

doi:10.1631/FITEE.1400323

**题目:** 基于无监督特征学习的语音情感识别方法

**目的:** 语音情感识别是人机交互的关键技术之一。同时,良好的情感特征对语音情感识别系统性能具有极大影响。目前的语音情感特征主要通过手工设计方法提取,对于其是否能够很好地刻画情感特性以及是否存在最优情感特征集,相关研究者并没有达成公认。所以有必要对语音情感特征提取进行进一步深入研究。

**创新点:** 提出一种基于数据驱动的非监督情感特征学习方法。该方法能够自动从非标注语音数据中学习产生与情感相关的特征映射函数,用于语音情感特征提取。

**方法:** 采用三种非监督学习算法( $K$ -均值聚类,稀疏自动编码器,稀疏受限玻尔兹曼机)从若干非标注语音块中学习产生与目标相关的特征提取器,继而对整个语音样本进行特征提取(卷积和池化),最后训练一个线性支持向量机对未知样本进行识别。同时对模型涉及的超参数(块大小和隐层结点数目)进行选择。

**结论:** 相对于传统原始特征,学习产生的特征具有一定的稀疏性并且对说话人及其他扰动因素具有一定鲁棒性。实验结果表明,尺寸较大的块和数量较多的隐层结点有助于提升系统性能(图4、5)。

**关键词组:** 语音情感识别;非监督特征学习;神经网络;情感计算