

基于小波变换与 BP 神经网络的多噪声类型识别与自适应音频降噪方法

王 深

暨南大学 广东广州 511443

摘 要: 本研究针对包含多类噪声, 如背景噪声、突发噪声与带状噪声的复杂音频信号, 提出了一种结合小波变换与 BP 神经网络的噪声识别方法, 并设计了具有自适应能力的多阶段降噪算法。通过连续小波变换与离散小波变换提取时频特征, 利用构建的 BP 神经网络对噪声类型进行多标签识别与量化。在此基础上, 根据不同噪声类型的特征参数, 分别应用谱减法、小波阈值去噪与带阻滤波器等方法实现自适应降噪。实验结果显示, 模型在处理两个音频样本时均能有效识别噪声类型, 降噪后信噪比分别得到提升, 频谱图与波形图对比进一步验证了降噪效果与语音清晰度的改善。本方法适用于复杂噪声环境下的语音增强任务, 具备良好的准确性与鲁棒性。

关键词: 音频去噪; 小波变换; BP 神经网络; 自适应算法; 信噪比提升

引言

在现代音频通信、语音识别与智能终端等应用场景中, 环境噪声对语音信号的干扰已成为影响系统性能的关键因素。尤其在开放或非受控环境下, 音频中往往混杂多种类型的噪声, 如稳定存在的背景噪声、瞬态突发噪声以及集中于某些频段的带状噪声等, 使得传统的单一降噪方法难以同时兼顾降噪效果与语音保真度。因此, 如何准确识别音频中所包含的噪声类型, 并针对性地选择最优的处理策略, 是

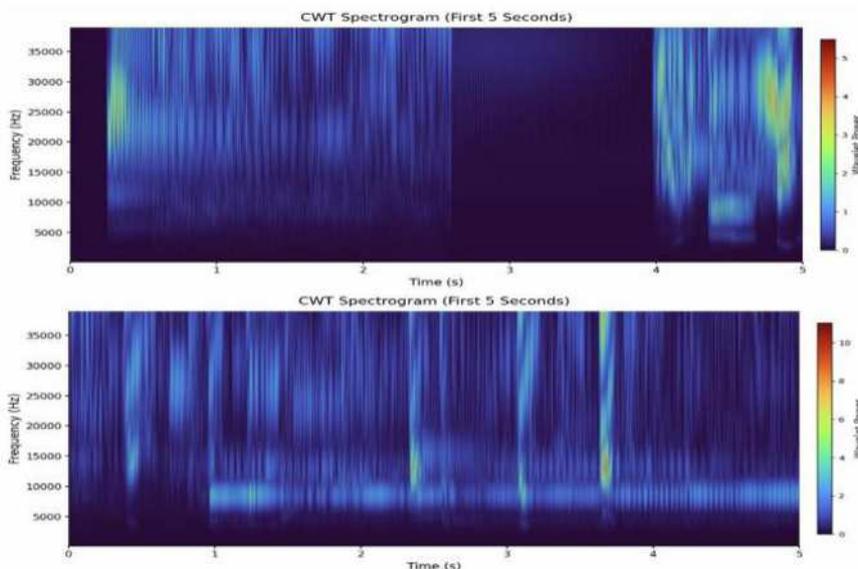
提升音频质量和系统鲁棒性的核心问题。

1 识别与量化噪声

1.1 小波变换处理

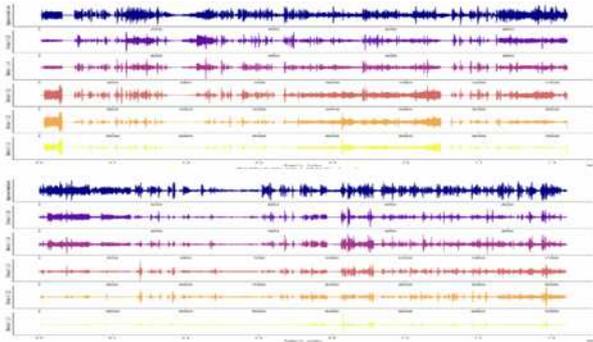
小波变换 (Wavelet Transform) 是一种具有良好时频局部化特性的信号处理方法, 能够在不同尺度上分析信号的高频细节与低频轮廓。

1.1.1. 连续小波变换 (CWT) 理论, 数学定义如下:



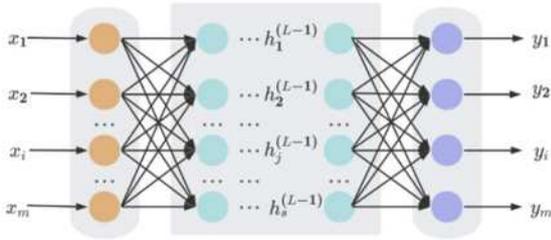
其中: $x(t)$ 是原始信号; $\psi(t)$ 是母小波; a 是尺度 (控制频率); b 是平移 (控制时间位置); $*$ 表示共轭。对本音频进行小波处理得到时频数据。由于数据过大下图仅可视化两端音频的前 5s 的小波变化 CWT 的处理结果。

1.1.2. 离散小波变换 (DWT)



除了连续小波变换用于整体时频结构分析外，我们还对音频信号进行了离散小波变换，以进一步提取多尺度的时间变化特征。DWT 能够将原始信号分解为不同频带的近似分量与细节分量。其中近似分量保留音频中的低频主要信息，反映信号整体轮廓。细节分量对应逐级提升的高频分量，敏感于突发噪声、快速跳变等结构。处理结果如下图所示。

1.2 构建 BP 神经网络结构



为了实现对音频中混合噪声的精准识别与多标签分类，本文采用 BP 神经网络作为噪声识别模型。该网络能够根据小波变换后提取的多维音频特征，实现对多种噪声类型的并发识别与量化输出。

2. 损失函数设置与网络结构设置

层级	节点数	激活函数	功能
输入层	1	—	输入小波变换后的时频特征
隐藏层 1	32	ReLU	提取非线性组合特征
隐藏层 2	16	ReLU	降维提取核心表示
输出层	3	Sigmoid	输出三类噪声的概率

选择损失函数采用逐标签独立的二元交叉熵损失。

1.3 应用结果分析

验证模型在实际音频中的噪声识别能力，我们选取了 part1.wav 与 part2.wav 两段含噪音频 (<https://github.com/6667ws/audio.git>) 分别进行了多标签识别分析，经过神

经网络训练处理后的结果为如下。

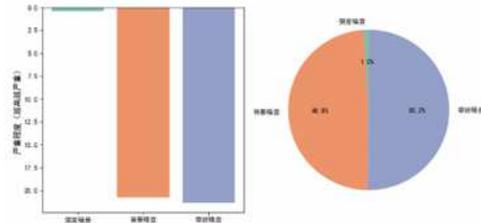


图 27 噪声严重程度图 噪声类型占比图

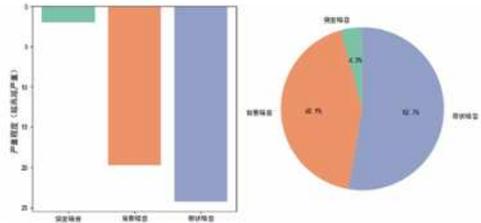


图 28 噪声严重程度图 噪声类型占比图

(1) part1.wav: 音频含三类噪声 (突发、背景、带状)。突发噪声占 4.3%，背景 42.9%，带状 52.7%，后两者为主要干扰。突发噪声短暂出现，背景和带状噪声持续存在。

(2) part2.wav: 同样检测到三类噪声。突发噪声占 1.0%，背景 48.8%，带状 50.2%，主要干扰来自背景和带状噪声，突发噪声影响极小。

2 自适应降噪处理

2.1 降噪算法基础

Step1 背景噪声处理

针对背景噪声，我们提取 ZCR 与谱平坦度 SF 作为量化参数。针对突发噪声，提取 RMS 方差与 RMS 峰值作为量化参数。针对带状噪声，提取频谱质心与频谱宽度值。利用各自特征来进行针对性的自适应处理方案。

Step1: 背景噪声降噪 对于平稳背景噪声，采用谱减法，其核心思想是在频域中从当前帧信号幅度谱中减去估计的背景噪声谱幅度：

$$|\hat{S}(f)| = \max(|X(f)| - \alpha \cdot |\hat{N}(f)|, \beta \cdot |\hat{N}(f)|)$$

其中 $|X(f)|$ 为当前帧频谱， $\hat{N}(f)$ 为噪声声谱， α 为过减因子， β 为下限残留控制因子。参数自适应设置， α 根据零交叉率 ZCR 动态调整。

$$\alpha = \alpha_0 + k_1 \cdot (1 - \text{ZCR}) \quad \beta = \beta_0 - k_2 \cdot \text{SF}$$

Step2 突发噪声处理

针对瞬态突发噪声，采用小波阈值去噪方法

1. 对高频系数应用阈值处理

$$d_j^{new} = \begin{cases} \text{sign}(d_j) \cdot (|d_j| - \lambda), & \text{if } |d_j| > \lambda \\ 0, & \text{otherwise} \end{cases}$$

2. 对处理后的小波系数重构音频信号。

$$\lambda = k_2 \cdot \sigma_{RMS} + k_3 \cdot \left(\frac{\text{Peak}}{\text{RMS}} - 1 \right)$$

其中 σ_{RMS} 表示能量波动程度, $\frac{\text{Peak}}{\text{RMS}}$ 反映突发强度。

Step3 带状噪音处理

对于集中出现在特定频段的带状噪声, 二阶带阻滤波器抑制其中心频率成分, 滤波器传递函数如下:

$$H(z) = \frac{z^2 - 2\cos(2\pi f_0/F_s)z + 1}{z^2 - 2r\cos(2\pi f_0/F_s)z + r^2}$$

2.2 处理结果分析

1. 信噪比结果分析

采用 SNR 作为主要性能指标, 公式如下:

$$\text{SNR} = 10 \cdot \log_{10} \left(\frac{\|s(t)\|^2}{\|s(t) - \hat{s}(t)\|^2} \right)$$

处理后结果如下:

part1.wav 降噪后信噪比: 16.578 dB

part2.wav 降噪后信噪比: 14.578 dB

结果显示, 本系统在复杂噪声环境中通过“识别 + 自适应降噪”显著提升信噪比。尤其 part1.wav 在多种高强度噪声下仍达 16.578 dB, 接近主观清晰度边界, 验证了模型的多噪声处理能力。

2. 频谱图结果分析

为评估降噪对语音频域特性的影响, 对 part1.wav 和 part2.wav 进行了降噪前后频谱图对比。频谱图展示时间-频率能量分布, 可直观反映背景、突发及带状噪声的抑制效果。每张图自上而下分别为原始音频和降噪后音频。

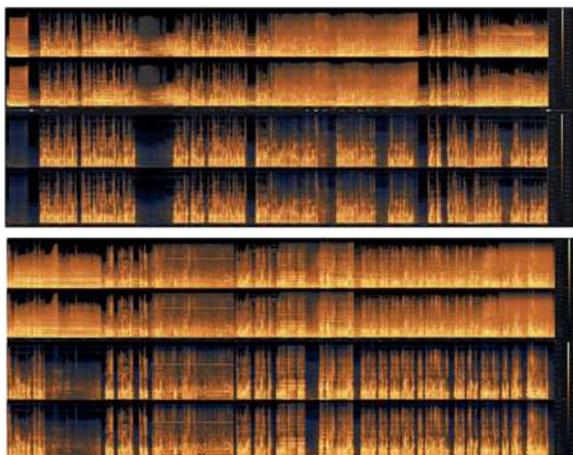


图 3

part1.wav: 降噪后频谱整体亮度降低, 低频背景能量显著减弱, 语音主频保留良好; 高频垂直条带明显变淡, 带状噪声得到有效抑制。

part2.wav: 降噪后低频噪声明显清除, 语音主频更清晰连续, 频谱由“整体泛亮”转为“结构清晰、语音主导”, 体现方法对复合噪声的良好适应性。

3. 音振幅对比

为验证降噪在时域的效果, 选取 part1.wav 前 10 秒绘制降噪前后波形图。原始波形图中噪声区域(人耳明显感知)已用圆圈标出, 处理后对应噪声显著消除, 语音信号更清晰。

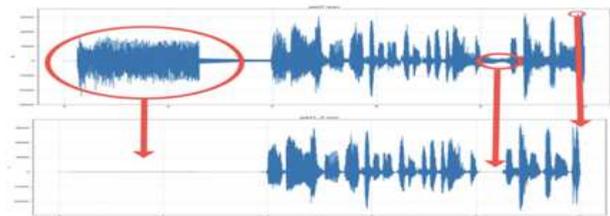


图 30 part1.wav 降噪前后的音频波形图

图 4

从对比图可见, 降噪后波形更平稳, 语音结构完整保留, 背景干扰显著减弱。系统基于三类噪声识别的自适应算法, 有效削弱甚至消除了红圈标注的高振幅噪声区域, 体现出良好的时域一致性和局部针对性。

3 模型总结

本文针对复杂噪声环境下的音频处理, 提出融合噪声识别与自适应降噪的方案, 重点应对背景噪声、突发噪声和带状噪声。实证结果表明, 该方法可有效识别并量化多类噪声, 显著提升信噪比, 并在时域与频域均表现出优良的去噪效果。尤其在多噪声共存场景下, 算法可根据噪声特征动态调整参数, 兼顾环境适应性与语音结构保真度。实验显示, 降噪后信噪比分别达 16.578 dB 和 14.578 dB, 主观听感更清晰自然, 验证了方法在多源噪声处理中的有效性与稳定性。方案结构灵活、可扩展, 具备工程应用潜力, 可用于语音通信、识别及智能终端的预处理环节。未来可探索更复杂的网络结构, 并通过大规模真实音频训练提高模型泛化与实用性。

参考文献:

- [1] 胡赵宇, 李喆, 蒙国勇等. 基于 Conv-TasNet 的变压器音频降噪识别网络 [J]. 电气自动化, 2024,46(6):82-85.
- [2] 李志远. 基于深度学习的广播电视音频降噪技术 [J]. 电声技术, 2024,48(12):122-125.

- [3] 刘豪. 音频降噪回声消除技术研究和应用 [J]. 通信电源技术, 2021, 38(5): 195-197.
- [4] 赵琳. 基于卷积神经网络的数字音频降噪技术研究 [J]. 电声技术, 2025, 49(5): 82-84.
- [5] 赵丹, 李蕊. 噪声环境中多轨道数字音频信号降噪方法 [J]. 现代电子技术, 2024, 47(13): 19-22.
- [6] 刘卫强, 林品荣, 陈儒军等. 基于深度学习与联合降噪的可控源音频大地电磁抗干扰数据处理方法研究 [J]. 地球物理学进展, 2024, 39(4): 1457-1473.
- [7] 樊一帆, 张丽丹. 强噪环境基于谱减法的录音数字音频信号降噪 [J]. 计算机仿真, 2023, 40(11): 433-436, 474.
- [8] 魏远征. 融合人工智能的音频信号降噪技术研究 [J]. 电声技术, 2024, 48(3): 144-146.
- [9] 陈伟. 音频信号处理中的自适应降噪算法 [J]. 电声技术, 2024, 48(3): 51-53.
- [10] 牛怡琴. 数字音频信号处理中的主动降噪方法 [J]. 卫星电视与宽带多媒体, 2024(11): 126-128.
- [11] 王政一, 侯婕. 基于人工智能技术的电视广播音频自动化降噪方法研究 [J]. 电声技术, 2025, 49(1): 119-121.