

图 5 纯高斯白噪声 N

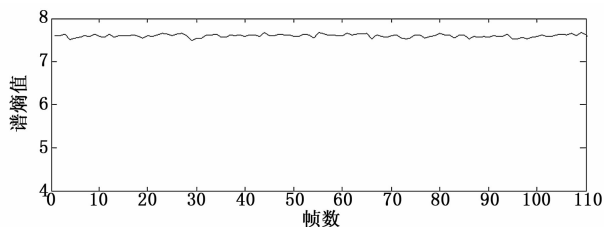


图 6 N 的谱熵值

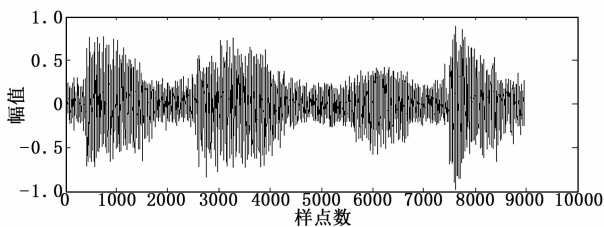


图 7 加噪语音 S₁

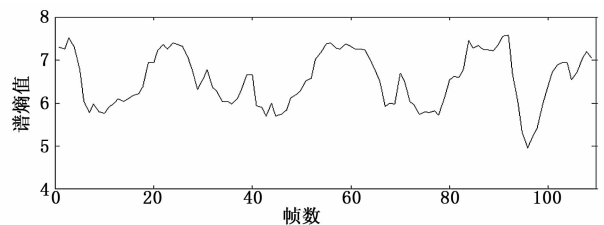


图 8 S₁ 的谱熵值

都比较小，在 5.8 上下波动，静音帧的谱熵值较大，在 7.0 之上波动；从图 6 和图 7 可以看出，纯高斯白噪声的谱熵值较大，在 7.5 之上波动；从图 8 和图 9 可以看出，在含噪语音中，语音帧和噪声帧的谱熵值分别在 6.0 和 7.2 左右。因此可据此将谱熵值最大的一帧作为噪声帧，并根据各帧的谱熵变化来确定并动态的更新噪声值。

实验二：低噪环境中 GMM 与 GMM-UBM 模型对比：

在未加噪情况下，采用未加入谱减法的 MFCC 作为特征参数，对比 GMM 与 GMM-UBM 模型识别效果，分别取 GMM 混合度为 8, 16, 32, 64 进行实验，得出识别率如下表：

表 1 不同混合度的 GMM 模型与 GMM-UBM 模型对比

识别模型	识别率/%	
不同混合度的 GMM 模型	8	75.45
	16	82.87
	32	88.32
	64	82.43
GMM-UBM 模型	97.48	

由表可见，在基于 GMM 模型中，混合度为 32 时的识别性能最好，而 GMM-UBM 模型识别性能远高于 GMM 模型。

实验三：不同参数在不同噪声环境中识别率对比：

分别在 SNR=10dB、5dB、0dB、-5dB 环境下，对未进行语音增强的 MFCC、采用基本谱减法 (SS) 得到的 MFCC 和采用改进谱减法 (ISS) 得到的 MFCC 进行说话人识别实验，得到识别结果如下：

表 2 不同信噪比下 3 种参数识别率 %

SNR/dB 特征参数	10	5	0	-5
MFCC	95.24	83.47	65.39	38.93
MFCC+SS	96.35	86.33	73.29	56.32
MFCC+ISS	96.83	91.58	84.71	75.43

由表可见，在相同信噪比环境下，基本谱减法所提取的 MFCC 性能好于未进行语音增强的 MFCC，而本文提出的改进谱减法提取的 MFCC 性能优于基本谱减法所提取的 MFCC。随着信噪比的降低，各方法提取的特征参数识别率均有所下降，本文提出的改进谱减法所提取的 MFCC 对识别性能的改善更为明显，在 SNR=-5dB 的环境中也能达到 75% 以上的识别率，较基本谱减法所提取的 MFCC 在识别率方面有 30% 以上的提高。

5 结束语

本文提出了改进谱减法用以改善加性噪声降低说话人识别系统性能的问题，首先介绍了传统的 MFCC 提取算法，随后对语音增强中的基本谱减法进行改进，并仿真实现了改进谱减法的 MFCC 提取。基于 GMM-UBM 系统经加噪测试实验，结果表明改进谱减法提取得到的 MFCC 较基本谱减法具有更强的鲁棒性，在低信噪比环境中有更好的表现。此外，采用 GMM-UBM 计算量较大，耗时较长，如何减少识别系统的计算量，提高识别效率将是接下来的研究工作。

参考文献：

- [1] 徐子豪, 张腾飞. 基于语音识别和无线传感网络的智能家居系统设计 [J]. 计算机测量与控制, 2012, 01: 180-182.
- [2] 张雪英. 数字语音处理及 MATLAB 仿真 [M]. 北京: 电子工业出版社, 2011.
- [3] Yu G H, Zhao Y L, Wei Z X. A descent nonlinear conjugate gradient method for large-scale unconstrained optimization [J]. Applied Mathematics and Computation, 2007, 187 (2): 636-642.
- [4] 曾祺, 甘涛, 曾红斌. 改进的多窗谱 MFCC 在说话人确认中的应用 [J]. 计算机系统应用, 2014, 23 (11): 192-195.
- [5] 王路露, 夏旭, 冯璐, 等. 基于频谱方差和谱减法的语音端点检测新算法 [J]. 计算机工程与应用, 2014, 50 (8): 194-197.
- [6] 李晔, 张仁智, 崔慧娟, 等. 低信噪比下基于谱熵的语音端点检测算法 [J]. 清华大学学报 (自然科学版), 2005, 45 (10): 1397-1400.
- [7] 周国鑫, 高勇. 基于 GMM-UBM 模型的说话人辨识研究 [J]. 无线电系统, 2014, 44 (12): 14-17.
- [8] 杨海燕, 景新幸, 曾招华. 基于 DSP 开发板的语音识别系统的研究 [J]. 计算机测量与控制, 2013 (01): 210-212, 220.