



双解码卷积循环网络风噪声有源控制

摘要

本文提出一种利用双解码卷积循环网络 (Dual-decoder Convolutional Recurrent Network, DCRN) 代替 FxLMS (Filtered-x Least Mean Square) 算法的有源噪声控制方法, 考虑到相位信息在有源噪声控制 (Active Noise Control, ANC) 中的重要性, DCRN 网络的输入特征为噪声信号的复数频谱 (包括实部谱和虚部谱). 网络结构中, 采用编码模块从噪声复数频谱中提取特征, 利用双解码模块分别估计网络输出的实部谱和虚部谱, 采用参数共享机制和组策略以降低训练参数的数量并提高网络的学习能力和泛化能力. 特别是针对风噪声, 选用新的损失函数以及对训练数据进行正则化处理以提升 DCRN 的性能. 实验结果表明, DCRN 方法在仿真环境与有源降噪耳机环境下对一般噪声和风噪声都表现出良好的降噪性能和鲁棒性.

关键词

双解码卷积循环网络; 有源噪声控制; FxLMS 算法; 复数频谱

中图分类号 TN912.3

文献标志码 A

收稿日期 2023-08-18

资助项目 国家自然科学基金 (12074192)

作者简介

吴礼福, 男, 博士, 教授, 研究方向为音频信号处理. wulifu@nuist.edu.cn

1 南京信息工程大学 电子与信息工程学院, 南京, 210044

2 南京信息工程大学 大气环境与装备技术协同创新中心, 南京, 210044

0 引言

嘈杂环境对人们心理健康和工作造成的负面影响使得噪声控制成为日益关注的焦点. 无源噪声控制利用物理材料与噪声声波相互作用消耗声能来衰减大部分中高频噪声^[1], 但其处理低频噪声的能力有限. 有源噪声控制 (Active Noise Control, ANC) 是一种基于声波干涉原理的噪声控制技术^[2], 它通过生成与初级噪声幅度相同、相位相反的抵消信号来衰减初级噪声, 已有了成功的应用, 如有源降噪耳机^[3-4].

风噪声作为低频非平稳噪声的典型代表, 在实际环境中广泛存在, 它是由空气湍流与障碍物 (如建筑物、人体、传声器腔体等) 之间相互作用产生的^[5-7], 其能量主要集中在 1 000 Hz 以下的低频, 并且随着频率的增加而滚降^[8-9]. 风噪声的非平稳特性使得传统的噪声抑制如基于维纳滤波解的降噪方法表现较差^[10-11]. 虽然有源噪声控制领域已经提出了一些针对风噪声的方法, 如 Chu 等^[12] 提出基于总最小二乘对滤波 x 最小均方算法进行修正, 提高风噪环境下噪声控制性能, 实验结果表明, 该算法在有风条件下可以提高约 3 dB 降噪量. 目前大多数针对风噪声的处理方法仍主要基于物理结构来削弱风噪影响, 例如: 吴睿等^[13] 基于风噪原理提出一种抗风噪结构抑制风噪对传声器采集信号的影响, 该设计由渐变管道和填充泡棉的腔体两部分组成, 风噪实验证明了该设计能有效地抑制风噪; 潘作峰等^[14] 提出一种密封状态下的车内风噪性能区间不确定性的分析和优化方法, 该方法可在保证相关零部件质量基本不变的前提下, 降低车内噪声水平及其波动幅度, 明显提升系统的稳健性.

近年来, 神经网络在有源噪声控制中的运用引起了关注. Jiang 等^[15] 首先利用经典 ANC 算法处理噪声, 然后将残余噪声和估计噪声送入神经网络以消除非线性失真. Mostafavi 等^[16] 提出一种基于深度学习的前馈 ANC 控制器, 通过考虑声学设备的延迟和非线性以衰减与施工相关的噪声. Im 等^[17] 提出一种深度学习辅助的次级路径更新技术, 训练深度神经网络根据不断变化的边界条件实时估计次级路径以应对时变环境. Cha 等^[18] 提出一种基于 DNoiseNet 的反馈有源降噪方法, 通过在各种环境中学习不同噪声的多级时间特征, 克服了传统有源降噪的局限性, 在此基础上, 还提出了基于多层感知机 (Multi-layer Perception, MLP) 神经网络的次级路径估计器以提高 DNoiseNet

的性能.Zhang 等^[19]提出一种 Deep-ANC 方法以解决非线性 ANC 问题,该方法训练网络以编码不同噪声环境下的最优控制参数,还引入延迟补偿策略解决 ANC 系统潜在的延迟问题.Kwon 等^[20]采用长短期记忆(Long Short-Term Memory, LSTM)层作为 ANC 控制器,生成控制信号以最小化残留噪声.

鉴于风噪声的非平稳、非线性特性和神经网络强大的非线性映射能力,本文研究了双解码卷积循环网络(Dual-decoder Convolutional Recurrent Network, DCRN)^[21]代替传统 FxLMS(Filtered-x Least Mean Square)算法的 ANC 方法.考虑到相位信息在 ANC 中的重要性,DCRN 网络的输入特征为噪声信号的复数频谱(包括实部谱和虚部谱).为了降低训练参数的数量并提高网络的学习能力和泛化能力,网络结构中采用了参数共享机制和组策略.

1 算法描述

1.1 FxLMS 算法

因其简单、稳定和相对低的计算复杂度, FxLMS 及其扩展算法在 ANC 中应用广泛. FxLMS 算法是以 LMS 算法为基础,在参考信号与滤波器之间放置估计的次级路径以补偿次级路径对有源噪声控制系统的影响.单通道前馈 FxLMS 算法框图如图 1 所示.

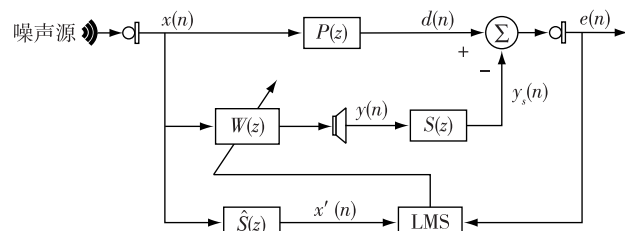


图 1 单通道前馈 FxLMS 算法框图

Fig. 1 Block diagram of single-channel feed-forward FxLMS

参考信号 $x(n)$ 由参考麦克风拾取并通过自适应滤波器 $W(z)$ 生成次级信号 $y(n)$, 如式(1). 参考信号和次级信号分别经过初级路径 $P(z)$ 和次级路径 $S(z)$ 得到期望信号 $d(n)$ 和抵消信号 $y_s(n)$, 噪声残留由误差麦克风拾取得到. FxLMS 算法对滤波器 $W(z)$ 进行自适应迭代, 生成抵消信号 $y_s(n)$ 以最小化噪声残留信号 $e(n)$, 如式(2). $W(z)$ 迭代更新方程如式(3).

$$y(n) = \mathbf{w}^T(n)x(n), \quad (1)$$

$$e(n) = d(n) - y_s(n), \quad (2)$$

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu e(n)[x(n) * \hat{s}(n)]. \quad (3)$$

式(3)中: $\mathbf{w}(n)$ 为自适应滤波器权重; μ 是自适应步

长参数; $\hat{s}(n)$ 为次级路径估计 $\hat{s}(z)$ 的脉冲响应; $*$ 表示线性卷积运算.

1.2 DCRN 方法

1.2.1 算法原理

图 2 所示为单通道前馈 DCRN 方法的系统框图. 相较于图 1, DCRN 方法采用 DCRN 代替 $W(z)$. 参考信号 $x(n)$ 经过短时傅里叶变换(Short-time Fourier Transform, STFT)后得到复频谱, 其实部和虚部频谱分别为 $X_r(t, f)$ 和 $X_i(t, f)$, 其中, t 和 f 表示为时频(T-F)单元内的时间和频率, 下标 r 和 i 分别表示信号频谱的实部和虚部. 将 $X_r(t, f)$ 和 $X_i(t, f)$ 作为输入送至 DCRN 并计算出次级信号的实部与虚部频谱即 $Y_r(t, f)$ 和 $Y_i(t, f)$. 次级信号的复频谱经逆短时傅里叶变换得到时域信号 $y(n)$, 与次级路径 $S(z)$ 卷积得到抵消信号 $y_s(n)$. $d(n)$ 和 $y_s(n)$ 用于计算损失函数, 训练 DCRN 参数以最小化误差信号 $e(n)$.

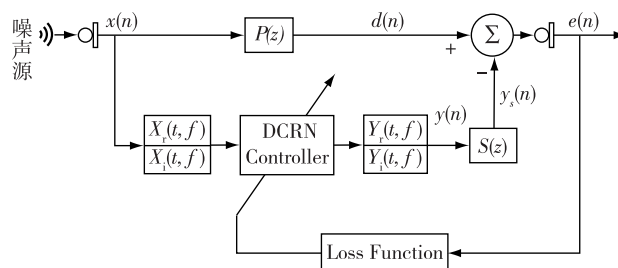


图 2 单通道前馈 DCRN 方法系统框图

Fig. 2 Block diagram of single-channel feed-forward DCRN

1.2.2 网络参数

图 2 中 DCRN 具体结构如图 3 所示. DCRN 结构包括 3 个部分: 编码模块、时序建模模块和解码模块. 编码模块和解码模块分别包含 5 层卷积层和 5 层反卷积层, 位于编码模块和解码模块之间的时序建模模块由采用组策略的 2 个长短期记忆层(LSTM)组成^[22-23], 其中组数设置为 2. 采用组策略的 LSTM 层有助于提高循环层的效率和降低模型的复杂度. 编码模块从输入信号复频谱中提取出高维特征, 而解码模块以逆向操作将编码模块提取的高维特征逐层恢复至目标信号.

图 3 中 $X_r(t, f)$ 和 $X_i(t, f)$ 被视为 2 个不同的输入通道, 编码模块和时序建模模块共享 $X_r(t, f)$ 和 $X_i(t, f)$, 而 2 个解码模块分别用于估计 DCRN 网络输出信号的实部和虚部频谱. 这样的设置使得 DCRN 可以同时处理实部和虚部的复频谱映射, 将估计实部和虚部视为 2 个相关子任务, 从而通过参数共享

鼓励学习和提高泛化能力.在模型训练过程中,采用了跳跃连接方法(skip connection)^[24],该方法将编码模块每个卷积层输出与双解码模块中对应的反卷积层的输入相连接,可以逐层传递编码模块中提取的特征到解码模块,避免输入信息在传递过程中丢失过多,从而改善通过DCRN的信息和梯度.DCRN网络在除输出层外的所有卷积层和反卷积层中均使用指数线性单元(Exponential Linear Units, ELU)作为

激活函数,并在每次卷积(或反卷积)之后采用批量归一化操作(Batch Normalization, BN)来加快训练速度和提高模型稳定性.此外,在2个解码模块的输出层之后使用线性层来预测输出信号的实部和虚部频谱.

表1给出了DCRN模型每层网络的具体参数设置,其中, T 表示STFT幅度谱中时间帧的数量,双解码模块的反卷积层参数相同.表1中:卷积层和反卷

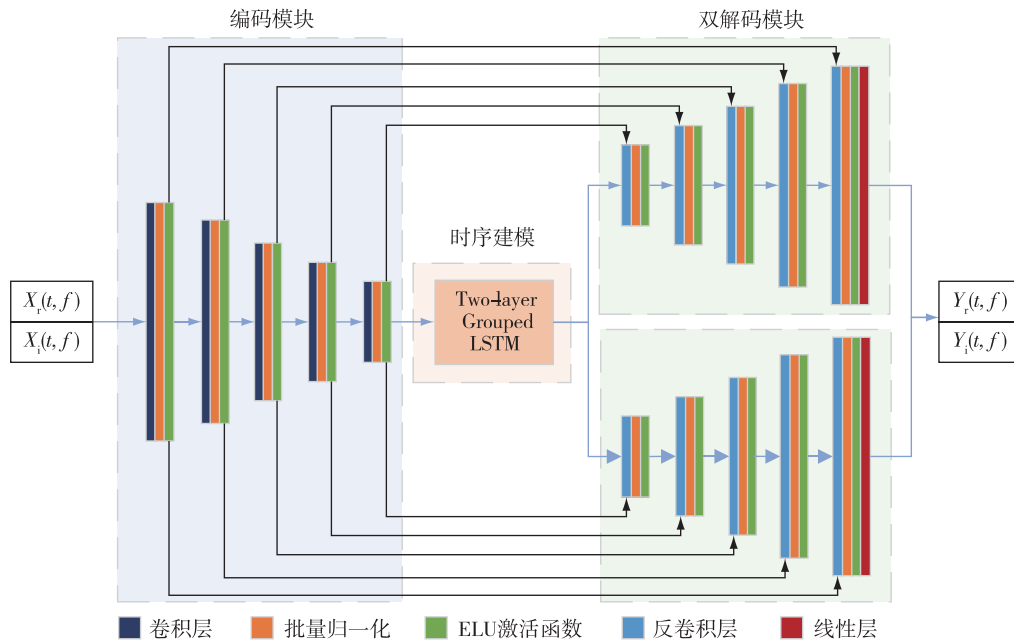


图3 DCRN 的具体结构

Fig. 3 Specific structure of DCRN

表1 DCRN 参数

Table 1 DCRN parameters

DCRN	层名称	输入维度	超参数	输出维度
编码模块	Conv2d_1	$(2, T, 161)$	$(1, 3), (1, 2), 8$	$(8, T, 80)$
	Conv2d_2	$(8, T, 80)$	$(1, 3), (1, 2), 16$	$(16, T, 39)$
	Conv2d_3	$(16, T, 39)$	$(1, 3), (1, 2), 32$	$(32, T, 19)$
	Conv2d_4	$(32, T, 19)$	$(1, 3), (1, 2), 64$	$(64, T, 9)$
	Conv2d_5	$(64, T, 9)$	$(1, 3), (1, 2), 128$	$(128, T, 4)$
时序建模模块	Reshape_1	$(128, T, 4)$		$(T, 512)$
	Grouped_LSTM_1	$(T, 512)$	512	$(T, 512)$
	Grouped_LSTM_2	$(T, 512)$	512	$(T, 512)$
	Reshape_2	$(T, 512)$		$(128, T, 4)$
解码模块	Deconv2d_5(x2)	$(256, T, 4)$	$(1, 3), (1, 2), 64$	$(64, T, 9)$
	Deconv2d_4(x2)	$(128, T, 9)$	$(1, 3), (1, 2), 32$	$(32, T, 19)$
	Deconv2d_3(x2)	$(64, T, 19)$	$(1, 3), (1, 2), 16$	$(16, T, 39)$
	Deconv2d_2(x2)	$(32, T, 39)$	$(1, 3), (1, 2), 8$	$(8, T, 80)$
	Deconv2d_1(x2)	$(16, T, 80)$	$(1, 3), (1, 2), 1$	$(1, T, 161)$
线性层	Linear(x2)	$(1, T, 161)$		$(1, T, 161)$
	concat	$(1, T, 161)(\times 2)$	161	$(2, T, 161)$

积层的输入维度和输出维度括号内 3 个参数分别表示特征通道数、时间总帧数、特征维度,卷积层和反卷积层的层间超参数分别对应卷积核、卷积步幅、输出通道数;5 层卷积层的输出通道数分别为 8、16、32、64、128,5 层反卷积层的输出通道数分别为 64、32、16、8、1;卷积层和反卷积层的时间维度和特征维度的卷积步幅设置为(1,2),卷积核均设置为(1,3).此外,为了聚合频谱的上下文,所有卷积层和反卷积层中沿频率维度采用的步长为 2.值得注意的是,由于采用了跳跃连接方法,解码模块中每层反卷积层的输入特征通道数是对应的卷积层输入特征通道数 2 倍.这样的设计使得编码模块提取的特征能够更加有效地传递到解码模块,从而有助于保持和恢复输入信号的重要特征.

1.2.3 特征提取和损失函数

网络输入参考信号 $x(n)$ 均以 16 kHz 采样,使用 20 ms 的汉明窗将参考信号 $x(n)$ 分割成一系列时间帧,相邻两帧之间的重叠为 10 ms.对每帧应用 320 点的短时傅里叶变换(STFT)生成 161 维的实部和虚部频谱.

本文选择复频谱映射是由于相位在有源噪声控制中的重要性.ANC 的目标是产生抵消信号以衰减误差麦克风处的初级噪声,均方误差(Mean Squared Error, MSE)损失函数是最广泛使用的处理回归问题的损失函数之一.式(4)为 MSE 损失函数的计算公式,式中, N 为样本数量即每次迭代数据中批处理量大小, $y_{k, \text{true}}$ 和 $y_{k, \text{predicted}}$ 分别为第 k 个样本的真实值和对应的预测值,在本文中分别是 $d(n)$ 与 $y_s(n)$ 经 STFT 变换后的复频谱.

$$e_{\text{MSE}} = \frac{1}{N} \sum_{k=0}^N (y_{k, \text{predicted}} - y_{k, \text{true}})^2. \quad (4)$$

2 实验结果与分析

2.1 实验配置

实验分为两部分,分别验证 DCRN 方法在一般噪声和风噪声条件下的降噪性能.第一部分实验使用 Sound-Effect 音效库 General Series 6000 系列(https://soundeffects.fandom.com/wiki/The_General_Series_6000_Sound_Effects_Library)中多种噪声信号创建训练集.验证集采用 NOISEX-92 数据集^[25]中的 Leopard 噪声,测试集使用 NOISEX-92 中的 babble 噪声、engine 噪声和 factory 噪声.第二部分实验针对风噪声进行训练,由于目前缺乏公开的风噪声数据集,

本文采用 Nelke 等^[26]提出的合成风噪声信号模型(后文称“Nelke 模型”)生成风噪声信号,包括阵风 and 恒定风两种类型.风噪声的数据集设置为: Nelke 模型生成约 10 h 的恒定风噪声作为训练集,约 1 h 的阵风风噪声作为验证集,在户外场景采集的正常风速和强风速两种条件下真实风噪声(<https://www.iks.rwth-aachen.de/en/research/tools-downloads/databases/wind-noise-database/>)作为测试集.

在每一部分实验中,为了验证网络的泛化性能,采用两类初级路径和次级路径进行实验.第一类采用镜像源法^[27]生成房间冲击响应(Room Impulse Responses, RIRs).本文模拟了一个 3 m×4 m×2 m 的房间,其中,参考麦克风位于(1.5, 1, 1) m 处,次级源(抵消扬声器)位于(1.5, 2.5, 1) m 处,误差麦克风位于(1.5, 3, 1) m 处.模拟生成的 RIRs 混响时间 $T_{60} \in \{0.15, 0.175, 0.2, 0.225, 0.25\}$ s^[28],阶数为 512.每种混响时间生成两个 RIRs,分别用作初级路径和次级路径.在训练阶段随机选择混响时间 $T_{60} \in \{0.15, 0.175, 0.225, 0.25\}$ s 中的一组 RIRs,测试阶段采用混响时间为 0.2 s 的 RIRs 作为默认测试 RIRs.第二类旨在验证算法在有源降噪耳机中的实际效果,其初级路径与次级路径来自于吴礼福等^[2]实测的有源降噪耳机场景,阶数分别为 512 和 256.

对于两部分实验,分别创建了 20 000 条训练数据、100 条验证数据和 100 条测试数据.每条训练数据、验证数据以及测试数据均是从对应数据集中随机裁剪的 6 s 信号创建.在训练模型时使用 AMSGard 优化器^[29]更新网络权重,初始学习率设置为 0.001,采用 StepLR 机制在训练过程中动态调整学习率,批大小设为 16,训练总轮数设为 30.

2.2 实验结果分析

本文以降噪量评价算法的降噪性能,其中,降噪量定义为算法作用前后残留噪声的功率谱密度差异.对比算法选择 FxLMS 算法与文献[19]中提出的 Deep-ANC 方法.FxLMS 算法滤波器长度设为 512; Deep-ANC 方法使用与 DCRN 方法相同的超参数和数据集.

式(4)中的平方运算会放大较大误差的影响(对异常值比较敏感),导致模型对异常值过度拟合.平均绝对值误差(Mean Absolute Error, MAE)损失函数由于使用绝对值计算使其对异常值的鲁棒性更好.由于风噪声信号呈现能量集中在低频部分并且随着频率升高能量滚降,而 MAE 对所有误差赋予相

同权重,可以更平衡地考虑低频与高频之间的差异.因此,在风噪声中优先采用 MAE 损失函数,同时将 MAE 损失函数用于一般噪声中进行对比.式(5)为 MAE 损失函数计算公式,式中各参数同式(4).后文中对使用 MSE 和 MAE 损失函数的 DCRN 方法分别命名为 DCRN-MSE 方法和 DCRN-MAE 方法.

$$e_{\text{MAE}} = \frac{1}{N} \sum_{k=0}^N |y_{k,\text{predicted}} - y_{k,\text{true}}|. \quad (5)$$

2.2.1 一般噪声

实验 1 采用镜像源法生成的 RIRs 作为初级路径和次级路径以验证 DCRN 在仿真环境中降噪性能.选择 4 种噪声情况进行测试,分别为 babble、engine、babble 转变为 engine 以及 babble 混合 factory.图 4a 和 4b 分别给出了 babble 和 engine 两种噪声下 FxLMS、Deep-ANC 和 DCRN 方法降噪后误差信号的功率谱对比.观察图 4 可以发现 DCRN 方法和 Deep-ANC 方法在 babble 和 engine 两种噪声下相较于 FxLMS 算法均表现出更好的降噪性能.传统的 FxLMS

算法降噪性能仅限于低频窄带部分,而 DCRN 方法在宽带和低频部分均表现出更高的降噪性能.

实际环境中可能存在噪声类型转换或多种噪声同时存在的情况,对 babble 转变为 engine 以及 babble 混合 factory 两种噪声情况进行测试,以验证 DCRN 方法的鲁棒性.图 5a 和 5b 分别给出了 babble 转变为 engine 噪声与 babble 混合 factory 噪声两种情况下, FxLMS、Deep-ANC 和 DCRN 方法降噪后误差信号的功率谱对比.观察图 5 可以发现,在噪声类型转变以及多种噪声混合的情况下,DCRN 方法相较于 FxLMS 算法在宽频带内降噪性能均有明显提升.实验 1 的结果表明,在仿真环境下 DCRN 方法在宽带噪声上均能够实现更好的降噪性能,并且在应对噪声变化的情况下表现出良好的鲁棒性.对比 DCRN-MAE 方法与 DCRN-MSE 方法可以发现在低频部分前者降噪量低于后者,表明在仿真环境中 DCRN-MAE 方法降噪性能略差于 DCRN-MSE 方法.

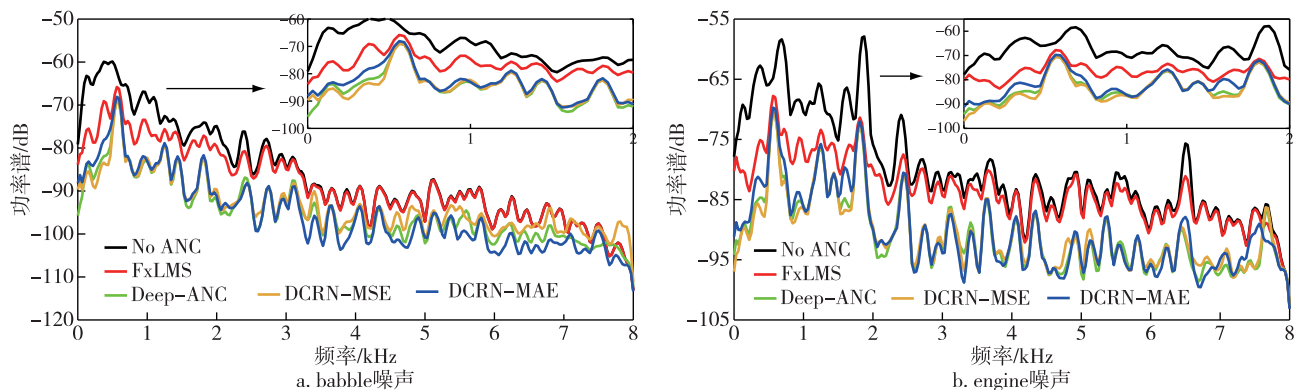


图 4 babble 与 engine 两种噪声降噪前后误差信号功率谱

Fig. 4 Error signal power spectrums before and after ANC a.babble noise;b.engine noise

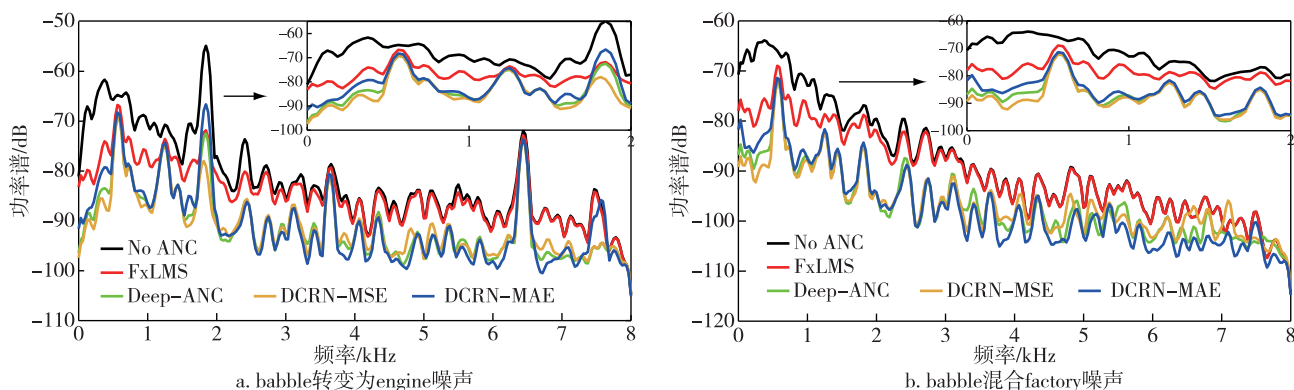


图 5 babble 转变为 engine 与 babble 混合 factory 两种情况降噪前后误差信号功率谱

Fig. 5 Error signal power spectrums before and after ANC a.from babble noise to engine noise;
b.babble noise mixed with factory noise

为了进一步验证 DCRN 方法在实际环境下的降噪性能, 实验 2 选择了有源降噪耳机场景进行验证. 图 6a 和 6b 分别给出了 babble 和 factory 两种噪声下, FxLMS、Deep-ANC 和 DCRN 方法降噪后误差信号的功率谱对比. 可以看出, 在有源降噪耳机环境下, DCRN 方法在 2 000 Hz 以内的低频部分相较于 FxLMS 算法降噪量提升 5~10 dB, 在整个宽频带内降噪性能略优于 Deep-ANC 方法. 图 7a 和 7b 分别给出了 babble 转变为 engine 噪声和 babble 混合 factory 噪声两种情况下, FxLMS、Deep-ANC 和 DCRN 方法降噪后误差信号的功率谱对比. 可以发现在有源降噪耳机环境下, 对于噪声类型转变和噪声混合的情况, DCRN 方法在 2 000 Hz 以内的低频部分相较于 FxLMS 算法降噪量均能提升 5~10 dB. 实验 2 的结果表明, 对于一般噪声, 在有源降噪耳机环境中 DCRN 方法在低频段相较于 FxLMS 算法和 Deep-ANC 方法均有降噪量的提升, 并且在应对噪声变化时表现出了较好的鲁棒性. 对比 DCRN-MAE 方法与 DCRN-MSE 方法可以发现在低频

段前者降噪量较于后者提升 5~10 dB, 并且解决了后者在高频部分能量不降反增的问题. 该结果表明, 在有源降噪耳机环境中对于一般噪声 DCRN-MAE 方法降噪性能优于 DCRN-MSE 方法.

2.2.2 风噪声

实验 3 采用模拟生成的房间脉冲响应 (RIRs) 作为初级路径和次级路径, 验证在仿真环境下, DCRN 方法降低风噪的性能. 选择户外采集的正常风速和强风速两种风噪声进行测试. 图 8a 和 8b 分别给出了正常风速和强风速两种风噪下, FxLMS、Deep-ANC 和两种 DCRN 方法降噪后误差信号的功率谱对比, 可以发现风噪声能量主要集中的 1 000 Hz 以内, 在正常风速和强风速两种风噪下两种 DCRN 方法相较于 FxLMS 算法降噪量均提高 10 dB 左右, DCRN-MAE 方法在整体降低风噪的能力上明显优于 Deep-ANC 和 DCRN-MSE 方法. 实验 3 验证了 DCRN 方法在风噪声环境中降低风噪声的可行性以及 DCRN-MAE 方法降低风噪的性能优于 DCRN-MSE 方法.

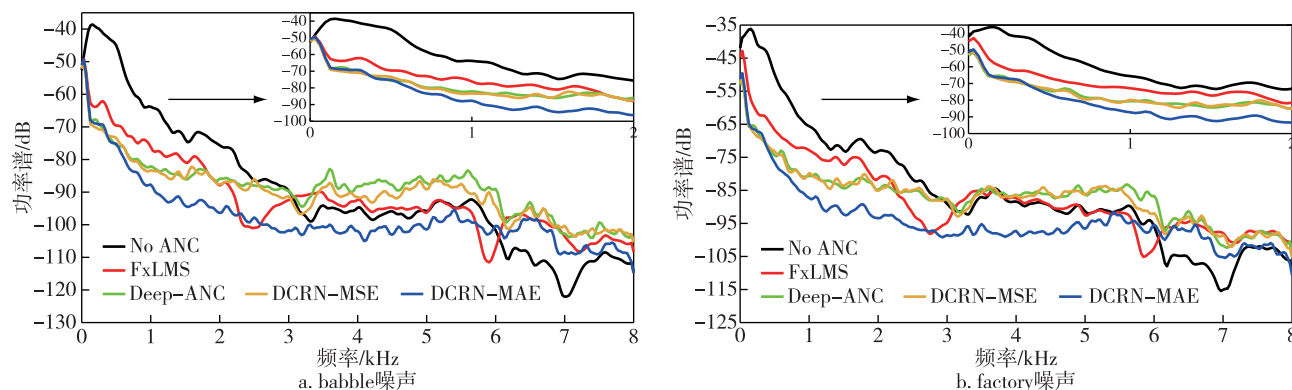


图 6 有源降噪耳机环境下 babble 与 factory 两种噪声降噪前后误差信号功率谱

Fig. 6 Error signal power spectrums before and after ANC in active headphone environment a. babble noise; b. factory noise

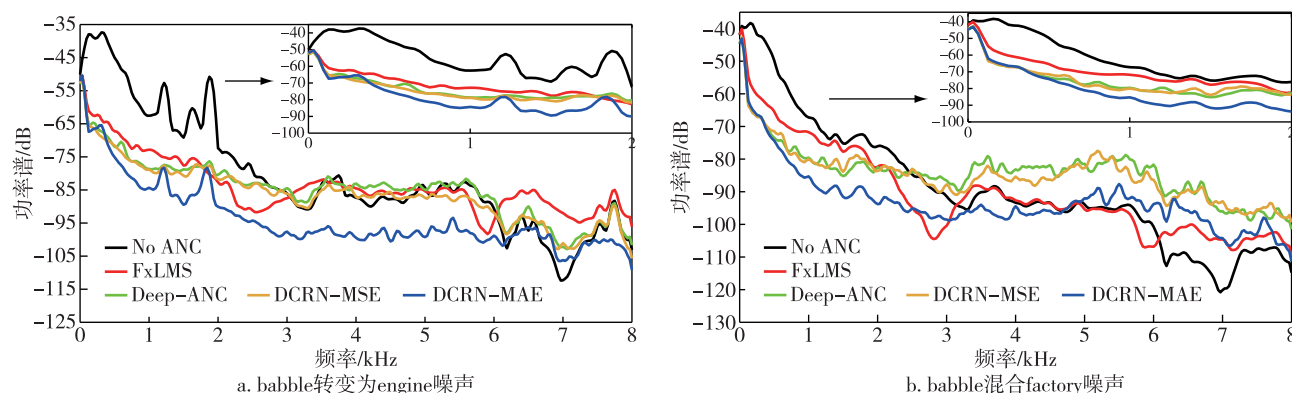


图 7 有源降噪耳机环境下 babble 变为 engine 与 babble 混合 factory 两种情况降噪前后误差信号功率谱

Fig. 7 Error signal power spectrums before and after ANC in active headphone environment

a. from babble noise to engine noise; b. babble noise mixed with factory noise

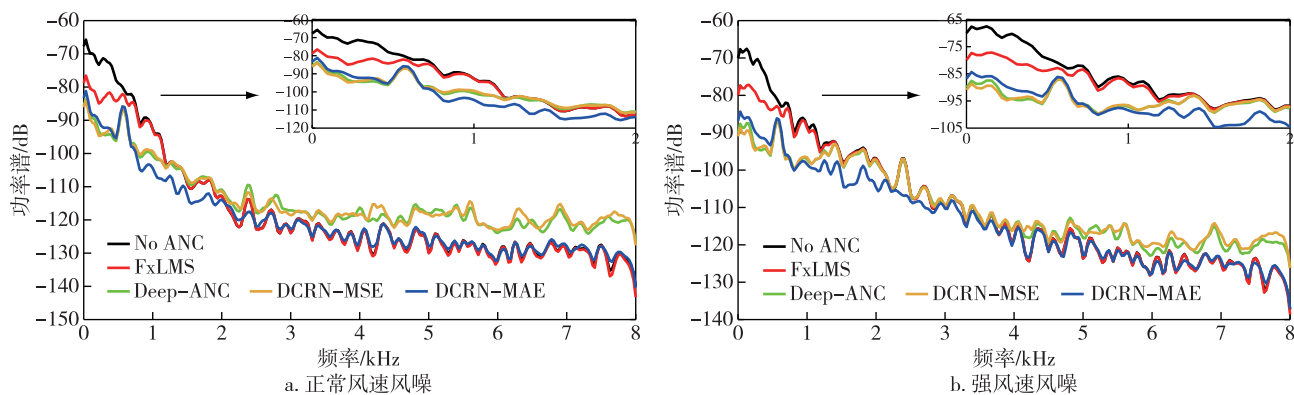


图8 仿真环境下正常风速和强风速两种风噪降噪前后误差信号功率谱

Fig. 8 Error signal power spectrums before and after ANC in simulated environment
a.normal speed wind noise;b.strong speed wind noise

从图8中可以观察到 DCRN-MSE 和 Deep-ANC 方法在风噪声的高频部分(1 000 Hz 以上)表现较差.这是由于风噪声信号在高频部分的能量滚降迅速,其能量较低频而言非常小,导致 DCRN-MSE 和 Deep-ANC 方法在训练模型时对风噪高频部分出现训练过拟合.对此可以通过正则化即提高风噪声信号高频部分的能量来解决模型出现过拟合的现象.考虑到风噪的特点,本文在风噪信号 1 000 Hz 频率以上叠加白噪声进行正则化处理.图9是风噪信号高频部分正则化处理前后的功率谱图.从图9中可以观察到正则化处理不改变风噪信号的主要结构,仅处理风噪信号高频部分不影响风噪能量主要集中在的低频段.

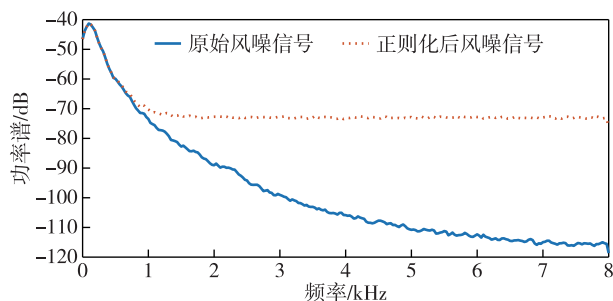


图9 风噪信号正则化处理前后功率谱

Fig. 9 Power spectrums of wind noise signal before and after regularization

实验4验证了 DCRN 方法在实际环境即有源降噪耳机中降低风噪的性能.风噪数据集的使用同实验3一致,不同的是对风噪数据集进行正则化处理.图10a和10b分别是在正常风速和强风速两种风噪

声情况下 FxLMS、Deep-ANC 和两种 DCRN 方法降噪后误差信号的功率谱对比.可以发现在风噪声能量主要集中在的 1 000 Hz 以内的低频段,DCRN-MSE 和 Deep-ANC 方法对于两种风速风噪情况仅在 600 Hz 以内相较于 FxLMS 算法有降噪量的提升,DCRN-MAE 方法在 1 000 Hz 以内相较于 DCRN-MSE 方法

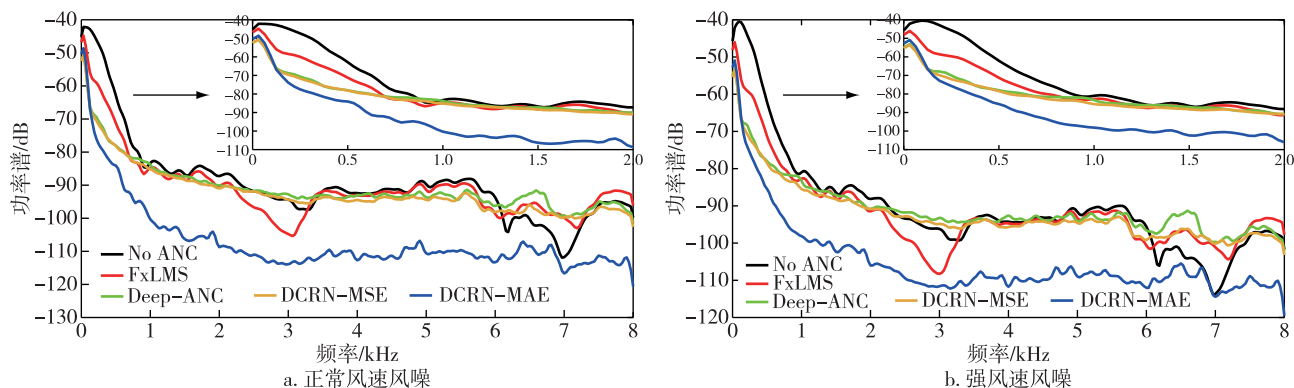


图10 有源降噪耳机环境下正常风速和强风速风噪降噪前后误差信号功率谱

Fig. 10 Error signal power spectrums before and after ANC in active headphone environment
a.normal speed wind noise;b.strong speed wind noise

降噪量均有显著提升.实验 4 结果表明,采用正则化训练数据后,在有源降噪耳机环境下 DCRN-MAE 方法在降低风噪性能上相较于 DCRN-MSE 方法具有明显优势.

实验 5 对比采用噪声信号目标幅度谱 (Target Magnitude Spectrum, TMS) 作为输入特征的 CRN 以验证融合相位估计的复频谱映射的优势.幅度谱映射的 CRN 具体结构如图 11 所示,相较于 DCRN 采用双解码模块分别估计网络输出信号的实部谱和虚部谱,幅度谱映射的 CRN 仅需一个解码模块即可. CRN 中编解码模块和时序模块中各层网络具体参数与 DCRN 相同.本文将采用幅度谱映射的 CRN 方法简称为 CRN-TMS 方法,训练 CRN 的数据集同实验 4 且采用 MAE 损失函数.图 12a 和 12b 分别是在正常风速和强风速两种风噪声情况下 FxLMS、Deep ANC、DCRN-MAE 方法和 CRN-TMS 方法降噪后误差信号的功率谱对比.可以观察到采用幅度谱映射的

CRN-TMS 方法降低风噪性能弱于传统 FxLMS 算法,表明 ANC 系统中相位信息的重要性,也表明采用复频谱的 DCRN-MAE 相较于仅采用幅度谱的 CRN-TMS 在融合相位信息后降低风噪具有明显的优势.

算法的计算复杂度在评估算法运行时消耗的计算资源量方面起主要作用.深度学习算法一般采用浮点数运算次数 (Floating Point Operations, FLOPs) (包括乘法运算次数和加法运算次数) 衡量模型计算复杂度.本文采用广泛使用的 Python 中 thop 库 (<https://pypi.org/project/thop/>) 计算 DCRN 的 FLOPs 和训练参数量.在计算复杂度方面,DCRN 模型中可训练参数约为 2.32×10^6 , 每个时间帧的浮点运算次数为 4.13×10^6 .对比算法 Deep-ANC 模型的可训练参数和浮点运算次数分别为 8.83×10^6 和 12.78×10^6 , DCRN 模型相较于 Deep-ANC 模型大幅降低了网络的可训练参数和计算复杂度.在同一硬件平台上,传统 FxLMS 算法、DCRN 方法与 Deep-

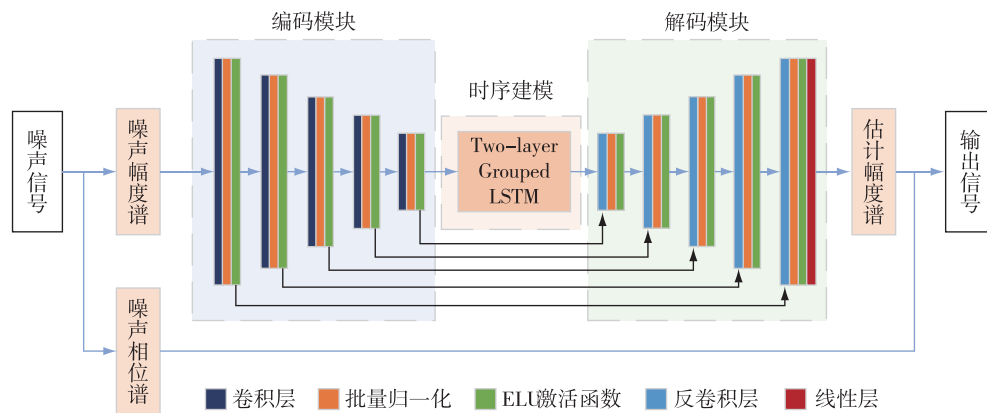


图 11 幅度谱映射的 CRN 具体结构

Fig. 11 Specific structure of CRN for magnitude mapping

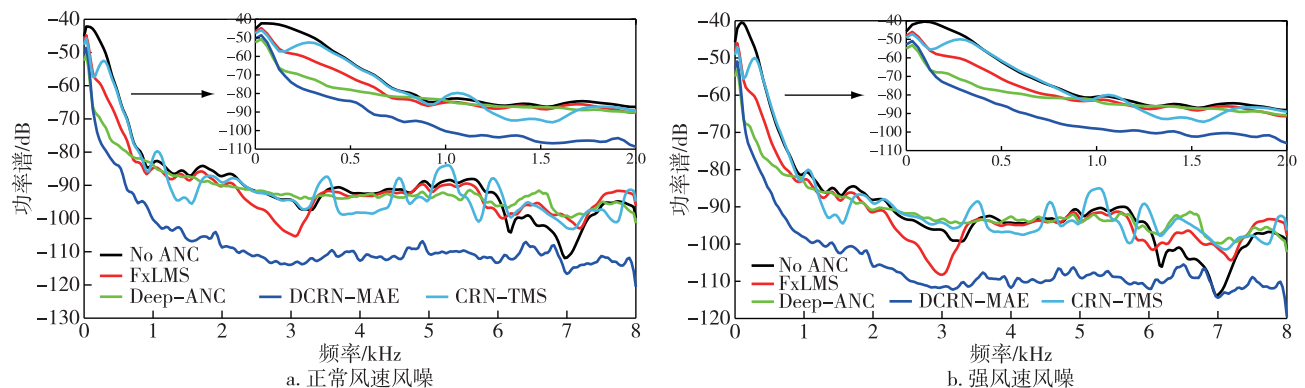


图 12 有源降噪耳机环境下正常风速和强风速风噪降噪前后误差信号功率谱

Fig. 12 Error signal power spectrums before and after ANC in active headphone environment

a. normal speed wind noise; b. strong speed wind noise

ANC方法对于1 s数据的平均处理时间(运行100次)如表2所示,DCRN方法与Deep-ANC方法比传统自适应滤波算法更快.这是由于深度学习网络训练完成后,其网络权重是固定的,因此可以利用硬件平台进行并行运算,从而减少运行时间.

表2 不同算法处理1 s数据的平均运行时间

Table 2 Average execution time for processing 1 second of data by different algorithms ms

	FxLMS	DCRN	Deep-ANC
运行时间	64	12	18

3 结论

本文研究了基于双解码卷积循环网络的DCRN有源噪声控制方法,利用DCRN代替自适应滤波器来编码,适用于不同噪声环境下ANC控制器最优控制参数,通过训练DCRN分别估计与参考信号对应的次级信号的实部和虚部,进而使抵消信号能够最大化衰减误差麦克风处的噪声残留.此外,本文采用复数频谱作为输入特征和训练目标,充分考虑了相位在有源噪声控制中的重要性.针对一般噪声与风噪声分别在仿真环境和有源降噪耳机环境下进行实验.实验结果表明,对于一般噪声DCRN方法具有良好的降噪性能和鲁棒性.在风噪声中,正则化训练数据后,采用MAE损失函数的DCRN方法相比于采用MSE损失函数的DCRN方法降噪风噪的性能明显提升.实验结果说明了损失函数对网络性能的显著影响.未来的研究将进一步探索采用不同形式的损失函数如基于对抗网络的损失函数,以尝试进一步提高网络性能.

参考文献

References

- [1] 张晓宇,仪垂杰.主、被动噪声控制的实验研究[J].噪声与振动控制,2011,31(3):145-148
ZHANG Xiaoyu, YI Chuijie. Experimental study on hybrid noise control[J]. Noise and Vibration Control, 2011, 31(3):145-148
- [2] 吴礼福,陈晶晶,郭业才.调节水床效应的双梯度有源噪声控制自适应算法[J].应用声学,2020,39(4):632-637
WU Lifu, CHEN Jingjing, GUO Yecai. Adaptive double-gradient active noise control algorithms for tuning waterbed effect[J]. Journal of Applied Acoustics, 2020, 39(4):632-637
- [3] Schumacher T, Krüger H, Jeub M, et al. Active noise control in headsets: a new approach for broadband feedback ANC [C]//2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). May 22 - 27, 2011, Prague, Czech Republic. IEEE, 2011: 417-420
- [4] Chang C Y, Li S T. Active noise control in headsets by using a low-cost microcontroller [J]. IEEE Transactions on Industrial Electronics, 2011, 58(5):1936-1942
- [5] Bradley S, Backman J, von Humerbein S, et al. The mechanisms creating wind noise in microphones [C]//Audio Engineering Society Convention 114. Audio Engineering Society, 2003:5718
- [6] Mirabilii D, Lodermeier A, Czwielong F, et al. Simulating wind noise with airflow speed-dependent characteristics [C]//2022 International Workshop on Acoustic Signal Enhancement (IWAENC). September 5 - 8, 2022, Bamberg, Germany. IEEE, 2022:1-5
- [7] 赖小强,李双田.双传声器系统中的风噪声抑制方法研究[J].信号处理,2013,29(4):436-442
LAI Xiaoqiang, LI Shuangtian. Research for wind noise reduction method in dual-microphone system [J]. Journal of Signal Processing, 2013, 29(4):436-442
- [8] Fisol U M M, Ripin Z M, Ismail N A, et al. Wind noise analysis of a two-way radio [C]//2013 IEEE International Conference on Smart Instrumentation, Measurement and Applications (ICSIMA). November 25-27, 2013, Kuala Lumpur, Malaysia. IEEE, 2014:1-5
- [9] King B, Atlas L. Coherent modulation comb filtering for enhancing speech in wind noise [C]//International Workshop on Acoustic Echo and Noise Control, 2008:14-17
- [10] 潘翔.基于风噪声抑制的语音信号增强研究[C]//信号处理在地球物理——浙江省信号处理学会2018年学术年会论文集.杭州,2018:100-105
- [11] Boll S. Suppression of acoustic noise in speech using spectral subtraction [J]. IEEE Transactions on Acoustics, Speech, and Signal Processing, 1979, 27(2):113-120
- [12] Chu Y J, Zhao S P, He L B, et al. Wind noise suppression in filtered-x least mean squares-based active noise control systems [J]. The Journal of the Acoustical Society of America, 2022, 152(6):3340-3345
- [13] 吴睿,金向锋.抑制风噪装置的实验[J].应用声学,2022,41(6):1010-1015
WU Rui, JIN Xiangfeng. Experiment on a structure for suppressing the wind noise [J]. Journal of Applied Acoustics, 2022, 41(6):1010-1015
- [14] 潘作峰,邓玉伟,郝耀东,等.动态密封状态下汽车风噪性能不确定性研究[J].汽车工程,2021,43(11):1645-1652,1661
PAN Zuofeng, DENG Yuwei, HAO Yaodong, et al. Study on uncertainty of vehicle wind noise performance under dynamic seal condition [J]. Automotive Engineering, 2021, 43(11):1645-1652,1661
- [15] Jiang Y H, Liu H Q, Zhou Y, et al. An integration development of traditional algorithm and neural network for active noise cancellation [C]//2022 IEEE 32nd International Workshop on Machine Learning for Signal Processing (MLSP). August 22-25, 2022, Xi'an, China. IEEE, 2022:1-6
- [16] Mostafavi A, Cha Y J. Deep learning-based active noise

- control on construction sites[J]. Automation in Construction, 2023, 151: 104885
- [17] Im S, Kim S, Woo S, et al. Deep learning-assisted active noise control in a time-varying environment[J]. Journal of Mechanical Science and Technology, 2023, 37(3): 1189-1196
- [18] Cha Y J, Mostafavi A, Benipal S S. DNoiseNet: deep learning-based feedback active noise control in various noisy environments[J]. Engineering Applications of Artificial Intelligence, 2023, 121: 105971
- [19] Zhang H, Wang D L. Deep ANC: a deep learning approach to active noise control[J]. Neural Networks, 2021, 141: 1-10
- [20] Kwon S, Kim B S, Park J. Active noise reduction with filtered least-mean-square algorithm improved by long short-term memory models for radiation noise of diesel engine[J]. Applied Sciences, 2022, 12(20): 10248
- [21] Tan K, Wang D L. Complex spectral mapping with a convolutional recurrent network for monaural speech enhancement [C]//I2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). May 12-17, 2019, Brighton, UK. IEEE, 2019: 6865-6869
- [22] 郭佳丽, 邢双云, 栾昊, 等. 基于改进的 LSTM 算法的时间序列流量预测[J]. 南京信息工程大学学报(自然科学版), 2021, 13(5): 571-575
GUO Jiali, XING Shuangyun, LUAN Hao, et al. Prediction of time series traffic based on improved LSTM algorithm [J]. Journal of Nanjing University of Information Science & Technology (Natural Science Edition), 2021, 13(5): 571-575
- [23] 韩莹, 管健, 曹允重, 等. LSTM-WBLS 模型在日降水量预测中的应用[J]. 南京信息工程大学学报(自然科学版), 2023, 15(2): 180-186
HAN Ying, GUAN Jian, CAO Yunzhong, et al. Application of improved LSTM-WBLS model in daily precipitation forecast[J]. Journal of Nanjing University of Information Science & Technology (Natural Science Edition), 2023, 15(2): 180-186
- [24] Park S R, Lee J. A fully convolutional neural network for speech enhancement [J]. arXiv e-Print, 2016, arXiv:1609.07132
- [25] Varga A, Steeneken H J M. Assessment for automatic speech recognition: II. NOISEX-92: a database and an experiment to study the effect of additive noise on speech recognition systems[J]. Speech Communication, 1993, 12(3): 247-251
- [26] Nelke C M, Vary P. Measurement, analysis and simulation of wind noise signals for mobile communication devices [C]//2014 14th International Workshop on Acoustic Signal Enhancement (IWAENC). September 8-11, 2014, Juan-les-Pins, France. IEEE, 2014: 327-331
- [27] Allen J B, Berkley D A. Image method for efficiently simulating small-room acoustics [J]. The Journal of the Acoustical Society of America, 1979, 65(4): 943-950
- [28] Breining C, Dreiscitel P, Hansler E, et al. Acoustic echo control. An application of very-high-order adaptive filters [J]. IEEE Signal Processing Magazine, 1999, 16(4): 42-69
- [29] Reddi S J, Kale S, Kumar S. On the convergence of Adam and beyond[J]. arXiv e-Print, 2019, arXiv:1904.09237

Active control of wind noise with dual-decoder convolutional recurrent network

WU Lifu^{1,2} GE Wenchang¹ CHEN Chen¹ WANG Shaobo¹

¹ School of Electronics & Information Engineering, Nanjing University of Information Science & Technology, Nanjing 210044, China

² Collaborative Innovation Center of Atmospheric Environment and Equipment Technology, Nanjing University of Information Science & Technology, Nanjing 210044, China

Abstract Here, an Active Noise Control (ANC) approach is proposed which replaces Filtered-x Least Mean Square (FxLMS) algorithm with Dual-decoder Convolutional Recurrent Network (DCRN). Due to the importance of phase information in ANC, the input feature of DCRN is the complex spectrogram of the noise signal (including real and imaginary spectrograms). In the network structure, a coding module is used to extract features from the noise complex spectrograms, and a dual-decoder module is used to estimate the real and imaginary spectrograms of the network output. Parameter sharing mechanism and group strategy are adopted to reduce the number of training parameters and improve the learning ability and generalization performance. Especially for wind noise, a new loss function is adopted and the training data are regularized to improve the performance of DCRN. Experiments in both simulation and ANC headphone environments show that the DCRN approach exhibits good noise reduction performance and robustness for both general noise and wind noise.

Key words dual-decoder convolutional recurrent network (DCRN); active noise control (ANC); filtered-x least mean square (FxLMS) algorithm; complex spectrogram