科学家提出一种新的语谱引导滤波语音去混响方法

作者：writer 来源：本站

本文原地址：https://www.iikx.com/news/progress/382.html

本文仅供学习交流之用，版权归原作者所有，请勿用于商业用途！

中国科学院声学研究所噪声与振动重点实验室副研究员郑成诗及其同事提出一种语谱引导滤波语音去混响方法，无需估计晚期混响功率谱和房间声学参数，运算量明显低于传统语音去混响方法，同时能显著提高语音质量。相关研究成果已在线发表于国际学术期刊Applied Acoustics。

已有研究表明，晚期混响声会不同程度地降低语音可懂度及其人机交互语音识别性能，如何有效抑制晚期混响声是近年的研究难点和热点。 单通道晚期混响抑制方法包括谱减法和基于高阶累量的方法等。基于高阶累量的方法，包括Kurtosis最大化和Skewness最大化这两种方法，运算量较大;基于谱减的方法，尽管运算量小、性能稳定，但是需要估计晚期混响功率谱或者房间声学参数。

双边滤波(Bilateral Filtering)和引导滤波(Guided Filtering)是两种比较经典的图像处理方法，都可以用于保持图像边缘，实现图像增强。相比较而言，引导滤波运算量更低、处理结果更为稳定。因此，引导滤波在图像去雾、边缘提取以及细节平滑等方面得到了广泛的应用。
相比于纯净语音语谱图，混响语音语谱图就像是给纯净语音语谱图蒙上了一层雾。

郑成诗在丹麦奥尔堡大学访问期间，曾与奥尔堡大学的教授一起尝试研究图像去雾算法应用于语音去混响，通过分析和实验证明了语谱引导滤波在去混响方面的性能。 语音语谱图有其自身特点，例如语谱在频率轴上的相关性一般要远小于语音在时间轴的相关性。在此基础上，郑成诗及其团队对应用于图像的引导滤波进行了拓展，使其更适用于语谱图增强。

相比传统的方法，该研究提出的语谱引导滤波语音去混响方法在主观语音质量评估(Perceptual Evaluation of Speech Quality,简称PESQ)得分和短时客观可懂度(Short-Time Objective Intelligibility,简称STOI)得分等方面优势显著，而且算法时间复杂度为O(1)。从语谱图看，晚期混响造成的拖尾效应也已得到有效抑制。
该研究得到了国家自然科学基金委的项目支持。(来源：中科院声学研究所) 论文链接

更多 科学进展 请访问 https://www.iikx.com/news/progress/

本文版权归原作者所有，请勿用于商业用途，爱科学iikx.com转发