目

大。减谱法及其变形方法总体上看运算量较小,易于实时实现,增强效果也较好,是目前常用的一类方法。

S.Boll^[1] 假设噪声是平稳的或变化缓慢的加性噪声,并且语音信号与噪声信号不相关的情况下,提出了谱减法(SS: Spectral Subtraction),该方法能够抑制背景噪声的影响,但由于其局部平稳性的假设与实际情况并不相符,因此效果不理想,残留的音乐噪声较大;Berouti^[2] 在传统谱减法的基础上增加了调节噪声功率谱大小的系数和增强语音功率谱的最小值限制,提高了谱减法的性能,但其修正系数和最小值是根据经验确定的,适应性较差;P. Lockwood & J. Boudy^[3] 在谱减法的基础上提出了非线性谱减法(NSS: Nonlinear Spectral Subtraction),它根据语音信号的信噪比自适应调整语音增强的增益函数,提高了语音的信噪比,众所周知,信噪比不能正确反映语音的听觉质量,因此,以信噪比作为调整估计参数的依据并不能提高信号的听觉质量;Boh Lim Sim^[4] 等人也提出了与此相似的改进算法,虽然提高了信号的信噪比,但残留的音乐噪声较大。Nathalie Virag^[5] 将人耳的掩蔽特性应用到非线性谱减法语音增强算法中,部分解决了谱减法残留音乐噪声大的问题,但在信噪比较低或非平稳噪声的情况下,其增强效果不理想。目前,对非平稳环境下的语音增强算法研究还较少。I. Cohen^[6] 等人首先估计语音信号概率密度分布函数,然后在此基础上改进了对数谱估计算法^[7],使得改进的算法对非平稳的噪声具有良好的抑制作用,该算法的缺点是语音信号的概率密度函数较难估计。

以上的语音增强算法主要缺点是对非平稳噪声适应性较差且残留音乐噪声比较大,针对以上存在的缺点,本文提出了一种基于听觉掩蔽效应和最小值统计特性的语音增强算法,该算法不但对非平稳的环境噪声有很好的抑制作用,亦能有效地抑制增强语音的残留音乐噪声。

2 算法描述

根据加性噪声模型,含噪语音信号一般可以写为:

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

式中, $x(n)$ 是纯净语音信号, $d(n)$ 是加性噪声,且假设 $x(n)$ 和 $d(n)$ 两者不相关。由于增强是按帧进行的,所以可把上述模型写成帧的形式:

$$y(m, n) = x(m, n) + d(m, n) \quad (2)$$

$$m = 1, 2, 3, \dots; n = 0, 1, \dots, N-1$$

式中, m 为帧号, N 为帧长。对(2)进行傅立叶变换得:

$$Y(m, k) = X(m, k) + D(m, k) \quad (3)$$

式中, k 是离散频率, $Y(m, k)$ 、 $X(m, k)$ 和 $D(m, k)$ 分别是含噪语音、纯净语音和噪声的傅立叶变换系数,其幅度谱分别记为 $Y_{m,k}$ 、 $X_{m,k}$ 和 $D_{m,k}$ 。

设增强的语音信号幅度的谱估计值为 $\hat{X}_{m,k}$, 我们定义了 $\hat{X}_{m,k}$ 的参数化估计公式,如(4)式所示:

$$\hat{X}_{m,k}^\gamma = \alpha(m, k) (Y_{m,k}^\gamma - \beta(m, k) E[D_{m,k}^\gamma]) \quad (4)$$

式中, m, k 分别是帧号和频率下标, $E[D_{m,k}^\gamma]$ 表示估计噪声的均值, γ 是常数指数,一般取 $1 \sim 2$, 在本文中取 $\gamma = 2$ 。 $\alpha(m, k)$ 和 $\beta(m, k)$ 是时间和频率的函数,其中我们在幅度谱的估计式中,系数 $\alpha(m, k)$ 的作用类似于 Crozier^[8] 提出的共振峰滤波器在各个频点的权重系数,这样可以控制增强语音的残留音乐噪声,其值根据人耳的听觉掩蔽特性确定;而系数 $\beta(m, k)$ 用以控制增强语音的信噪比,其值根据最小均方误差原则确定。

首先定义如下的误差函数:

$$\delta(m, k) = X_{m,k}^\gamma - \hat{X}_{m,k}^\gamma \quad (5)$$

由于我们不可能知道真实的纯净语音的功率谱,因此,我们可用一个理想的广义谱减法模型表示,即假设语音信号与噪声信号完全独立,则有下式成立:

$$Y_{m,k}^\gamma = X_{m,k}^\gamma + D_{m,k}^\gamma \quad (6)$$

在上式中,若 $\gamma = 2$,则此时求得的参数 $\alpha(m, k)$ 和 $\beta(m, k)$ 在均方意义上是最优的;如果 $\gamma \neq 2$,则(6)在严格意义上并不相等,此时求得的参数也不是最优的,但用于语音增强时,仍能得到较好的效果。将(4)式和(6)式代入(5)式得:

$$\delta(m, k) = (1 - \alpha(m, k)) X_{m,k}^\gamma - \alpha(m, k) D_{m,k}^\gamma + \alpha(m, k) \beta(m, k) E[D_{m,k}^\gamma] \quad (7)$$

根据(7)式,可得最小均方误差函数为:

$$E[\delta(m, k)^2] = (1 - \alpha(m, k))^2 E[X_{m,k}^{2\gamma}] + \alpha^2(m, k) E[D_{m,k}^{2\gamma}] + \alpha^2(m, k) (\beta^2(m, k) - 2\beta(m, k)) E[D_{m,k}^\gamma]^2 + 2(\alpha(m, k) - \alpha^2(m, k)) (\beta(m, k) - 1) E[X_{m,k}^\gamma] E[D_{m,k}^\gamma] \quad (8)$$

为了简化(8)式,我们进一步假设 $X(m, k)$ 和 $D(m, k)$ 分别是统计独立的且是零均值的复高斯型随机变量,则 $X_{m,k}$ 和 $D_{m,k}$ 服从瑞利分布^[4],因此可得:

$$E[D_{m,k}^\gamma] = \Gamma(\gamma/2 + 1) [E[X_{m,k}^2]]^{\gamma/2} \quad (9)$$

同理可得:

$$E[X_{m,k}^\gamma] = \Gamma(\gamma/2 + 1) [E[X_{m,k}^2]]^{\gamma/2} \quad (10)$$

其中, $\Gamma(\cdot)$ 表示 Γ 函数。令 $\frac{\partial E[\delta(m, k)^2]}{\partial \beta(m, k)} = 0$, 则由(8)、

(9)和(10)式可得到参数 $\beta(m, k)$ 的估计式如(11)式所示:

$$\beta(m, k) = 1 + \left(\frac{\alpha(m, k) - 1}{\alpha(m, k)} \right)^{\frac{2}{\gamma}} \quad (11)$$

其中, $\xi(m, k) = E[X_{m,k}^2] / E[D_{m,k}^2]$, 称为语音信号的先验信噪比, 后文将给出其估计式。

然后再根据听觉掩蔽特性确定 $\alpha(m, k)$ 的估计式。根据

(7) 式, 可得增强语音与纯净语音的误差为:

$$E[\delta(m, k)] = \Gamma(\gamma/2 + 1) E[D_{m,k}^2] \{ \xi(m, k) - \alpha(m, k) \rho(m, k) + \alpha(m, k) \beta(m, k) \} \quad (12)$$

式中 $\rho(m, k) = E[Y_{m,k}^2] / E[D_{m,k}^2]$ 称为后验概率。由 (12) 式求得的误差实际就是残留噪声 (包括音乐噪声和语音失真), 为了进一步抑制该噪声, 使残留噪声不被人耳觉察, 我们引入了人耳的听觉掩蔽特性, 则基于掩蔽效应的语音增强准则为:

$$E[\delta(m, k)] \leq T(m, k) \quad (13)$$

式中, $T(m, k)$ 为语音信号的掩蔽值。由 (11)、(12) 和 (13) 式可得:

$$\alpha(m, k) = \frac{T(m, k)}{\Gamma(\gamma/2 + 1) \{ \xi(m, k) - \rho(m, k) + 1 \} E[D_{m,k}^2]} \quad (14)$$

将 $\alpha(m, k)$ 和 $\beta(m, k)$ 按照不同准则进行估计, 可以使语音增强算法在增强语音的信噪比和语音信号的失真之间得到较好的折衷, 使增强的语音既具有较高的信噪比, 又具有较好的听觉质量。

在实际的环境中, 噪声往往是非平稳的。为了适应噪声的非平稳特性, 在估计噪声时不应仅局限于静音段, 最优的方法是按帧估计噪声。通过分别观察含噪语音和噪声的功率谱, 发现含噪语音的功率谱是以较快地速度衰减到噪声功率谱的水平。因此, 我们可以通过估计含噪语音功率谱的最小值并对其进行修正的方法来估计噪声的功率谱。该算法的主要优点是能较快地跟踪噪声的变化且无需进行静音段的判断, 提高了算法的鲁棒性。首先按照 (15) 式对含噪语音信号的功率谱进行平滑处理:

$$P(m, k) = \lambda P(m-1, k) + (1-\lambda) |Y(m, k)|^2 \quad (15)$$

式中, $P(m, k)$ 表示平滑后的幅度谱, λ 为平滑系数, 一般取 0.94~0.98。然后在长为 τ 的时间窗内, 找出平滑后功率谱的最小值 $P_{\min}(m, k)$, 每输入一帧语音, 都更新此最小值。考虑到该最小值和真实噪声之间的偏差, 我们采用文献^[9]的偏差补偿算法, 通过该算法对 $P_{\min}(m, k)$ 进行补偿, 即可得到噪声的估计值 $E[D_{m,k}^2]$ 。通过对多种噪声类型进行实验分析, 当 τ 取 0.8s 时, 可得到较好的实验效果。

估计出噪声后, 语音信号的后验信噪比 $\rho(m, k) = E[Y_{m,k}^2] / E[D_{m,k}^2]$ 可直接计算。由于不能求得纯净语音信号的功率谱, 因此先验信噪比 $\xi(m, k)$ 的估计由 (16) 式确定^[7]:

$$\xi(m, k) = \eta \xi(m-1, k) + (1-\eta) \frac{\max(Y_{m,k}^2 - E[D_{m,k}^2], 0)}{E[D_{m,k}^2]} \quad (16)$$

式中, $0 < \eta < 1$, 称为平滑系数。

3 掩蔽阈值的计算

掩蔽阈值在感知音频编码中获得了广泛的应用^[10], 本文提出的语音增强算法中引入信号的掩蔽特性, 可以减小增强语音的残留音乐噪声的影响。由于不可能得到纯净的语音信号, 因此, 在计算语音信号的掩蔽阈值时, 首先利用传统的谱减法^[1]近似估计语音信号, 然后利用估计的语音信号计算其掩蔽阈值。

首先将语音信号进行临界频带分析或 Bark 分析。利用 FFT 求得语音信号功率谱 $E(m, k)$, 将语音信号的频率域划分为不同的临界频带, 语音信号在各个临界带的能量由 (17) 式求得:

$$B_i = \sum_{k=bl_i}^{bh_i} E(m, k) \quad (17)$$

其中, B_i 表示各个临界频带的能量, bl_i 和 bh_i 分别表示临界频带 i 的上限和下限, $i = 1, 2, \dots, i_{\max}$ 是临界频带号, i_{\max} 的取值与采样频率有关。考虑到各个频带间相互掩蔽的影响, 定义如下的传播函数^[10]:

$$SF_{ij} = 15.81 + 7.5(\Delta + 0.474) - 17.5\sqrt{1 + (\Delta + 0.474)^2} \text{ dB} \quad (18)$$

式中 $\Delta = i - j$ 表示两个频带的频带号之差, $i, j = 1, 2, \dots, i_{\max}$ 且 $|\Delta| \leq i_{\max}$ 。则考虑频带间的影响后的能量为:

$$C_j = \sum_{i=1}^{i_{\max}} B_i * SF_{ij}, \quad j = 1, 2, \dots, i_{\max} \quad (19)$$

由于噪声和音调的掩蔽特性不同, 因此, 首先判断各个频带是噪声特性还是音调特性。可以根据谱平坦度度量 (SFM: Spectral Flatness Measure) 判断, 其定义如下:

$$SFM = \frac{\mu_j}{\mu_a} \quad (20)$$

其中, μ_g 和 μ_a 分别表示各个频带功率谱的几何平均和算术平均, 且 $SFM \in [0, 1]$, 其值为 0 时, 表示该频带的信号具有纯音调特性, 纯音调的掩蔽阈值偏移量为 $(14.5 + i) \text{ dB}$, 为 1 时, 具有白噪声特性, 白噪声的掩蔽阈值偏移量为 5.5 dB 。根据 (20) 式, 定义如下的音调系数:

$$\phi = \min\left(\frac{SFM_{\text{dB}}}{-60}, 1\right) \quad (21)$$

则相对的掩蔽阈值的偏移量为:

$$O_i = \phi(14.5 + i) + 5.5(1 - \phi) \text{ dB} \quad (22)$$

确定了相对的掩蔽阈值的偏移量后, 则掩蔽阈值由下式

计算:

$$T(m,i)=10^{\log_{10}(C_i)-(O_i/10)}$$

(23)

在各个临界频带上, 语音信号具有相同的掩蔽特性, 即其掩蔽阈值相等。将 $T(m,i)$ 扩展到各个频谱上, 记为 $T'(m,k)$ 。最终的掩蔽阈值由下式确定:

$$T(m,k)=\max(T'(m,k),T_a(m,k))$$

(24)

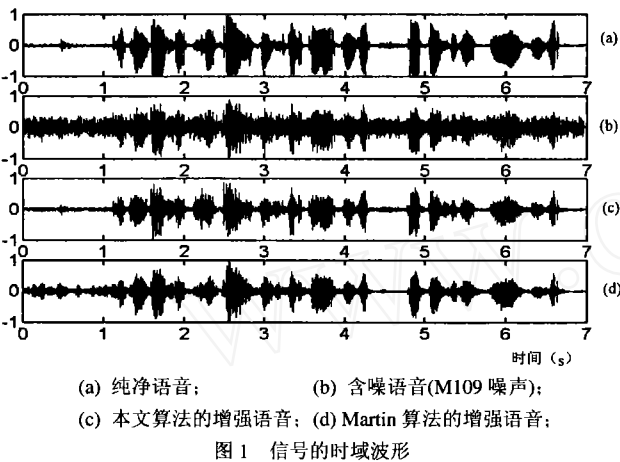
式中, $T_a(m,k)$ 是绝对听阈, 定义如 (25) 式:

$$T_a(m,k)=3.64f^{-0.8}-6.5\exp(f-3.3)^2+10^{-3}f^4$$

(25)

其中, f 和 $T_a(m,k)$ 的单位分别是 kHz 和 dB SPL (Sound Pressure Level)。

4 实验结果及性能评价



根据上文的分析, 对所提出的语音增强算法进行性能评测, 并与 R. Martin 提出的语音增强算法^[9]进行比较。D. L. Wang^[11]指出相位对语音信号的感知不重要, 因此可以利用含噪语音信号的相位代替增强语音信号的相位, 并根据 (4) 式估计出语音信号的幅度谱, 然后利用逆傅立叶变换, 即可得到增强的语音信号。

在本文中, 所用的试验语音数据的采样率为 16kHz , 帧长为 512, 重叠 $1/2$, 数据窗是汉明窗, 计算掩蔽阈值时的临界频带数为 22, 通过实验分析, η 取 0.96。在实验过程中, 选用的噪声取自 Noisex-92^[12]数据库, 分别是白噪声 (White), 似语音噪声 (Babble), 某一厂房的噪声 (Factory) 和 M109 坦克的噪声 (M109)。按照 0dB 的信噪比, 将上述的噪声数据和干净的语音数据 (在安静的实验室条件下录得的语音数据) 混合, 得到所需的含噪语音数据。图 1 分别是纯净语音、含噪语音和各种算法的增强语音的波形图。

语音增强算法的性能评测采用客观测试结合非正式听音测试来进行。采用的客观测试指标包括全局信噪比 (SNR) 和 Itakura-Saito 语音失真测度 (IS), SNR 和 IS 的主要区别是 SNR 不能反映人们对语音信号的听觉质量而 IS 语音失真

测度在一定程度上可以反映人们的对语音信号的主观感受^[13], IS 系数越小, 说明语音信号的品质越好。实验结果如表 1 所示。可以看出, 本文所提出的算法的 IS 都远远小于 Martin 所提出的算法, 而 Martin 所提出的算法的增强语音的信噪比较大。这是因为, 在本文的算法中, 我们引入了人耳的听觉掩蔽特性, 对参数的优化不仅考虑了信噪比, 而且考虑了人耳的听觉掩蔽特性, 从而使得增强语音的信噪比略有下降, 但其主观品质却大为提高。

表 1 算法比较

		White	Factory	M109	Babble
Itakura-Saito 失真测度	I	1.32	0.68	0.47	0.52
	II	4.14	3.30	1.44	1.90
	III	5.55	2.58	1.70	1.37
全局信噪比 SNR	I	4.48	2.99	2.89	2.81
	II	6.51	3.20	3.37	2.72
	III	0	0	0	0

I: 本文算法的增强语音; II: Martin 算法的增强语音; III: 含噪语音;

客观测试的缺点是不能很好地反映语音信号的感知质量, 同时也不能反应语音信号和噪声信号在频率上的分布情况。对含噪语音

进行增强时, 往往会引入音乐噪声和语音失真, 常常会引起人们听觉的不适, 客观性指标不能刻画音乐噪声的大小和语音失真的大小。因此, 在本文的实验中, 我们将分析增强语音信号的语谱图, 从而在直观上了解音乐噪声和语音的失真大小。图 2 是纯净语音的语谱图, 图 3、4、5、6 分别是在各种噪声情况下增强语音和含噪语音的语谱图。从各个语谱图可以看出, 本文提出的算法的增强语音在时频分布上接近于纯净的语音,

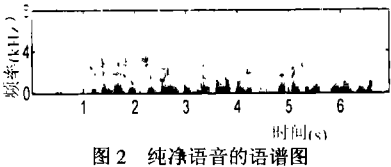


图 2 纯净语音的语谱图

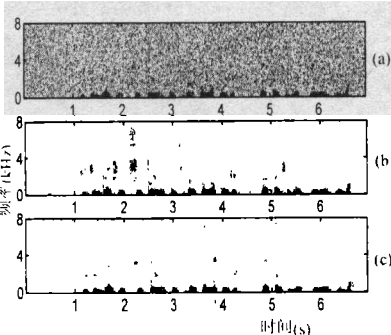


图 3 (a) 含噪语音的语谱图(White 噪声)
(b) 本文算法的增强语音语谱图
(c) Martin 算法的增强语音语谱图

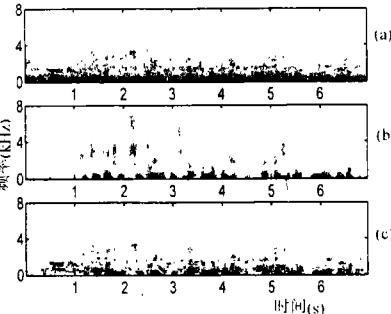


图 4 (a) 含噪语音的语谱图(Babble 噪声)
(b) 本文算法的增强语音语谱图
(c) Martin 算法的增强语音语谱图

而 Martin 的算法的残留噪声较大。

最后,对增强语音进行非正式听音测试表明,本方法增强语音的残留噪声在听觉上比 Martin 所提出的算法要小得多,这与 IS 失真测度的测量评测结果一致,部分证实了文^[13]中提到的 IS 失真测度与音质的主观评测有较高度相关性的结论。需要指出的是,尽管本方法的增强语音的全局信噪比稍低于 Martin 的算法,非正式听音测试却表明其增

强语音的音质比它好,这说明全局信噪比的大小与主观评测并不一致。总之,客观测试和非正式听音测试表明,在各种环境情况下,本文基于听觉掩蔽效应的语音增强方法能在听觉上更好地抑制语音中的音乐噪声。音乐噪声的出现既与噪声谱的估计有关,又与语音谱的估计有关。在本文的实验中, Babble 噪声情况下产生的音乐噪声较大,这是由于 Babble 噪声较难估计的缘故。由本文的实验可知,合理地设定语音功率谱的估计准则,并准确地估计噪声谱,是可以减小音乐噪声影响的。

在计算量方面,和常规谱减法相比,本文的算法需要估计听觉掩蔽阈值,并且噪声是按帧估计的,因此计算量有所增加。但在估计听觉掩蔽阈值和噪声估计时,主要是进行频谱分析,而频谱分析可以通过快速傅立叶变换实现,因此本文的计算量是可以接受的。

5 结论

由于环境的复杂多变,语音增强技术是一项具有挑战性的课题。任何增强算法都有其优点和局限性。本文所提出的算法中,提出一个功率谱估计算法,并根据语音信号的信噪比和其听觉掩蔽特性自适应地调整其参数,从而使得增强语音在保证一定的信噪比的情况下,大大减小了增强语音的残留音乐噪声,提高了增强语音的感知质量。

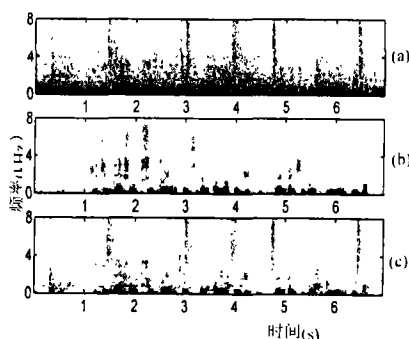


图5 (a) 含噪语音的语谱图(Factory 噪声)
(b) 本文算法的增强语音语谱图
(c) Martin 算法的增强语音语谱图

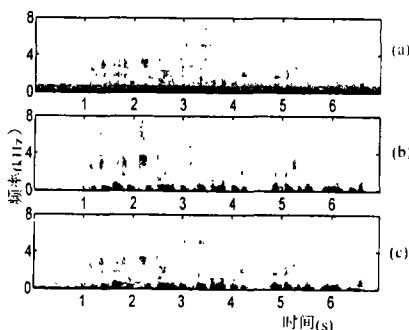


图6 (a) 含噪语音的语谱图(M109 噪声)
(b) 本文算法的增强语音语谱图
(c) Martin 算法的增强语音语谱图

参考文献

- [1] S. Boll. Suppression of acoustic noise in speech using spectral subtraction. *IEEE Transactions on Acoustics, Speech, and Signal Processing*. 1979 (2): 113-120.
- [2] M. Berouti, R. Schwartz, J. Makhoul. Enhancement of speech corrupted by acoustic noise. *Proc. IEEE ICASSP*, Washington, DC, Apr. 1979; 208-211.
- [3] P. Lockwood, J. Boudy. Experiments with a nonlinear spectral subtractor(NSS), hidden Markov models and projection for robust recognition in cars. *Speech Communication*. 1992; 11: 215-228.
- [4] Boh Lim Sim, Yit Chow Tong etc.. A parametric formulation of the generalized spectral subtraction method. *IEEE Transaction on Speech and Audio Processing*. 1998; 6(4): 328-337.
- [5] Nathalie Virag. Single channel speech enhancement based on masking properties of human auditory system. *IEEE Transactions on Speech and Audio Processing*. 1999; 7(2): 126-137.
- [6] I. Cohen, B. Berdugo. Speech enhancement for nonstationary noise environments. *Signal Processing*. 2001; 81: 2403-2418.
- [7] Y. Epharim, D. Malah. Speech enhancement using a minimum mean square log-spectral amplitude estimator. *IEEE. Transactions on Acoustics, Speech, and Signal Processing*. 1984; 32(6): 1109-1121.
- [8] P. M. Crozier, B.M.G. Cheetham etc. Speech enhancement employing spectral subtraction and linear predictive analysis. *Electronics Letters*. 1993; 29 (12):1094-1095.
- [9] R. Martin. Noise power spectral density estimation based on optimal smoothing and minimum statistics. *IEEE. Transactions on Speech and Audio Processing*. 2001; 9(5): 504-512.
- [10] T. Painter, A. Spanias. Perceptual coding of digital audio. *Proc. Of the IEEE*. 2000; 88(4): 451-512.
- [11] D. L. Wang, J. S. Lim. The unimportance of phase in speech enhancement. *IEEE. Transactions. Acoustics, Speech, and Audio Processing*. 1982; 30(8): 679-681.
- [12] http://spib.rice.edu/spib/select_noise.html.
- [13] 杨行峻,迟惠生.语音信号数字处理. 北京: 电子工业出版社, 1999.

作者简介

1974 年生, 博士生。研究方向为语音增强、噪声环境下的语音识别等。