ISSN 1004-9037,CODEN SCYCE4
Journal of Data Acquisition and Processing Vol. 39,No. 5,Sep. 2024,pp. 1043－1043
DOI:10. 16337/j. 1004-9037. 2024. 05. 001
http://sjcj. nuaa. edu. cn
E-mail:sjcj@nuaa. edu. cn
Tel/Fax:＋86-025-84892742

“语音/音频信号智能处理”专栏序言

近年来,深度学习和大模型的快速发展极大地推动了语音/音频信号智能化处理技术的进步。以语音/音频信号智能处理为代表的新一代智能信息技术,逐渐成为我国推动经济高质量发展、建设创新型国家的重要技术保障和核心驱动力。语音识别、声纹识别、语音合成、声源定位、情感分析、噪声消除等语音/音频信号智能处理技术已在智能助手、智能家居、人机交互、自动驾驶、医疗诊断等众多领域得到了广泛应用。随着生成式人工智能的进一步发展,语音/音频信号智能处理技术将会在更多领域发挥重要作用。

语音/音频信号智能处理的关键在于对信号规律进行建模并实现智能化处理。相比于早期使用的以HMM模型为代表的统计建模技术,基于大模型的深度学习技术通过端到端的学习方式,直接从原始音频数据中提取特征,避免了繁琐的特征工程过程。这种技术的进步使得语音识别、说话人识别和语音合成等应用的性能得到显著提升,大大提升了用户的体验感,极大地推动了相关产业的发展。

随着以大模型为基础的新一代人工智能技术的发展,基于大模型的语音/音频信号智能处理引起了国际学术界和产业界的广泛关注。以ChatGPT为代表的大规模语言模型在自然语言生成和理解方面表现出色,这些模型同样在语音处理领域展现了巨大的潜力。2024年5月,美国OpenAI公司发布的GPT-4o大模型,为ChatGPT注入了更加强大的语音信号处理能力和快速的响应能力,使其能够在短至232 ms时间内对音频输入做出反应,引起了整个语音/音频信号智能处理领域的广泛关注。此外,基于深度学习的多模态音频处理技术也被用于声源信息和视频场景分析,可以实现更高质量的音频体验。语音/音频信号智能处理技术在高质量音频采集、声场重建和智能语音交互等场景中的应用也取得了显著成果。

为推动语音/音频信号智能处理领域高质量发展,《数据采集与处理》2024年第5期特别策划“语音/音频信号智能处理”专栏,集中刊发北京工业大学、哈尔滨工业大学、中国科学技术大学、南京大学、南京信息工程大学、中国科学院等单位知名学者团队的6篇高质量论文。专栏论文聚焦语音/音频信号智能处理中的关键技术问题,系统总结了当前研究现状,深入分析了未来发展趋势,结合深度学习方法优化了典型应用,展示了本领域研究的深度和应用的广度。期待本期专栏论文能为语音和音频信号处理领域研究人员开展核心技术研究提供有益启发和参考。

专栏责任编委

专栏责任编委、副主编

张雄伟,陆军工程大学教授、博士生导师。全国优秀教师、国务院政府特殊津贴获得者、百千万人才工程国家级人选。兼任《数据采集与处理》《陆军工程大学学报》编委。主要从事语音信号处理领域的研究,具体研究方向包括:语音编码、语音增强、说话人识别、语音转换、语音鉴伪、语音隐藏等。获国家科技进步二等奖1项、省部级科技进步一等奖3项,国家级教学成果二等奖1项、省部级教学成果一等奖2项,发表学术论文100余篇,出版《智能语音处理》等著作教材10余部,其中《DSP芯片的原理与开发应用》被评为高等教育“十一五”国家级规划教材。

刘晓峻,南京大学教授、博士生导师,国务院政府特殊津贴获得者,首批教育部新世纪优秀人才。先后担任南京大学电子科学与工程学院副院长、近代声学教育部重点实验室主任、南京大学声学研究所所长;先后兼任中国声学学会常务理事、副秘书长,江苏省声学学会理事长,《数据采集与处理》《应用声学》副主编和《声学学报》等刊物编委。主要从事声学人工材料及新原理声学功能器件、光声检测及声信号处理等方面的研究,在Nature等国际重要学术刊物上发表论文400余篇,他引10 000余次,多篇论文入选ESI热点论文和高被引论文。曾获教育部自然科学奖、江苏省科学技术进步奖及中国声学学会魏荣爵奖等。