

面向船载远程会议的麦克风阵列高精度 DOA 估计 *

刘雨佶^{1,2,3},童 峰^{1,2,3},陈东升^{1,2,3},卢荣富⁴,冯万健⁴

(1.厦门大学 水声通信与海洋信息技术教育部重点实验室,福建 厦门 361002;2.厦门大学 海洋与地球学院,福建 厦门 361002;
3.厦门大学深圳研究院,广东 深圳 518000;4.厦门亿联网络技术股份有限公司,福建 厦门 361000)

摘要: 随着船舶智能化水平提高,船载远程会议系统对提高应急处理能力、推进船岸一体化网络建设有重要意义,麦克风阵列是保证远程会议系统语音效果和支持多模态交互的重要语音前端。但船舶舱室狭小尺寸一方面导致只能采用小尺寸麦阵,另一方面小舱室导致的强混响以及嘈杂舱室噪声也使传统麦克风阵列算法性能严重下降。考虑船舶舱室复杂环境下小尺寸麦阵 DOA 估计场景,提出了一种轻量级 Mask–DOA 估计神经网络模型。该方法在 DOA 估计神经网络引入 Mask 算法降低噪声和混响的干扰,并提取增强后的 GCC–PHAT 作为网络特征,从而在小尺寸麦阵上实现高精度 DOA 估计。仿真和实验结果表明,所提出的 Mask–DOA 模型面对复杂的船舶舱室环境更鲁棒,泛化能力更强。

关键词: 波达方向估计;船舶舱室噪声混响环境;神经网络;时频掩蔽

中图分类号: TN912.3

文献标识码: A

DOI:10.16157/j.issn.0258-7998.212108

中文引用格式: 刘雨佶,童峰,陈东升,等. 面向船载远程会议的麦克风阵列高精度 DOA 估计[J]. 电子技术应用,2022,48(3):32–36,77.

英文引用格式: Liu Yuji, Tong Feng, Chen Dongsheng, et al. High precision DOA estimation of microphone array for shipboard teleconferencing[J]. Application of Electronic Technique, 2022, 48(3): 32–36, 77.

High precision DOA estimation of microphone array for shipboard teleconferencing

Liu Yuji^{1,2,3}, Tong Feng^{1,2,3}, Chen Dongsheng^{1,2,3}, Lu Rongfu⁴, Feng Wanjian⁴

(1.Key Laboratory of Underwater Acoustic Communication and Marine Information Technique of the Ministry of Education,
Xiamen University, Xiamen 361002, China;
2.College of Earth and Ocean Sciences, Xiamen University, Xiamen 361002, China;
3.Shenzhen Research Institute of Xiamen University, Shenzhen 518000, China;
4.Xiamen Yilian Network Technology Co., Ltd., Xiamen 361000, China)

Abstract: With the improvement of ship intelligence level, shipboard teleconferencing system is of great significance to improve the emergency handling capacity and promote the construction of shipboard integrated network. Microphone array is an important voice front-end to ensure the voice effect as well as the multi-mode interaction of teleconferencing system. However, while the small size of ship cabins leads to the adoption of small-size array, strong reverberation caused by small cabins and noisy cabin noise also seriously degrade the performance of traditional microphone array algorithm. Considering the direction of arrival(DOA) estimation scenario of small-size array in complex environment of ship cabin, a lightweight Mask–DOA estimation neural network model is proposed in this paper. With this method, Mask algorithm is introduced into the DOA estimation neural network to reduce the noise and reverberation interference, then the enhanced GCC–PHAT is extracted as the network feature, so as to realize the high-precision DOA estimation on the small-size microphone array. Simulation and experimental results show that the Mask–DOA model proposed in this paper is more robust and has better generalization ability in the complex environment of ship cabin.

Key words: direction of arrival estimation; ship cabin noise and reverberation environment; neural network; time-frequency masking

0 引言

船载远程会议系统在船舶智能化方面发挥着显著作用,特别是可提高应急处理能力,推进船岸一体化网

* 基金项目:国家自然科学基金(11274259);深圳虚拟大学园扶持经费研发机构建设项目(YFJGJS1.0)

络建设。近些年来,船载远程会议监测系统发展迅速^[1-3]。麦克风阵列通过提供准确波达方向(Direction Of Arrival, DOA)估计可实现语音增强处理,同时还可以为远程会议系统摄像机提供说话人方位信息,实现多模态交互,已成为远程会议系统的重要语音前端^[4-5]。

一般远程会议场所较为理想,因此往往采用较大的麦克风阵列以保证DOA估计,提高语音增强性能和多模态交互效果。但是,船载远程会议所在船舶舱室属于非常典型复杂声学场景。一方面,舱室尺寸狭小,既造成严重混响,也导致无法方便容纳尺寸较大的远程会议麦克风阵列;另一方面,受严重船舶舱室噪声干扰^[6],包括由各个舱室有限的空间里集中了非常多的电气设备与发动机等设备造成严重的内部噪声,以及其他舰船噪声、海浪等导致的外部噪声。这些都将使得船舶舱室声学特性变得复杂,对麦阵DOA估计提出了更高的挑战。

近些年,随着人工智能的发展,Xiao等人提出利用多层次感知机(Multilayer Perceptron,MLP)来进行DOA估计^[7],利用深层网络与大数据来提高DOA估计准确率,远远超过传统DOA估计算法。Diaz-Guerra等人利用带相位变换导向响应功率特征作为特征,建立神经网络将DOA估计任务转化为回归问题^[8]。Nguyen等人使用具有多任务学习功能的2D卷积神经网络从短时空间伪谱鲁棒地估计声源的数量和到达方法^[9],这种方法减少了神经网络学习声音类别和方向信息之间不必要的关联,加速模型的收敛。

Varzandeh等人使用GCC-PHAT和听觉启发特征周期度(Periodicity Degree,PD)作为特征,使用多输入单输出的卷积神经网络(CNN)进行训练,利用PD功能,CNN能够进行语音活动检测(Voice Activity Detection,VAD),从而大大降低混响与噪声对DOA估计的影响^[10]。这些智能算法更多地利用了复杂网络的强大非线性能力与输入特征信息的冗余来提高DOA估计的准确率,且对于阵列的尺寸也没有严格的限制。因此,通常特征维度大,网络结构复杂,计算量大,虽然DOA效果好,但受限于计算量无法真正在移动端实现。

基于上述考虑,本文提出一种轻量级Mask-DOA估计神经网络模型,该方法在DOA估计神经网络的基础上引入Mask算法降低噪声和混响的干扰,提取增强后的GCC-PHAT作为网络特征,从而在小尺寸阵列上实现高精度DOA估计。该方法不仅对环境的鲁棒性更高,且计算复杂度较低,适合在船舶舱室内实用。轻量级Mask-DOA估计神经网络模型分为3个步骤,如图1所示,分别是Mask网络降噪、DOA特征提取以及DOA估计网络。



图1 轻量级 Mask-DOA 估计神经网络

1 轻量级 Mask-DOA 估计神经网络模型

1.1 浅层 Mask-NN

理想二值掩蔽(Ideal Binary Mask, IBM)^[11-12]可以充分利用语音在时频域中的稀疏性。对于每一个时频单元,

语音与噪声能量差异性一般很大,即大部分的时频单元上信噪比是极大或极小的。IBM将所有的时频点分为两种极端情况,即将连续的时频单元信噪比离散化为1和0两种状态:若该时频单元信噪比大(语音占主导),则标记为1;若信噪比低(噪声占主导),则标记为0。如此,计算每一个时频单元得到Mask谱。通过式(1)得到降噪信号的时频域。

每一个时频单元上信噪比计算公式为:

$$\text{SNR}(t,f)=10 \times \log_{10} \left(\frac{|S(t,f)|^2}{|N(t,f)|^2} \right) \quad (1)$$

其中,S(t,f)为信号,N(t,f)为噪声。

Mask谱是通过每一个时频单元的信噪比与所设置的阈值决定的,公式如下:

$$\text{Mask}_{\text{IBM}}(t,f)= \begin{cases} 1 & \text{SNR}(t,f)>\text{LC} \\ 0 & \text{其他} \end{cases} \quad (2)$$

其中,LC是信噪比判断阈值,通常情况下取0。

Mask-NN降噪部分目标是以IBM准则来训练神经网络使得其拥有生成Mask谱的能力。由于语音具有短时相关性,通常采用扩帧技术来使神经网络学习语音帧之间的相关性,一定程度上也有抑制混响的作用。取阵列接收信号时频域的第*i*帧和前后*k*帧作为神经网络的输入特征,取求到的Mask谱第*i*帧作为训练目标。由于采用了扩帧处理,最终降噪后信号帧数会减少2*k*。若对信号短时傅里叶变换的点数为NFFT,则Mask-NN输入特征大小为(2*k*+1,NFFT+1),其输出大小为(1,NFFT+1)。

TF-Mask的方法将语音降噪问题转化为每一个时频点的二分类问题,因此搭建的Mask-NN实际是一个多目标二分类的模型。神经网络隐藏层的层数、每层的个数应该根据NFFT值、扩帧的*k*值而选择。隐藏层每一层激活函数推荐使用ReLU:

$$y_{\text{ReLU}}=\max(x,0) \quad (3)$$

其中,*x*是输入。由于ReLU计算简单,收敛速度快,有效避免梯度消失,对于小于0的值直接清0,大于0的值直接输出本身,与IBM形成Mask的方法是类似的。

Mask-NN的输出层的神经元个数与空间分类的个数相同,激活函数采用sigmoid函数:

$$y_{\text{sigmoid}}=\frac{1}{1+e^{-x}} \quad (4)$$

1.2 GCC-PHAT 特征提取

广义互相关函数时延估计算法根据两个麦克风信号的互相关函数峰值来估计时延值。在声源定位系统中,麦克风阵列的每个阵元接收到的目标信号都来自于同一个声源。因此,各通道信号之间具有较强的相关性。理想情况下,通过计算每两路信号之间的相关函数,就可以确定两个麦克风观测信号之间的时延。

互相关函数可以表示为:

$$R_{x_1x_2}(\tau)=E(x_1(t)x_2(t-\tau)) \quad (5)$$

式中,*x*₁和*x*₂分别是两个通道接收信号。根据维纳辛钦

定理可得：

$$R_{x_1 x_2}(\tau) = \int X_1(\omega) X_2^*(\omega) e^{-j\omega\tau} d\omega \quad (6)$$

其中, $X(\omega)$ 是频域形式, $R_{x_1 x_2}(\tau)$ 波峰直接影响着时延估计的准确性。为了锐化 $R_{x_1 x_2}(\tau)$ 的峰值, 可以根据信号和噪声的先验知识, 在频域内对互功率谱进行加权, 从而能抑制噪声和混响干扰。最后进行傅里叶逆变换, 得到广义互相关函数 $R_{x_1 x_2}(\tau)$:

$$R_{x_1 x_2}(\tau) = \int \varphi_{12}(\omega) X_1(\omega) X_2^*(\omega) e^{-j\omega\tau} d\omega \quad (7)$$

其中, $\varphi_{12}(\omega)$ 表示频域加权函数。本文采用的是 PHAT 加权, 其形式如下:

$$\varphi_{\text{PHAT}}(\omega) = \frac{1}{|X_1(\omega) X_2^*(\omega)|} \quad (8)$$

相位变换加权函数实质上是一个白化滤波器, 使得信号间的互功率谱更加平滑, 从而锐化广义互相关函数。经过 PHAT 加权的互功率谱近似于单位冲激响应, 突出了时延的峰值, 能够有效抑制混响噪声, 提高时延估计的精度和准确度。特别是针对舰船噪声, 通过 PHAT 加权的 GCC 特征抗噪抗混响能力更强。

1.3 浅层 DOA-NN

使用神经网络来进行 DOA 估计其实就是将空间进行等间隔划分成 N 份, 将 DOA 估计问题转换成一个 N 任务的二分类问题。因此, DOA 估计的精度与划分的空间数量直接相关, 将空间划分越细, 精度越高, 但分类的难度也就越大, 对 DOA 估计的准确性提出了挑战; 反之, 精度越低, 但是 DOA 估计的准确率越高。

由于本文提取的 GCC-PHAT 特征在不同入射角度阵列接收信号上体现出较大的差异性, 这是使用神经网络模型进行 DOA 估计的关键所在。由于前期有 Mask 网络进行降噪处理, 以及 GCC-PHAT 本身对混响噪声抑制性, 使得利用浅层神经网络实现 DOA 估计成为可能, 仅需少量的非线性神经元即可实现特征到角度的映射关系。在 DOA-NN 网络结构中推荐使用单隐藏层的全连接神经网络。唯一需要注意的是 DOA-NN 的输出层神经元的数量应该等于空间分割数, 输出层的激活函数应该使用 sigmoid 函数。

2 仿真和分析

2.1 仿真环境设置

本文采用 IMAGE 算法对船舶舱室声学特性进行建模, 为了更加真实地模拟船舶舱室, 仿真中建立一个空间大小合适的舱室模型, 规格为 $6.5 \text{ m} \times 6.5 \text{ m} \times 2.7 \text{ m}$ 。由于船舶舱室内的布局是无法确定的, 仿真在同一房间内通过调整墙壁的反射系数来得到不同混响程度以适配更广泛多样的真实多样的舱室模型。本文设置了混响时间分别为 $0.2 \text{ s}, 0.4 \text{ s}, 0.6 \text{ s}, 0.8 \text{ s}$ 的 4 种不同混响环境, 噪声比分别为 $10 \text{ dB}, 5 \text{ dB}, 0 \text{ dB}, -5 \text{ dB}$ 的 4 种信噪比环境。

在仿真中设置的 DOA 估计的精度为 5° , 因此需要将空间划分为 72 等份。将一个直径为 7 cm 的 6 麦均匀圆阵置于房间中间, 周围等间隔 72 个角度放置扬声器, 扬声器到麦克风阵列圆心距离为 $0 \sim 3 \text{ m}$ 均匀分布, 来模拟不同角度人声的入射。语音采用 THCHS-30 数据集, 添加的噪声信号采用实测的舰船各种噪声, 包括船舶沉闷轰鸣声、船舶发电机引擎噪声、船舶鸣喇叭声。每一种不同混响不同信噪比的船舶舱室中对每一个角度不同距离进行 30 次入射实验。每一个不同混响的船舶舱室内样本数为 8 640 条, 最终样本总数为 34 560 条。

2.2 仿真设计与结果

基于上文对轻量级 Mask-DOA 估计神经网络模型的描述, 在仿真中实例化该网络, 网络结构以及具体流程如图 2 所示。

面对船舶舱室复杂的声学环境, 本文使用麦克风阵列 DOA 估计的常规波束形成法(CBF)作为基线, 主要对比目前主流的 MLP 算法^[2]。DOA 估计效果用准确率和均方误差(RMSE)来衡量, 公式如式(9)和式(10)所示:

$$\text{Acc} = \frac{L_{\text{acc}}}{L_{\text{all}}} \times 100\% \quad (9)$$

$$\text{RMSE} = \sqrt{\frac{1}{L_{\text{all}}} \sum_{i=0}^{L-1} (\theta_{\text{predict}}^i - \theta_{\text{true}}^i)^2} \quad (10)$$

式中, L_{acc} 为正确的样本数, L_{all} 为所有样本数, θ_{true}^i 表示第 i 个实验正确方向, $\theta_{\text{predict}}^i$ 表示第 i 个实验预测方向。表 1 是 CBF 算法在不同混响等级下多个信噪比的平均结果。

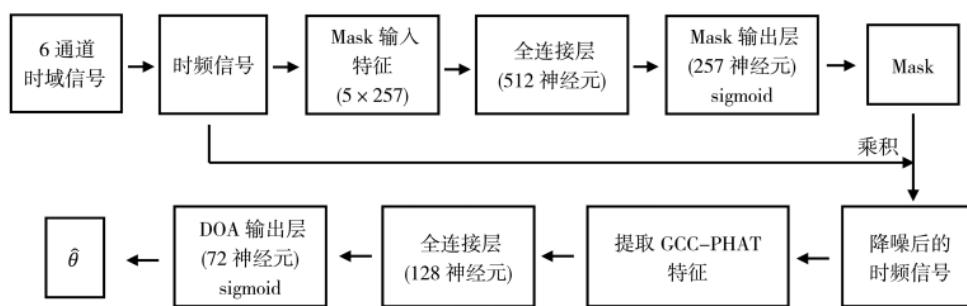


图 2 Mask-DOA 具体网络结构与流程图

表 1 CBF 作为基线的船舶舱室内 DOA 估计效果

RT60/s	准确率/%	均方误差(RMSE)/(°)
0.2	69.14	63.19
0.4	48.09	69.31
0.6	34.83	83.21
0.8	23.31	97.29

为了对比本文算法与 MLP 算法的准确率、均方误差以及鲁棒性,本文采用了相同的数据集进行对照实验。采用单个混响程度船舶舱室的 80% 进行模型的训练,剩余 20% 的数据进行该舱室的测试,用其余各个房间所有的数据对训练好的模型进行测试,准确率对比结果如图 3 和图 4 所示,均方误差的对比结果如图 5 和图 6 所示。

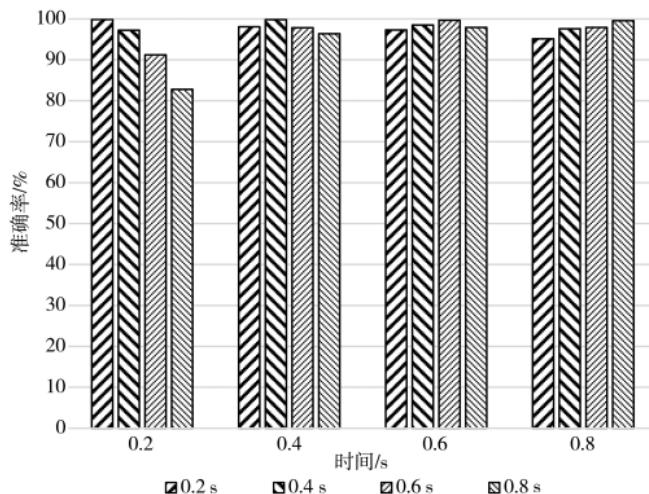


图 3 Mask-DOA 训练模型准确率

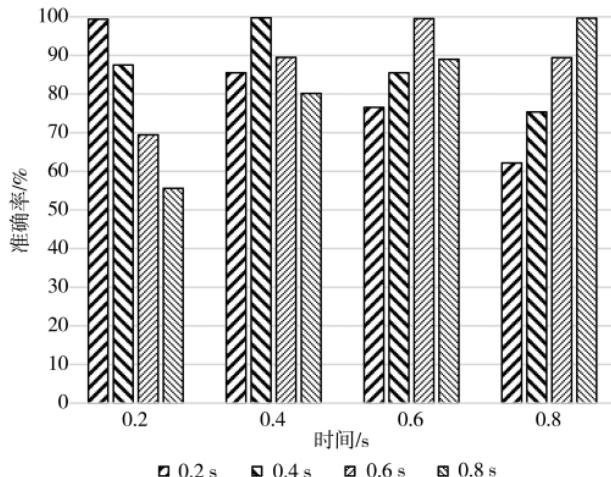


图 4 MLP 训练模型准确率

对比图 3 和图 4 以及图 5 和图 6,无论是从准确率还是均方误差的角度来看,本文提出的 Mask-DOA 算法的效果要明显优于 MLP 算法,更是远远超过基线算法。具体表现在,Mask-DOA 算法训练的模型无论在环境是否与训练集匹配的条件下都有较高的 DOA 估计效果,

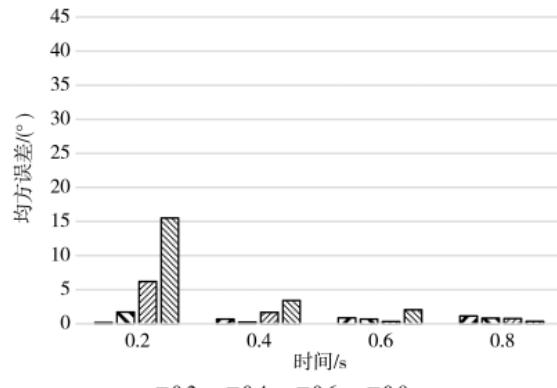


图 5 Mask-DOA 训练模型均方误差

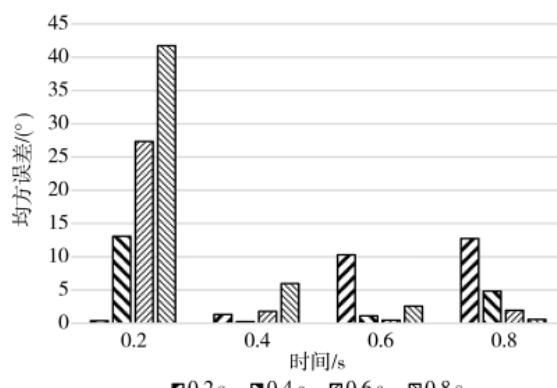


图 6 MLP 训练模型均方误差

即模型的泛化能力强。而 MLP 算法在环境匹配的条件下可以达到较高的准确率,但是一旦环境失配,其准确率下降明显,模型泛化能力较差,抗混响能力较差。

3 实验和分析

3.1 实验环境设置

实验场地选择某公司的可调混响室作为舰船模拟的场景,其混响室尺寸为 $6.5 \text{ m} \times 5.7 \text{ m} \times 2.7 \text{ m}$,其混响可调范围为 0.2~0.6 s。将一个直径为 7 cm 的六麦圆形阵列放置在旋转台上,将一个扬声器与麦阵等高放置,播放 THCHS-30 语音,将其他多个扬声器放置在房间的任意位置并播放舰船噪声来模拟船舶环境。通过不断地调节室内的混响环境,每一个方向不同距离进行 10 次,对 72 个方向进行实验。

3.2 实验结果与分析

挑选 0.2 s 混响调节下采集到的数据作为训练集,对 0.4 s 采集到的数据进行测试。分别采用所提 Mask-DOA 方法和 MLP 算法与 CBF 算法进行对比实验,结果如表 2 所示。通过表 2 中 3 种算法的对比,传统的 CBF 算法在小信噪比、混响复杂的情况下效果非常差,因此在真实的船舶舱室调节下效果只会更差。Mask-DOA 算法的效果要优于 MLP 算法,体现出其较强的泛化能力,比较适用于低信噪比、声学特性复杂的舰船舱室场景。

表 2 3 种算法实验结果对比

算法	准确率/%	均方误差/(°)
CBF	53.45	68.8
MLP	73.53	15.43
Mask-DOA	93.32	5.89

4 结论

面向船载远程会议系统语音前端需求,本文提出一种适用于船舶舱室环境的小阵列高精度 DOA 估计算法,该方法在 DOA 估计神经网络的基础上引入 Mask 算法降低噪声和混响干扰,从而用较为纯净的语音信号提高 DOA 估计性能。在直径为 7 cm 的六元圆阵上进行的仿真与实验表明,DOA 估计精度达到 5°,且算法模型有着较强的泛化能力,证明了本文算法在船舶舱室内使用的有效性。

同时,需要指出:本文采用深度学习方式构建了信号预处理与 DOA 估计模型系统,虽通过模拟舱室环境实验验证了性能改善,但是考虑到深度学习类算法的效

果与鲁棒性受到数据集大小与数据类型适配的影响,算法的评估验证尚不够充分。特别是,由于船舶舱室属于带行业背景的特殊类型场景,噪声及混响特性复杂、差异性大,无法直接利用通用语音领域的丰富开源数据集、混响特性资源,本文方法训练出的模型不可避免存在数据失配条件下性能下降的不足。

因此,面向智能船舶、智慧海洋发展的大趋势,下一步计划结合实际场景下多样化的船舶舱室特性、背景噪声进一步开展所提方法的性能评估、验证。

参考文献

- [1] 长航局.长江航道船载视频会议及远程视频监控系统通过验收[EB/OL].(2011-08-12)[2021-08-30].<http://news.hexun.com/2011-08-12/132386171.html>.
- [2] 苑立娟,秦芹.云平台船舶状态实时监控系统[J].舰船科学技术,2021,43(10):178-180.
- [3] 王栋.科考船船岸一体化网络设计及发展趋势[J].船舶,(下转第 77 页)

(上接第 31 页)

- ence, 1996, 18(10): 959-971.
- [9] VIOLA P A, JONES M J. Rapid object detection using a boosted cascade of simple features[C]//IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2001: 511-518.
- [10] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//2014 IEEE Conference on Computer Vision and Pattern Recognition, 2014: 580-587.
- [11] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.
- [12] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2016: 779-788.
- [13] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector[C]//European Conference on Computer Vision. Berlin: Springer, 2016: 21-37.
- [14] MOGELMOSE A, TRIVEDI M M, MOESLUND T B. Vision-based traffic sign detection and analysis for intelligent driver assistance systems: Perspectives and survey[J]. IEEE Transactions on Intelligent Transportation Systems, 2012, 13(4): 1484-1497.
- [15] Houben S, Stal Kamp J, Salmen J, et al. Detection of traffic signs in real-world images: the german traffic sign detection benchmark[C]//The 2013 International Joint

Conference on Neural Networks(IJCNN). IEEE, 2013: 1-8.

- [16] TIMOFTE R, ZIMMERMANN K, VAN GOOL L. Multi-view traffic sign detection, recognition, and 3D localisation[J]. Machine Vision and Applications, 2014, 25(3): 633-647.
- [17] ZHU Z, LIANG D, ZHANG S, et al. Traffic-sign detection and classification in the wild[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 2110-2118.
- [18] RAJENDRAN S P, SHINE L, PRADEEP R, et al. Fast and accurate traffic sign recognition for self driving cars using RetinaNet based detector[C]//2019 International Conference on Communication and Electronics Systems. Piscataway: IEEE Press, 2019: 784-790.
- [19] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2016: 770-778.
- [20] AGHDAM H H, HERAVI E J, PUIG D. A practical approach for detection and classification of traffic signs using convolutional neural networks[J]. Robotics and Autonomous Systems, 2016, 84: 97-112.

(收稿日期:2021-06-07)

作者简介:

张成(1990-),男,硕士,助教,主要研究方向:人工智能技术、自动化驾驶。

张瑞宾(1984-),男,博士,副教授,主要研究方向:人工智能技术、自动化驾驶。

王曙道(1973-),男,硕士,高级工程师,主要研究方向:汽车结构分析。



扫码下载电子文档

略。同时根据实际业务运营的成效，对客户群体的模拟运营进行干预和调优。通过这种模拟运营指导实际执行、实际成效优化模拟运营方案的方式，可极大降低风险，提升整体运营服务质量。

5 结论

本文主要论述了在数据中台理念下的智能标签建设和应用方向，提出了一套较为完善的智能标签构建方法，为高效构建智能标签及应用提供了方向。在数字化转型的大趋势下，下一步应深入探究标签与具体业务融合，在具体业务场景下提高标签落地应用效果。

参考文献

- [1] 林森,欧阳柳.基于大数据理论的电力客户标签体系构建[J].电气技术,2016,17(12):98–101.
- [2] 韩向东,季献忠.企业数字化转型八大发展趋势[J].管理会计研究,2021,4(Z1):21–26,101.
- [3] 崔恒志,王翀,吴健.基于数据中台的数据资产管理体系[J].计算机系统应用,2021,30(3):33–42.
- [4] 姚洪.基于数据中台的数据治理系统的设计与实现[J].科学技术创新,2020(35):74–75.
- [5] 吕辉,许道强,仲春林,等.基于电力大数据的标签画像技术与应用研究[J].电力信息与通信技术,2017,15(2):43–48.
- [6] 张娜,马燕,崔桓睿,等.大数据技术进展与发展趋势[J].通讯世界,2019,26(2):78–79.
- [7] 郑钤.基于 MapReduce 模式的大数据分布式计算态势分析[J].通讯世界,2018(6):102–104.
- [8] 阿里巴巴数据技术及产品部.大数据之路:阿里巴巴大数据实践[M].北京:电子工业出版社,2017.
- [9] 钟华.企业 IT 架构转型之道:阿里巴巴中台战略思想与架构实战[M].北京:机械工业出版社,2017.
- [10] 邓中华.大数据大创新—阿里巴巴云上数据中台之道[M].北京:电子工业出版社,2018.
- [11] 施巍巍.江苏联通对大数据精准营销模型的探讨[J].信息系统工程,2015(1):112–114.
- [12] 印斯佳.电力客户差异化服务探析[J].城市建设理论研究,2011(34):2095–2104.
- [13] 宋亮.用智能标签发现未来[J].标签技术,2020(3):16–18.
- [14] 李海峰,王炜.数字孪生智慧学习空间:内涵、模型及策略[J].现代远程教育研究,2021,33(3):73–80,90.
- [15] 侯欣洁,戚德祥.数字技术驱动下的出版“走出去”数字平台建设[J].科技与出版,2021(5):60–65.

(收稿日期:2021-02-05)

作者简介:

盛妍(1967-),女,硕士,高级工程师,主要研究方向:数据挖掘分析。

朱青(1990-),通信作者,女,硕士,中级工程师,主要研究方向:数据分析及应用,E-mail:zhuqing2020@foxmail.com。



扫码下载电子文档

(上接第 36 页)

2017,28(S1):234–241.

- [4] 舒雪姣,陈景发,敬代波.语音 DOA 技术在视频会议系统中的应用[J].电子技术与软件工程,2014(22):138–139.
- [5] 李伟.智能环境下基于音频视频特征融合的多说话人跟踪研究[D].兰州:兰州理工大学,2012.
- [6] 李鹏,高小平,陈梅清,等.主动降噪技术应用于船舶舱室的可行性分析[J].舰船科学技术,2020,42(3):69–71.
- [7] 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 (ICASSP).South Brisbane:IEEE,2015:2814–2818.
- [8] DIAZ-GUERRA D,BELTRAN J R.Direction of arrival estimation with microphone arrays using SRP-PHAT and neural networks[C]//2018 IEEE 10th Sensor Array and Multichannel Signal Processing Workshop(SAM),Sheffield,UK,2018:617–621.
- [9] NGUYEN T N T,GAN W S,RANJAN R,et al.Robust source counting and DOA estimation using spatial pseudo-spectrum and convolutional neural network[J].IEEE/ACM Transactions on Audio,Speech ,and Language Processing,

2020,28:2626–2637.

- [10] VARZANDEH R,ADILOGLU K,DOCLO S,et al.Exploiting periodicity features for joint detection and DOA estimation of speech sources using convolutional neural networks[C]//ICASSP 2020 – 2020 IEEE International Conference on Acoustics,Speech and Signal Processing(ICASSP),Barcelona,Spain,2020:566–570.
- [11] WANG Y,WANG D L.Towards scaling up classification-based speech separation[J].IEEE Transactions on Audio,Speech ,and Language Processing,2013,21(7):1381–1390.
- [12] SRINIVASAN S,ROMAN N,WANG D.Binary and ratio time-frequency masks for robust speech recognition[J].Speech Communication,2006,48(11):1486–1501.

(收稿日期:2021-08-30)

作者简介:

刘雨佶(1997-),男,硕士研究生,主要研究方向:麦克风阵列、语音增强。

童峰(1973-),通信作者,男,博士,教授,主要研究方向:水声通信与网络、声探测与感知、智能语音处理,E-mail:ftong@xmu.edu.cn。

陈东升(1975-),男,硕士,助理教授,主要研究方向:水声通信、声信号处理。



扫码下载电子文档

版权声明

经作者授权，本论文版权和信息网络传播权归属于《电子技术应用》杂志，凡未经本刊书面同意任何机构、组织和个人不得擅自复印、汇编、翻译和进行信息网络传播。未经本刊书面同意，禁止一切互联网论文资源平台非法上传、收录本论文。

截至目前，本论文已经授权被中国期刊全文数据库（CNKI）、万方数据知识服务平台、中文科技期刊数据库（维普网）、DOAJ、美国《乌利希期刊指南》、JST 日本科技技术振兴机构数据库等数据库全文收录。

对于违反上述禁止行为并违法使用本论文的机构、组织和个人，本刊将采取一切必要法律行动来维护正当权益。

特此声明！

《电子技术应用》编辑部

中国电子信息产业集团有限公司第六研究所