基于改进语音特征与极限学习机的语音端点检测

罗　庆，包亚萍，俞　强

(南京工业大学 计算机科学与技术系，江苏 南京 211816)

摘　要：　语音端点检测（Voice Activity Detection,VAD），是指在给定语音信号帧中判别语音是否存在，鲁棒的 VAD 有助于提高语音应用的自动化效率，例如语音增强、说话人识别、助听器等.为了提高低信噪比下语音端点检测的精度以及效率，提出了一种新的语音特征—低频消噪能量（Low Frequency De-noising Energy，LFDE），将其应用于 VAD 中，并利用 LFDE 与现有的声学特征（梅尔频率倒谱参数、共振峰频率）结合训练极限学习机（Extreme Learning Machine, ELM）分类器.仿真实验发现，端点检测的精度与效率都有提高.

关键词：　低频消噪能量；梅尔倒谱参数；共振峰频率；极限学习机

Voice activity detection based on improved speech features

and extreme learning machine

LUO Qing，BAO Ya-ping，YU Qiang

(Department of Computer Science and Technology, Nanjing Tech University, Nanjing 211816, China)

Abstract：　Voice Activity Detection (VAD) refers to the determination of the existence of speech in a given speech signal frame. Robust VAD helps to improve the automation efficiency of speech applications, such as speech enhancement, speaker recognition, and hearing aids and so on. In order to improve the accuracy and efficiency of voice activity detection under low SNR, a new speech feature—Low Frequency De-noising Energy (LFDE) is proposed, which is applied to VAD and utilizes LFDE and existing acoustic features (Mel frequency cepstrum parameters, formants Frequency) combined with the Extreme Learning Machine (ELM) classifier.Simulation experiments show that the accuracy and efficiency of voice activity detection are improved.

Key words：　low frequency de-noising energy; mel-frequency cepstrumcoefficient; formant frequency; extreme learning machine

作者简介：

罗　庆　男，（1994-），硕士.研究方向为语音信号处理与 SOC 设计.E-mail:luoqingjiangnan@163.com.

包亚萍　女，（1965-），教授.研究方向为语音信号处理、认知无线电、图像处理、嵌入式设计.

俞　强　男，（1993-），硕士.研究方向为语音信号处理、麦克风阵列、嵌入式系统设计与应用.