

doi:10.1631/FITEE.1700814

**题目：**鸡尾酒会问题的技术回顾、当前进展及未来挑战

**概要：**鸡尾酒会问题即在多人同时说话的场景下追踪并识别某一个特定说话人的语音。在自动语音识别技术大规模推广应用中，鸡尾酒会问题是亟待解决的关键问题之一。本文回顾了在过去 20 多年中针对鸡尾酒会问题提出的相关技术。主要讨论在鸡尾酒会问题中扮演中心角色的语音分离问题。介绍了以下内容：传统的单通道情况下的技术，如计算听觉场景分析（computational auditory scene analysis, CASA）、非负矩阵分解（non-negative matrix factorization, NMF）以及生成式模型建模；传统的多通道情况下的技术，如波束成形和多通道盲源分离；一些基于深度学习的最新技术，如深度聚类（deep clustering, DPCL）、深度吸引网络（deep attractor network, DANet）以及排列不变性训练（permutation invariant training, PIT）。此外，介绍了在鸡尾酒会环境下针对改善多说话人语音识别和说话人识别精度的相关技术。笔者认为，利用一个更加强大的模型来有效地开发和利用来自麦克风阵列、声学训练集合以及语言本身的知识非常重要。更好的优化策略和技术的提出会逐步解决鸡尾酒会问题。

**关键词：**鸡尾酒会问题；计算听觉场景分析；非负矩阵分解；排列不变性训练；多说话人语音处理