基于多窗谱估计和几何谱减的低信噪比语音增强方法

李淯 1,2,3，胡俊 1,2，刘新 2,3,4, 黄石磊 2

(1 重庆邮电大学 光通信与网络重点实验室，重庆 400065；2 深港产学研基地 深圳市智能媒体和语音重点实验室 广东 深圳 518057；3 北京大学 深圳研究院，广东 深圳 518057；4 安徽大学 计算智能与信号处理教育部重点实验室，安徽 合肥 230601)

摘要： 在低信噪比条件下，传统谱减算法对语音和噪声的独立性假设会产生一定的“音乐噪声”，从而降低了语音的可懂度.对此，本文提出了一种结合多窗谱估计和几何谱减的语音增强算法.本算法利用多窗谱估计带噪语音信号的功率谱，采用改进最小值控制递归平均算法实时跟踪估计噪声谱，应用几何谱减算法求解增益函数，从而恢复出精确的语音信号.通过在 IEEE 数据集上进行实验，并以 PESQ 和 LSD 作为评价指标，结果表明本算法在低信噪比条件下，缩短了带噪语音与纯净语音之间的频谱距离，并有效的抑制了增强语音的背景噪声，提高了语音可懂度.

关键词： 语音增强；多窗谱；改进的最小控制递归平均；几何谱减

The Low SNR Speech Enhancement Method Based on Multi-Taper Spectrum Estimation and Spectral Subtraction of Geometric

LI Xu1,2,3, HU Jun1,2, LIU Xin2,3,4, HUANG Shi-lei2

(1 Chongqing Key Laboratory of Optical Communication and Networks, Chongqing University of Posts and Telecommunications, Chongqing 400065, China; 2 IMSL Shenzhen Key Lab, PKU-HKUST Shenzhen Hong Kong Institution, Shenzhen 518057, China; 3 Peking University Shenzhen institute, Shenzhen 518057, China; 4 Key Laboratory of Intelligent Computing & Signal Processing, Ministry of Education, Anhui University, Hefei 230601,China;)

Abstract：The traditional spectral subtraction methods tend to produce“music noise”under low signal-to-noise ratio (SNR) due to the independence assumption of speech and noise. The intelligibility of speech is reduced greatly. In this paper, we proposed a speech enhancement method via integrating multi-taper spectrum (MTM) estimation and spectral subtraction of geometric (GA). Our method employs the MTM to estimate the power spectrum of the noisy speech, and the improved minima controlled recursive average method to track the estimated noise spectrum in real time. Furthermore, the GA is used to calculate the gain function and speech signal is recovered accurately. Our method can reduce the spectral distance between the noisy speech and the clean speech under the low signal-to-noise ratio prominently, and restrain the background noise of the enhanced speech effectively. Experiments on the IEEE dataset, PESQ and LSD as the evaluation metrics, show that our method improves the speech intelligibility significantly.

Key words： speech enhancement; multi-taper spectrum; improved minima controlled recursive average; spectral subtraction of geometric;

作者简介：

李淯男，（1993-），硕士研究生.研究方向为语音增强、语音端点检测.E-mail:loveli5991529@163.com.

胡俊男，（1990-），硕士研究生.研究方向为音频检索、语音识别.

刘新男，（1992-），硕士研究生.研究方向为语音标点预测、语音识别.

黄石磊男，（1979-），博士，副研究员，高级工程师.研究方向为语音信号处理.