

文献引用格式: 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.

WEI Y Z. Research on audio signal denoising technology integrating artificial intelligence [J]. Audio Engineering, 2024, 48(3): 144 – 146.

中图分类号: TN912.3

文献标识码: A

DOI: 10.16311/j.audiee.2024.03.043

## 融合人工智能的音频信号降噪技术研究

魏远征

(河南医药健康技师学院, 河南 开封 475000)

**摘要:** 文章研究了一种融合人工智能的音频信号降噪技术。首先, 构建一个包含滤波方法、人工智能等关键组成部分的音频降噪框架。其次, 针对音频降噪框架中自适应滤波与长短期记忆 (Long Short Term Memory, LSTM) 网络相结合的声音降噪方法, 详细介绍其数学原理和实现过程。最后, 采用 Noisex-92 数据集对所提方法进行全面的测试与评估。结果表明, 文章提出的方法在信噪比 (Signal to Noise Ratio, SNR) 和信号失真比 (Signal Distortion Ratio, SDR) 上均取得了显著提升。

**关键词:** 人工智能; 长短期记忆 (LSTM); 自适应滤波; 音频降噪

### Research on Audio Signal Denoising Technology Integrating Artificial Intelligence

WEI Yuanzheng

(Henan Technician College of Medicine and Health, Kaifeng 475000, China)

**Abstract:** In this paper, an audio signal denoising technology with artificial intelligence is studied. Firstly, an audio denoising framework is constructed, which includes key components such as filtering method and artificial intelligence. Secondly, the mathematical principle and implementation process of the method of noise reduction combining adaptive filtering with Long Short Term Memory (LSTM) network in the audio noise reduction framework are introduced in detail. Finally, Noisex-92 data set is used to test and evaluate the proposed method comprehensively. The results show that the method proposed in this paper has significantly improved the Signal to Noise Ratio (SNR) and Signal Distortion Ratio (SDR).

**Keywords:** artificial intelligence; Long Short Term Memory (LSTM); adaptive filtering; audio denoising

## 0 引言

随着人工智能技术的不断发展和普及, 其在各领域的应用展现出巨大的潜力<sup>[1-2]</sup>。音频信号降噪技术旨在提高音频质量、降低噪声干扰, 从而满足人们日益增长的音频处理需求, 如语音识别、音乐重建等。在此背景下, 融合人工智能与传统滤波方法的音频信号降噪技术成为当前研究的热点之一<sup>[3-4]</sup>。目前, 传统的滤波方法, 如均值滤波、中值滤波等, 被广泛应用于去噪和信号增强任务中。然而, 这些方法往往难以处理复杂的噪声环境和动态信号变化, 效果有限。为此, 越来越多的研究开

始将人工智能技术引入音频信号增强领域, 如深度学习算法、循环神经网络等, 以提高音频增强的效果和性能。文章探讨了一种融合自适应滤波<sup>[5-6]</sup>与长短期记忆 (Long Short Term Memory, LSTM) 网络<sup>[7-8]</sup>的声音降噪方法, 并将其应用于音频信号降噪任务, 采用 Noisex-92 数据集对该方法进行测试与评估<sup>[9-10]</sup>。通过本研究, 旨在为音频信号降噪技术的进一步发展和应用提供新的思路和方法。

## 1 融合人工智能的音频降噪框架

文章将传统的滤波方法与人工智能技术相结合, 构建音频降噪总体框架, 如图 1 所示。

作者简介: 魏远征 (1992—), 男, 本科, 助理讲师, 研究方向为计算机应用。

信号预处理模块用于对输入音频信号进行预处理,包括降噪、降采样等操作,以提高后续处理的效果。在特征提取模块中,利用特征提取方法对音频信号进行特征提取,以捕获信号的关键信息。常用的特征提取方法包括时频分析、小波变换等。引入传统的滤波方法,如卷积、均值滤波等,用于对音频信号进行初步处理和去除噪声。人工智能处理模块采用人工智能技术,如长短期记忆网络、卷积神经网络等,对音频信号进行更高级的处理和特征学习,以进一步提升音频降噪的效果。对于经过人工智能处理后的音频信号,需要进行一些后处理操作,如去混响、音量平衡等,以使其更符合实际应用需求。

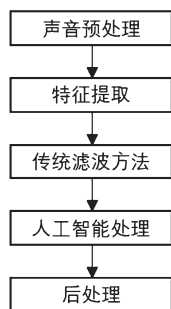


图1 融合人工智能的音频降噪框架

## 2 自适应滤波与 LSTM 相结合的音频降噪

自适应滤波是一种基于信号处理和统计理论的方法,主要根据输入信号的特性动态调整滤波器的参数,以适应信号环境的变化。自适应滤波的核心思想是通过最小化预测误差来优化滤波器的参数,从而实现对输入信号的准确预测和适应性调节。在自适应滤波中,预测误差是指实际输出信号与期望输出信号之间的差值,其大小反映了滤波器对输入信号的处理效果。自适应滤波的应用领域广泛,包括语音处理、图像处理、通信系统等。在语音处理中,自适应滤波常用于去除噪声、降低回声等,有利于提高语音质量和清晰度。在自适应滤波中,设一个长度为  $N$  的输入信号序列为  $x(n)$ ,其经过一个自适应滤波器  $h(n)$  处理后得到输出信号  $y(n)$  为

$$y(n) = \sum_{k=0}^M h(k)x(n-k) \quad (1)$$

式中:  $M$  为自适应滤波器的长度,  $h(k)$  为滤波器的

系数。

通常情况下,自适应滤波器的系数  $h(k)$  通过最小均方 (Least Mean Square, LMS) 算法进行更新,其更新规则为

$$h(k+1) = h(k) + \mu e(n)x(n-k) \quad (2)$$

式中:  $e(n)$  为预测误差,  $\mu$  为学习率参数。

$e(n)$  为实际输出信号与期望输出信号之间的差值,即

$$e(n) = d(n) - y(n) \quad (3)$$

设  $y(n)$  为经过自适应滤波处理后得到的输出信号序列,将该序列作为 LSTM 的输入。LSTM 模型包括输入门、遗忘门以及输出门等。假设  $y(n)$  为 LSTM 的输入序列,  $h'(n)$  为隐藏层状态序列,  $c(n)$  为细胞状态序列,则输入门  $i(n)$  为

$$i(n) = \sigma[W_{xi}y(n) + W_{hi}h'(n-1) + W_{ci}c(n-1) + b_i] \quad (4)$$

式中:  $\sigma$  为 sigmoid 函数,  $W_{xi}$ 、 $W_{hi}$ 、 $W_{ci}$  为输入门的权重矩阵,  $b_i$  为偏置向量。

计算遗忘门  $f(n)$  的值,公式为

$$f(n) = \sigma[W_{xf}y(n) + W_{hf}h'(n-1) + W_{cf}c(n-1) + b_f] \quad (5)$$

接着更新细胞状态  $c(n)$  的值,公式为

$$c(n) = f(n) \odot c(n-1) + i(n) \odot \tanh[W_{xc}y(n) + W_{hc}h'(n-1) + b_c] \quad (6)$$

式中:  $\odot$  表示逐元素相乘,  $\tanh(\cdot)$  为双曲正切函数,  $W_{xc}$ 、 $W_{hc}$  为细胞状态的权重矩阵,  $b_c$  为偏置向量。

输出门  $o(n)$  的值为

$$o(n) = \sigma[W_{xo}y(n) + W_{ho}h'(n-1) + W_{co}c(n-1) + b_o] \quad (7)$$

更新隐藏状态  $h(n)$  的值,计算公式为

$$h'(n) = o(n) \odot \tanh[c(n)] \quad (8)$$

通过以上方法,可以得到经过 LSTM 处理后的声音降噪结果序列  $h'(n)$ 。

## 3 实验与分析

为验证所提方法的正确性,文章采用 Noisex-92 数据集对方法进行测试,具体步骤如下。第一,数据准备。从 Noisex-92 数据集中获取干净音频信号和各类噪声样本,并划分数据集为训练集和测试集。第二,预处理。对训练集中的干净音频信号和噪声样本进行预处理,包括采样率调整、标准化、去除静音段等操作。第三,模型选择与训练。基于 MATLAB 构建结合自适应滤波与 LSTM 的声

表 1 实验结果

实验序号	传统自适应滤波方法		文章所提方法	
	SNR/dB	SDR/dB	SNR/dB	SDR/dB
1	15.2	13.5	18.7	16.3
2	13.8	12.2	17.5	15.1
3	16.5	14.8	19.2	17.0
4	14.3	12.7	18.0	15.6
5	17.2	15.5	20.1	18.0

音降噪模型,并在训练集上对所选模型进行训练。第四,测试模型。在测试集上对训练好的模型进行验证,评估模型在不同噪声环境下的性能表现,如信噪比(Signal to Noise Ratio, SNR)、信号失真比(Signal Distortion Ratio, SDR)等,结果如表 1 所示。

从表 1 可以看出,相较于传统自适应滤波方法,文章所提方法在 SNR 和 SDR 指标上均取得了显著的提升。在 SNR 指标上,文章所提方法比传统自适应滤波方法平均提升了 3.1 dB,在 SDR 指标上平均提升了 2.7 dB。这表明文章所采用的方法能够有效提高音频信号的质量和清晰度,显著降低了信号中的噪声水平。此外,通过对比多个样本实验,可以发现文章所提方法在不同样本上的效果一致且稳定,表明该方法具有较好的泛化能力,适用于不同的音频信号环境,并且能够在各种情况下取得较好的降噪效果。文章所提方法在音频信号降噪任务上具有较高的正确性和可行性,能够有效提高音频信号的质量和清晰度,为相关领域的研究和应用提供了有力支持。

#### 4 结语

文章主要研究了融合人工智能的音频信号降噪技术,重点分析自适应滤波与 LSTM 相结合的音频降噪方法。实验结果表明,所提方法在 SNR 和 SDR 指标上均取得了显著提升,证明了其在音频信号降噪任务中的有效性和可靠性。该研究成果为音频处理领域的发展和应用提供了新的思路和方法,具有重要的理论和实践意义。未来需要进一步

优化算法性能,探索更多的音频降噪手段,以满足人们日益增长的音频处理需求。

#### 参考文献:

- [1] 张妮,徐文尚,王文文.人工智能技术发展及应用研究综述[J].煤矿机械,2009,30(2):4-7.
- [2] 邹蕾,张先锋.人工智能及其发展应用[J].信息安全,2012(2):11-13.
- [3] 史秋莹,郑铁然.基于深度学习的环境声音识别[J].智能计算机与应用,2018,8(5):34-37.
- [4] 杨帆,李祎男,乔涵,等.基于深度卷积神经网络的语音信号去噪关键技术研究[J].计算机与数字工程,2022,50(2):344-349.
- [5] 曹晏飞,滕光辉,余礼根,等.含风机噪声的蛋鸡声音信号去噪方法比较[J].农业工程学报,2014,30(2):212-218.
- [6] 米建伟,方晓莉,仇原鹰.非平稳背景噪声下声音信号增强技术[J].仪器仪表学报,2017,38(1):17-22.
- [7] 朱乔木,李弘毅,王子琪,等.基于长短期记忆网络的风电场发电功率超短期预测[J].电网技术,2017,41(12):3797-3802.
- [8] 杨丽,吴雨茜,王俊丽,等.循环神经网络研究综述[J].计算机应用,2018,38(增刊2):1-6.
- [9] 褚伟.基于条件深度卷积生成对抗网络的语音增强研究[J].智能计算机与应用,2019,9(4):82-86.
- [10] 李银国,欧阳希子,郑方.语音识别中听觉特征的噪声鲁棒性分析[J].清华大学学报(自然科学版),2013,53(8):1082-1086.

编辑:郭芳园