

一种基于深度神经网络的基音检测算法

曹 猛, 李鸿燕, 赵蓉蓉

(太原理工大学 信息工程学院, 山西 太原 030024)

摘 要: 针对强噪声干扰环境下的基音检测这一挑战性问题, 提出了一种基于深度神经网络的基音检测算法. 该算法运用监督学习方法估计基音, 在混合语音数据中得到概率性的基音状态. 通过静态帧级声音特征训练, 采用深度神经网络模拟每个基音状态产生的后验概率, 最后执行维特比算法将基音状态连接成基音曲线. 实验结果表明: 提出的基音检测算法在不同强噪声甚至混响环境中具有很强的鲁棒性, 而且明显胜过其他的基音检测算法.

关键词: 基音检测; 深度神经网络; 监督学习; 维特比算法

A Pitch Detection Method Based on Deep Neural Network

CAO Meng, LI Hong-yan, ZHAO Rong-rong

(College of Information Engineering, Taiyuan University of Technology, Taiyuan 030024, China)

Abstract: Based on deep neural network (DNN), a pitch detection algorithm is proposed to solve the challenging problem of pitch detecting in strong noisy background. In this paper, we estimate pitch using supervised learning, where the probabilistic pitch states are directly learned from noisy speech data. The feedforward deep neural network is trained on static frame-level acoustic features to produce accurate probabilistic outputs of pitch states, which are then connected into pitch contours by Viterbi decoding. The experimental results show that the proposed algorithm are robust to different noise conditions and can even be applied to reverberant speech. The proposed approach outperforms other state-of-art pitch tracking algorithms.

Key words: pitch detection; deep neural network; supervised learning; Viterbi decoding

作者简介:

曹 猛 男, (1988-), 硕士研究生. 研究方向为盲语音分离, 语音识别.

李鸿燕 女, (1973-), 博士后, 副教授. 研究方向为盲信号处理, 模式识别. E-mail: tylihy@163.com.