

基于 ResNet 的稳健语音 DOA 估计算法

郭业才^{1,2} 刘力玮¹ 顾弘毅¹

(1. 南京信息工程大学电子与信息工程学院, 南京, 210044; 2. 南京信息工程大学滨江学院, 无锡, 214105)

摘要: 针对传统波达方向(Direction of arrival, DOA)估计算法在阵列模型误差条件下估计性能下降的问题, 提出了一种基于残差网络(Residual network, ResNet)的 DOA 估计算法。该算法根据神经网络数据驱动不依赖阵列流型的特点, 从广义互相关(Generalized cross-correlation, GCC)中提取特征, 将提取的特征作为神经网络深层分类器的输入, 对信号分类; 根据分类结果选取对应子区间数据进行训练, 建立 ResNet 学习特征与 DOA 估计之间的非线性映射关系, 形成一个数据驱动的稳健 DOA 估计系统。仿真与实验验证结果表明, 该算法有效地解决了传统 DOA 算法在阵列模型误差条件下无法准确得到 DOA 结果的问题。

关键词: DOA 估计; 阵列模型误差; 广义互相关; 残差神经网络

中图分类号: TP391 **文献标志码:** A

ResNet-Based DOA Algorithm Speech Estimation with Robustness

Guo Yecai^{1,2}, Liu Liwei¹, Gu Hongyi¹

(1. College of Electronic and Information Engineering, Nanjing University of Information Science and Technology, Nanjing, 210044, China; 2. Binjiang College, Nanjing University of Information Science and Technology, Wuxi, 214105, China)

Abstract: Aiming at the performance degradation of traditional DOA (direction of arrival) estimation algorithm under the condition of array model error, a DOA estimation algorithm based on ResNet (residual network) is proposed. According to the characteristics of data-driven neural network that it does not depend on array flow pattern, the proposed algorithm extracts features from generalized cross-correlation (GCC), and takes the extracted features as input of neural network deep classifier to classify signals. On the basis of the classification results, the corresponding sub-interval data are selected for training and the non-linear mapping relationship between ResNet learning features and DOA estimation to form a data-driven robust DOA estimation system. The simulation and experimental results show that the proposed algorithm can effectively solve the problem that the traditional DOA algorithm cannot accurately obtain the DOA results under the condition of array model error.

Key words: DOA estimation; array imperfection; generalized cross-correlation; residual network

引言

波达方向(Direction of arrival, DOA)估计是阵列信号处理的重要方向之一,其广泛应用在远程自动语音识别^[1],电话会议和自动摄像机转向^[2]等方面。然而,当阵列流型存在缺陷时,要得出准确的DOA估计是很困难的。因此,需要一种可以在缺少阵列流型误差的先验信息时进行鲁棒DOA估计的方法。

传统DOA估计方法主要可以分为:(1)子空间方法,例如多重信号分类(Multiple signal classification, MUSIC)^[3]和借助旋转不变技术估计信号参数技术(Estimating signal parameters via rotational invariance techniques)^[4];(2)广义互相关(Generalized cross-correlation, GCC)^[5]和最小二乘法(Least squares, LS)^[6];(3)信号同步方法,例如基于联合可控响应功率和相位变换(Steered response power using the phase transform weight, SRP-PHAT)方法^[7]和多通道互相关(Multichannel cross-correlation coefficient, MCCC)方法^[8];(4)脉冲响应的盲识别方法,例如自适应特征值分解(Adaptive eigenvalue decomposition, AED)算法^[9]和独立分量分析方法^[10];(5)基于 l_1 范数惩罚的稀疏信号表示方法^[11];(6)基于模型的方法,例如最大似然法^[12]。上述方法计算成本高,并且依赖对阵列流型的假设。

针对上述传统方法存在的问题,近几年出现了基于神经网络的DOA估计方法^[13-14]。这些方法只对语音信号做简单的特征处理,如计算信号的协方差矩阵或广义互相关,利用神经网络完成信号特征到DOA结果之间的映射关系。这些方法将DOA视作分类问题,没有对信号做过强的假设,一定程度上克服了传统方法的缺点,取得了良好的DOA估计效果。

近年来,深度卷积神经网络在图像处理研究中取得一系列突破。网络深度是深度卷积网络的一个重要参数,更深的网络能强化特征提取和非线性映射能力,但堆叠过深会产生梯度爆炸、消失的问题。尽管此问题可以通过增加初始归一化^[15]和中间归一化层解决,但随着网络深度的进一步增加,拟合精度变饱和后迅速退化。残差卷积网络^[16]一定程度上解决了这一问题,通过更深的网络提高拟合精度。

结合以上背景,本文在文献[13]的基础上,将残差网络(Residual network, ResNet)应用于声源DOA估计中,并根据神经网络数据驱动,不依赖阵列流型的特点,提出了一种在阵列误差条件下基于ResNet的稳健语音DOA估计算法。与其他基于神经网络的DOA估计方法不同,该算法没有将DOA估计视作分类问题,使用基于欧几里德距离的代价函数对DOA进行回归建模。将整个定位范围分若干个子区间,使用一个基于神经网络的深层分类器对信号分类,根据分类结果选取对应子区间数据训练得到的ResNet对信号DOA估计,即为本文的基于ResNet的缺陷阵列DOA语音估计。

1 阵列模型

1.1 理想阵列模型

设均匀线型麦克风阵列有 M 个阵元,阵元间距为 d ,每个阵元都是相同的全向麦克风,远场信号以 θ 入射。假设噪声为与入射信号独立的高斯白噪声,均值为0,方差为 σ^2 ,则阵列在 k 时刻的输出为

$$\mathbf{x}(k) = \mathbf{a}(\theta)\mathbf{s}(k) + \mathbf{n}(k) \quad (1)$$

式中: $\mathbf{a}(\theta)$ 为 $M \times 1$ 维的阵列流型矩阵, $\mathbf{s}(k)$ 为 $1 \times K$ 维的目标信源复振幅矢量, $\mathbf{n}(k)$ 为 $M \times K$ 维的加性噪声复矢量。

$$\mathbf{a}(\theta) = \frac{1}{\sqrt{M}} [1 \quad e^{j\varphi(\theta)} \quad \dots \quad e^{j(M-1)\varphi(\theta)}] \quad (2)$$

式中: $\varphi(\theta) = \frac{2\pi d}{\lambda} \sin\theta, \theta \in \left[-\frac{\pi}{2}, \frac{\pi}{2} \right]$ 。

1.2 考虑误差的阵列模型

考虑2种典型的阵列误差,增益、相位误差以及麦克风位置误差。本文使用的增益误差矢量为

$$\mathbf{e}_{\text{gain}} = \rho \times [0, 0.2, \dots, 0.2, -0.2, \dots, -0.2]_{1 \times M}^T \quad (3)$$

式中: $\rho \in [0, 1]$ 为控制阵列误差程度的系数, M 为阵元个数。相位误差矢量为

$$\mathbf{e}_{\text{phase}} = \rho \times [0, -30^\circ, \dots, -30^\circ, 30^\circ, \dots, 30^\circ]_{1 \times M}^T \quad (4)$$

位置误差矢量为

$$\mathbf{e}_{\text{pos}} = \rho \times [0, -0.2, \dots, -0.2, 0.2, \dots, 0.2]_{1 \times M}^T \times d \quad (5)$$

式中, d 为均匀线阵的间距。加入阵列误差的阵列流型为

$$\mathbf{a}(\theta, \mathbf{e}_{\text{gain}}, \mathbf{e}_{\text{phase}}, \mathbf{e}_{\text{pos}}) = (\mathbf{I}_M + \text{diag}(\delta_{\text{gain}} \mathbf{e}_{\text{gain}})) \times \text{diag}(\exp(\mathbf{j} \delta_{\text{phase}} \mathbf{e}_{\text{phase}})) \times \mathbf{a}(\theta, \delta_{\text{pos}} \mathbf{e}_{\text{pos}}) \quad (6)$$

式中: $\delta(\cdot)$ 为控制误差是否存在的参数, \mathbf{I}_M 为 $M \times M$ 的酉矩阵, $\text{diag}(\cdot)$ 为对角线设置为给定矢量的对角矩阵。虽然式(6)的阵列流型与实际情况相比简化了很多,但在利用神经网络进行DOA估计的过程中整个系统没有用到有关阵列误差的先验信息,这里的阵列流型仅用来生成仿真数据,因此神经网络DOA估计的稳健性并不是针对特定的阵列误差,可以认为这里的简化是合理的。

2 基于ResNet的语音DOA估计

基于广义互相关的传统DOA估计算法都存在过强的假设,例如不存在混响、噪声与信号之间互不相关、阵列模型为均匀线阵或圆阵等。因此,极端环境下经典DOA估计算法往往会有较大偏差甚至失效。针对此问题,本文使用ResNet完成基于GCC的DOA估计结果的映射。由于ResNet的卷积结构,将输入特征以一对麦克风对的GCC为一行,堆叠成矩阵,卷积时可以获取不用麦克风对之间的特征信息。

传统的基于GCC的DOA估计算法需要先通过GCC估计时延,再通过时延得到信号入射角度。通过时延估计入射角的过程依赖于阵列流型,因此当阵列模型存在与理想模型偏差较大时,DOA估计误差会明显增大。而神经网络由数据驱动,DOA估计的过程不依赖于阵列流型,仍然可以在阵列模型误差条件下获得良好的DOA估计效果。

2.1 特征提取

选取GCC-PHAT作为输入特征的基础。先将麦克风阵列接收的语音信号分为长0.1 s的语音帧,对每帧语音做GCC-PHAT处理。根据阵列模型参数和环境参数对每帧语音尺寸裁剪,得到子特征矩阵。最后,将所有子特征矩阵加权求和得到语音信号的特征矩阵。

2.1.1 GCC-PHAT

设阵列模型中的第 m 和 n 个麦克风阵元接收的信号经过离散采样后分别为 $x_m(k)$ 和 $x_n(k)$,则

$$x_m(k) = a_m s(k - \tau_m) + n_m(k) \quad (7)$$

$$x_n(k) = a_n s(k - \tau_n) + n_n(k) \quad (8)$$

式中: $n_m(k), n_n(k)$ 分别为两个麦克风接收信号中的噪声; a_m, a_n 为信号幅度衰减因子; τ_m, τ_n 为声源信号传播到2个麦克风所用的时间采样点数量; $s(k)$ 为声源信号。

忽略混响的影响, $x_m(k)$ 和 $x_n(k)$ 的相关函数为

$$R_{mn}(\tau) = E[x_m(k)x_n(k - \tau)] \quad (9)$$

将式(7,8)代入式(9)得

$$\begin{aligned} R_{mn}(\tau) = E\{[a_m s(k - \tau_m) + n_m(k)][a_n s(k - \tau_n - \tau) + n_n(k - \tau)]\} = \\ a_m a_n R_{ss}[\tau - (\tau_m - \tau_n)] + a_m R_{s n_m}(\tau - \tau_m) + a_n R_{s n_n}(-\tau - \tau_n) + R_{n_m n_n}(\tau) \end{aligned} \quad (10)$$

设 $s(k)$, $n_m(k)$ 和 $n_n(k)$ 互不相关, 则式(3)可写为

$$R_{mn}(\tau) = a_m a_n R_{ss}(\tau - \tau_{mn}) \quad (11)$$

式中, $\tau_{mn} = \tau_m - \tau_n$, $R_{ss}(\tau)$ 为声源 $s(k)$ 的自相关函数。

当 $\tau - \tau_{mn} = 0$ 时, $R_{mn}(\tau)$ 取最大值。因此可由 $R_{mn}(\tau)$ 的最大值估计两个麦克风阵元接收信号的时延采样点 τ_{mn} 。

由互相关函数和互功率谱的关系, 可得

$$R_{mn}(\tau) = \frac{1}{N} \sum_{n=0}^{N-1} G_{mn}(n) e^{j\frac{2\pi}{N}\tau n} = \frac{1}{N} \sum_{n=0}^{N-1} X_m(n) X_n^*(n) e^{j\frac{2\pi}{N}\tau n} \quad (12)$$

GCC 通过在式(4)上加权函数, 得

$$R_{mn}(\tau) = \frac{1}{N} \sum_{n=0}^{N-1} W_{mn}(n) G_{mn}(n) e^{j\frac{2\pi}{N}\tau n} \quad (13)$$

式中: $W_{mn}(n)$ 为加权函数, $G_{mn}(n)$ 为 2 个信号的互功率谱。PHAT 加权函数 $W_{mn}(n) = \frac{1}{|G_{mn}(n)|}$, 其效果等价于白化滤波。

2.1.2 特征矩阵构建

设均匀线阵由 M 个阵元, 阵列间距为 d , 阵元麦克风之间两两组合总共有 $m = C_M^2$ 对麦克风对, 对每 0.1 s 信号计算 GCC 并做相位变换(GCC-PHAT), 2 个麦克风阵元之间的最大间隔为 $(M-1)d$, 阵元之间存在的最大时延为 $\tau = \frac{(M-1)d}{c}$ (设声速为 c), 设入射信号采样率为 f_s (单位 Hz), 代表时延的 GCC 波峰一定会出现在中间 $n = f_s \times \tau$ 个点中。因此, 输入矩阵维度为 $m \times n$ 。

由于语音信号的特点, 在对语音信号分帧处理时可能有一些非语音帧, 为了更强的鲁棒性对每个语音帧加权^[16]

$$\mathbf{o} = \sum_{m=1}^L \omega_m \mathbf{o}_m \quad (14)$$

$$\omega_m = \frac{\sum_{d=1}^D |o_m(d)|^a}{\sum_{m=1}^L \sum_{d=1}^D |o_m(d)|^a} \quad (15)$$

式中: \mathbf{o}_m 为第 m 个语音帧, D 为 GCC 矩阵的元素数, $o_m(d)$ 为 GCC 矩阵的第 d 个元素, $|\cdot|$ 为取绝对值, a 为控制参数, 若 $a = 0$ 则获得 GCC 向量的平均值。使用大的 a 可有效降低静音帧对 GCC 矩阵的影响。

2.2 深层分类器结构

深层分类器是一个神经网络结构, 如图 1 所示。分类器输入为基于 GCC 的特征矩阵, 分类器的前向传播为

$$\text{out}_{i+1} = f(W_i \text{out}_i + b_{i+1}), i = 1, 2, 3, \dots, p-1 \quad (16)$$

式中: out_i 为第 i 层神经元的输出, W_i 为第 i 层与第 $i+1$ 层之间的神经元连接权重, b_i 为第 $i+1$ 层的神经元偏置, p 为神经网络层数, $f(\cdot)$ 为神经元激活函数。 out_1 为网络输入, 即基于 GCC-PHAT 的特征矩阵, out_p 为网络的输出, 即信号所在的子区间编号。

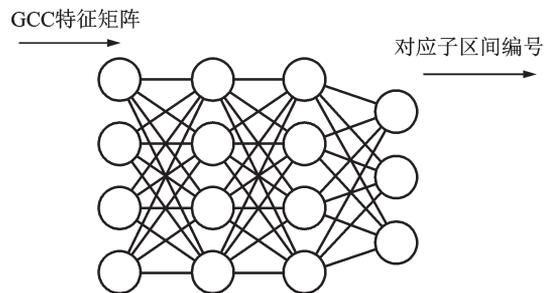


图 1 深层分类器的结构

Fig.1 Structure of deep classifier

输出层的激活函数取 softmax 函数,损失函数为

$$J(\theta) = -\frac{1}{m} \left[\sum_{i=1}^m \sum_{j=1}^n 1\{y^{(i)}=j\} \log \frac{e^{\theta_j^T x}}{\sum_{l=1}^n e^{\theta_l^T x}} \right] \quad (17)$$

式中: m 为样本个数, n 为网络输出层神经元个数(与子区间个数相等), $1\{\cdot\}$ 为示性函数, $y^{(i)}$ 为第*i*个样本的标签, $\theta_j^T x$ 为输出层第*j*个神经元的输入。再通过误差的反向传播,更新整个分类器网络的权重。

2.3 ResNet 网络结构

ResNet 与其他卷积神经网络(Convolutional neural network, CNN)的区别在于其残差结构。残差结构的典型结构如图 2 所示,设期望映射为 $H(x)$,则残差结构的期望映射为 $F(x)=H(x)-x$,最终的输出结果为 $F(x)+x$ 。

$F(x)+x$ 通过外加求和单元的前馈神经网络实现,增加求和单元不会增加参数,整个网络可以按照原来的反向传播算法进行训练。

改良的 ResNet 将批归一化层(Batch normalization, BN)和激活层(ReLu)的位置由非线性层的后面转移到前面^[17],改良的 ResNet 结构,如图 3 所示。图中除第一个卷积块以外所有卷积块都由 BN、Re-lu 和卷积层依次连接得到。

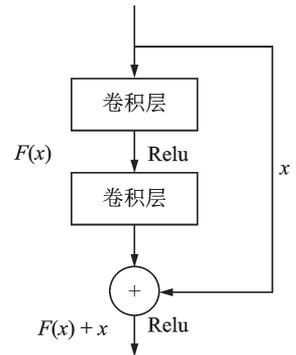


图 2 残差单元结构
Fig.2 Residual element structure

每个残差求和层(Sum)前有 2 个 3×3 的卷积块,每次卷积时加上尺寸为 1 的 Pad 保持尺寸不变。

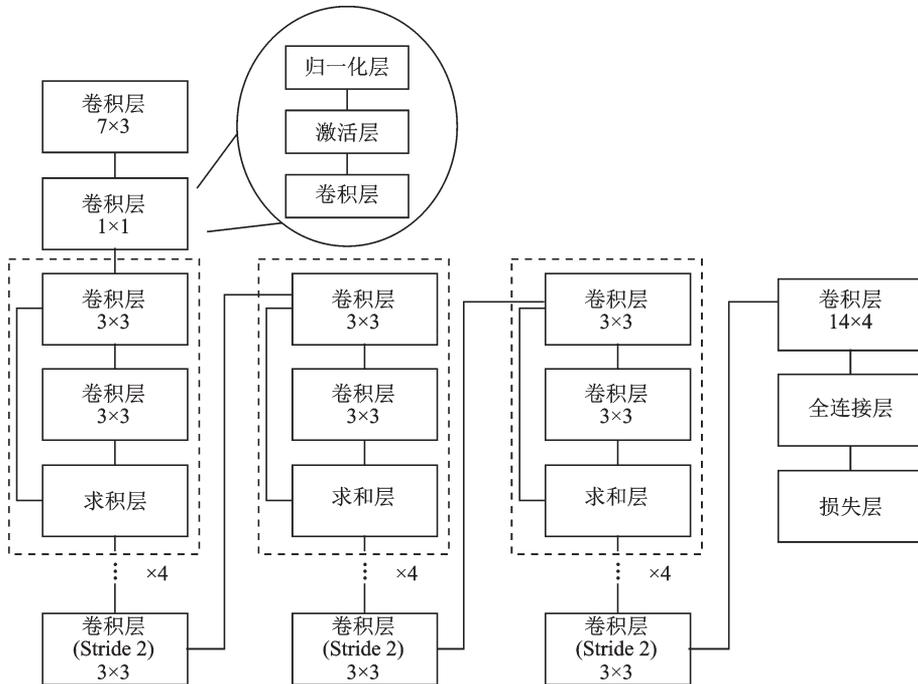


图 3 ResNet 结构
Fig.3 Structure of ResNet

每4次残差以后的一个卷积块选用 3×3 的卷积核,加1尺寸的Pad,以2为步长对输入进行卷积,缩小尺寸1/2,通道数加倍。缩小3次后在经过一个卷积块将尺寸收缩到 1×1 ,通过全连接层,送入Loss层。

网络使用基于欧几里德距离的损失函数为

$$J = \frac{1}{m} \sum_i^m \frac{1}{2} \|y_i - Y_i\|_2^2 \quad (18)$$

式中: y_i 为第*i*个样本的标签, Y_i 为第*i*个样本特征输入网络后网络的输出, m 为样本个数。

2.4 系统性能指标

采用均方根误差(Root mean square error, RMSE)来比较本文算法、基于多层感知器(Multi-layer perceptron, MLP)的算法和LS-TDOA算法的准确性和稳定性。RMSE定义为

$$\text{RMSE} = \sqrt{\frac{\sum_{r=1}^R (\theta_s - \hat{\theta}_s^r)^2}{R}} \quad (19)$$

式中: R 为测试使用的样本数, θ_s 为信号的真实入射角度, $\hat{\theta}_s^r$ 为DOA系统得出的估计入射角度。

3 实验和仿真

3.1 仿真与实验条件

仿真实验条件,如图4所示,房间尺寸为 $5.5 \text{ m} \times 3.3 \text{ m} \times 2.3 \text{ m}$,均匀阵列放置在高度为1 m位置,阵列的阵元间隔为0.2 m,阵元数量为8个。房间的四面墙壁和天花板反射系数为0.95(普通石灰墙),地板的反射系数为0.90。

数据集声源信号选取纯净的语音信号,入射角度以 0.1° 遍历各个分组的定位范围,按7:3的比例随机抽取训练集和测试集。数据集规模约为4.5万个。将音响设置在参考阵元的不同角度作为声源,实验数据由一个8阵元线型麦克风阵列采集,阵元间距为0.2 m。

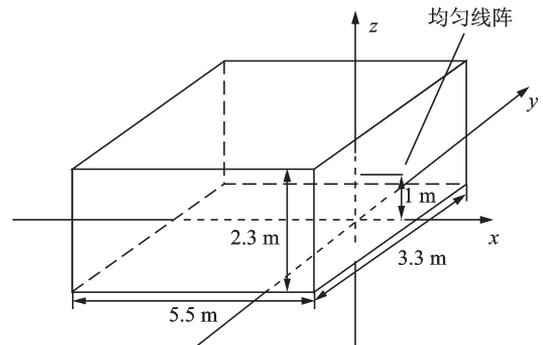


图4 仿真条件示意图

Fig.4 Schematic of simulation conditions

3.2 结果分析

3.2.1 子区间个数对DOA性能的影响

划分的子区间个数对系统性能的影响,如图5所示。图中横坐标代表测试信号的信噪比,纵坐标为估计结果与实际结果之间的均方根误差。图5表明,随着*P*的增大,每个子区间的宽度减小,用于DOA估计的ResNet的泛化压力随之降低,因此系统性能提升。当*P*大于6以后,继续增大*P*对DOA系统的提升不再明显。

3.2.2 阵列误差对DOA估计性能的影响

阵列误差的DOA估计性能的影响,如图6所示。图中横坐标为控制阵列误差大小的参数 ρ ,纵坐标为估计结果与实际结果之间的均方根误差,3条曲线分别为MUSIC算法、基于MLP的语音DOA估计算法^[13]和本文提出的基于ResNet的语音估计算法。可以看出MUSIC算法虽然在 ρ 较小时性能更强,但随着阵列流型与理想流型相差增大,估计性能显著下降。基于MLP的语音估计算法虽然表现出了一定的稳健性,但由于其结构简单,泛化能力不够强,性能仍有一定下降。基于ResNet的DOA估计算法由于增加了分类器结构,降低了泛化压力,又得益于ResNet比MLP更强的建模能力,对阵列误差的

稳健性更强。

3.2.3 信噪比对性能的影响

信噪比对系统性能的影响,如图7所示。图中将本文算法、基于MLP的DOA估计算法以及MUSIC算法在不同信噪比下的性能进行对比,图中实线为使用约3000个仿真测试信号在随机角度位置的测试结果,图中虚线为使用约100个语音信号在消声室3个固定角度的测试结果。

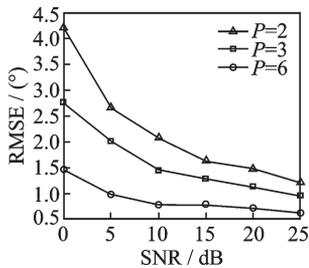


图5 划分的子区间个数对系统性能的影响

Fig.5 Influence of the number of subintervals on system performance

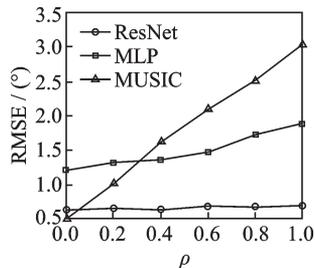


图6 阵列误差对DOA估计性能的影响

Fig.6 Effect of array error on DOA estimation performance

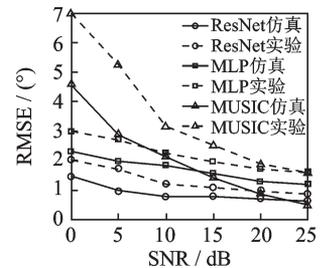


图7 信噪比对系统性能的影响

Fig.7 Influence of signal to noise ratio on system performance

由图7表明,在高信噪比时几种算法之间的差异不是很大,但信噪比在10 dB以下时,MUSIC算法性能显著下降,而2种基于神经网络的DOA算法仍能有效工作,且本文提出的算法性能更佳。由于实验器材的电噪声和测量时引入的不可避免的误差等原因,几种算法的RMSE均有所上升,但本文算法在低信噪比下的优势仍十分明显。

4 结束语

本文首次使用ResNet解决了阵列误差条件下语音信号的定位问题。通过增加分类器结构,降低了神经网络的泛化压力。利用ResNet强大的非线性映射能力,在普通石灰墙房间和消声室环境中均取得了良好的效果,证明了基于ResNet的DOA估计算法对阵列误差的稳健性。

但在极低信噪比下,神经网络仍出现了一定程度的过拟合现象,如果扩大数据集的同时使用更深的网络结构可能会获得更好的效果。

参考文献:

- [1] Xiao X, Zhao S, Nguyen D H H, et al. The NTU-ADSC systems for reverberation challenge 2014[C]// Proc REVERB challenge workshop. Florence: [s.n.], 2014.
- [2] Zhao S, Chng E S, Hieu N T, et al. A robust real-time sound source localization system for olivia robot[C]// 2010 APSIPA Annual Summit and Conference. Singapore: [s.n.], 2010.
- [3] Schmidt R. Multiple emitter location and signal parameter estimation[J]. IEEE Transactions on Antennas and Propagation, 1986, 34(3): 276-280.
- [4] Roy R, Kailath T. ESPRIT-estimation of signal parameters via rotational invariance techniques[J]. IEEE Transactions on Acoustics, Speech, and Signal Processing, 1989, 37(7): 984-995.
- [5] Knapp C, Carter G. The generalized correlation method for estimation of time delay[J]. IEEE Transactions on Acoustics, Speech, and Signal Processing, 1976, 24(4): 320-327.
- [6] Huang Y, Benesty J, Elko G W, et al. Real-time passive source localization: A practical linear-correction least-squares

- approach[J]. *IEEE transactions on Speech and Audio Processing*, 2001, 9(8): 943-956.
- [7] DiBiase J H. A high-accuracy, low-latency technique for talker localization in reverberant environments using microphone arrays [M]. Providence: Brown University, 2000: 73-83.
- [8] Benesty J, Chen J, Huang Y. Time-delay estimation via linear interpolation and cross correlation[J]. *IEEE Transactions on Speech and Audio Processing*, 2004, 12(5): 509-519.
- [9] Benesty J. Adaptive eigenvalue decomposition algorithm for passive acoustic source localization[J]. *The Journal of the Acoustical Society of America*, 2000, 107(1): 384-391.
- [10] Lombard A, Zheng Y, Buchner H, et al. TDOA estimation for multiple sound sources in noisy and reverberant environments using broadband independent component analysis[J]. *IEEE Transactions on Audio, Speech, and Language Processing*, 2011, 19(6): 1490-1503.
- [11] Malioutov D, Cetin M, Willsky A S. A sparse signal reconstruction perspective for source localization with sensor arrays[J]. *IEEE Transactions on Signal Processing*, 2005, 53(8): 3010-3022.
- [12] Stoica P, Sharman K C. Maximum likelihood methods for direction-of-arrival estimation[J]. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 1990, 38(7): 1132-1143.
- [13] Xiao X, Zhao S, Zhong X, et al. A learning-based approach to direction of arrival estimation in noisy and reverberant environments[C]// *IEEE International Conference on Acoustics, Speech and Signal Processing*. South Brisbane: IEEE, 2015: 2814-2818.
- [14] Faye A, Youm A B, Ndaw J D. LVQ based DOA estimation[C]// *Computational Intelligence, Communication Systems and Networks (CICSyN), 2013 Fifth International Conference on*. [S.l.]: IEEE, 2013: 245-250.
- [15] Glorot X, Bengio Y. Understanding the difficulty of training deep feedforward neural networks[C]// *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*. Chia Laguna: [s.n.], 2010: 249-256.
- [16] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]// *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Seattle: [s.n.], 2016: 770-778.
- [17] He K, Zhang X, Ren S, et al. Identity mappings in deep residual networks[C]// *European Conference on Computer Vision*. Cham: Springer, 2016: 630-645.

作者简介:



郭业才(1962-),男,博士,教授,博士生导师,研究方向:通信信号处理、水声信号处理等,E-mail: guo-ye-cai@163.com。

刘力玮(1996-),男,硕士研究生,研究方向:阵列信号处理。

顾弘毅(1993-),男,硕士研究生,研究方向:阵列信号处理。

(编辑:张彤)