Technical Program

会议程序册

2025 IEEE International Symposium on Machine Learning and Media Computing

2025 年 IEEE 机器学习与媒体计算国际研讨会

25-28 July, 2025

2025 年 7 月 25-28 日

Wanda Vista Hotel, 87 Shimao Blvd, Songbei District, Harbin, Heilongjiang, China

黑龙江省哈尔滨市松北区世茂大道 87 号万达文华酒店

Technically Sponsored By (技术支持)

Institute of Electrical and Electronics Engineers

(国际电气电子工程师协会)

IEEE Systems, Man and Cybernetics Society

(IEEE 系统、人、控制论学会)

Organized By (主办单位)

Lingnan University, Hong Kong SAR, China (岭南大学)

Shenzhen University, China (深圳大学)

Harbin Institute of Technology, China (哈尔滨工业大学)

Ningbo University, China (宁波大学)

Sun Yat-sen University, China (中山大学)

Southeast University, China (东南大学)

Tongji University, China (同济大学)

Fuzhou University, China (福州大学)

Nankai University, China (南开大学)

Tianjin University, China (天津大学)

Local Organizer (本地单位)

Harbin Institute of Technology, China (哈尔滨工业大学)

July 25, 2025

12:00-21:00	Registration
19:00-21:00	VIP Reception

July 26, 2025

8:30-9:10	Opening Ceremony @多功能 2 厅	
9:10-10:00	Keynote Speech A @多功能 2 厅 Title: E-CARGO/RBC: Enabling Research Innovations in the Era of AI Speaker: Prof. Haibin Zhu (朱海滨 教授) Host: Yuheng Jia (贾育衡)	
10:00-10:20	Coffee / Tea Break	
10:20-11:10	Keynote Speech B @多功能 2 厅 Title: 深伪鉴伪与网络认知安全 Speaker: Prof. Xiangyang Luo (罗向阳 教授) Host: Guopu Zhu (朱国普)	
11:10-12:00	Keynote Speech C @多功能 2 厅 Title: 复杂三维模型的处理、理解与生成技术及应用 Speaker: Prof. Xiaopeng Fan (范晓鹏 教授) Host: Long Xu (徐龙)	
12:00-14:00	Lunch Buffet @美食汇西餐厅	
14:00-15:40	Oral Session 1 @多功能 3 厅	Oral Session 2 @董事会厅
	Poster Session @多功能 2 厅外侧	Poster Session @多功能 2 厅外侧
15:40-16:00	Coffee / Tea Break	
16:00-17:40	Oral Session 3 @多功能 3 厅	Oral Session 4 @董事会厅
	Poster Session @多功能 2 厅外侧	Poster Session @多功能 2 厅外侧
18:30-21:00	Banquet @多功能 2 厅	

July 27, 2025

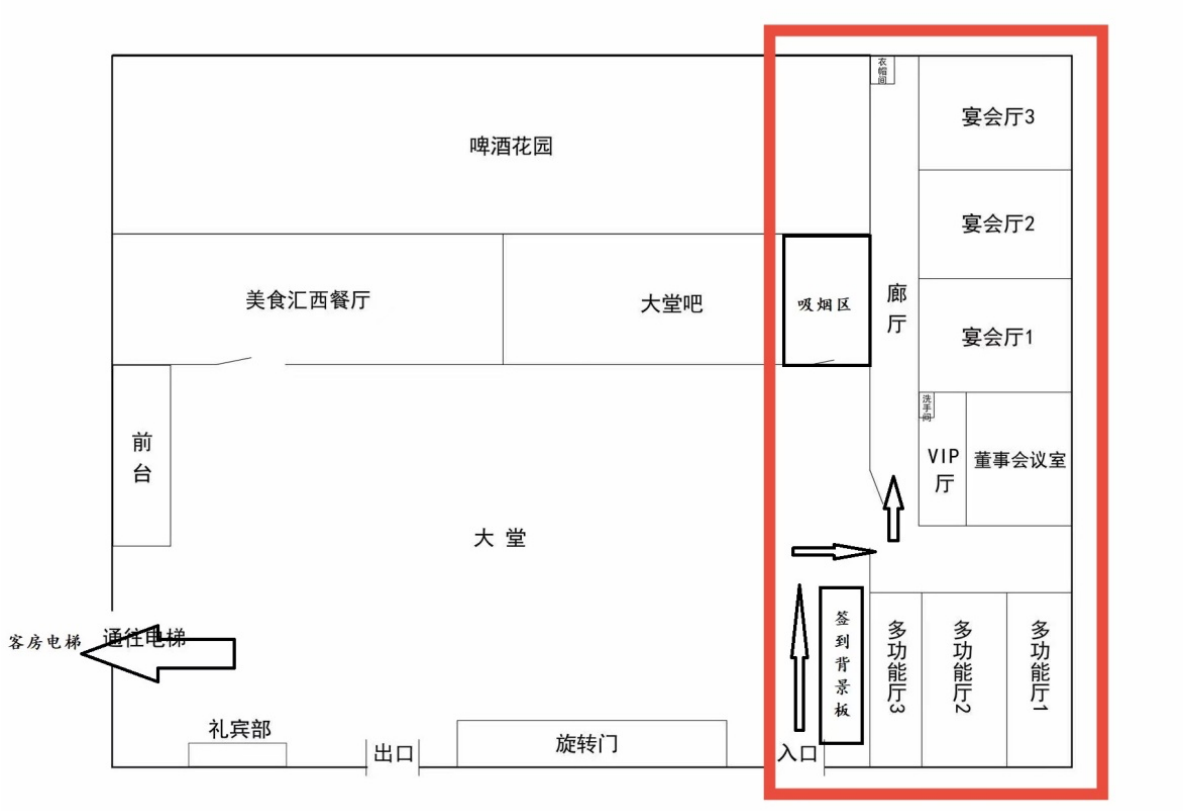
8:30-9:20	Keynote Speech D @多功能 2 厅 Title: Generative Visual Data Compression and Quality Assessment Speaker: Prof. Shiqi Wang (王诗淇 教授) Host: Yun Zhang (张云)

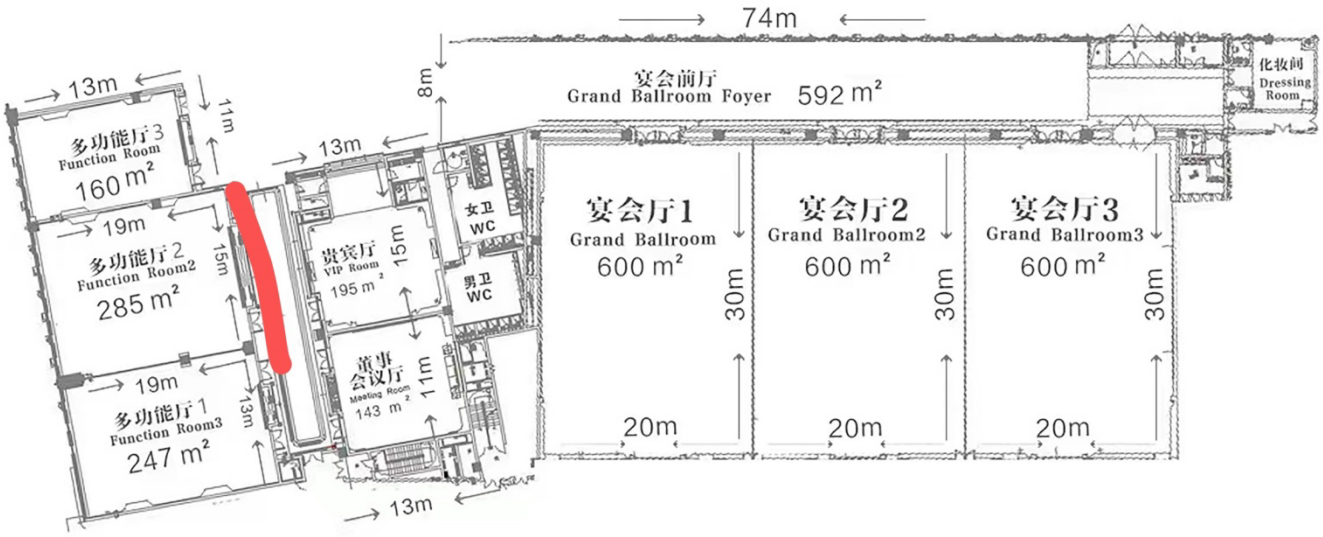
9:20-10:10	Keynote Speech E @多功能 2 厅 Title: Implicit Neural Representation for Video Coding Speaker: Prof. Xinfeng Zhang (张新峰 副教授) Host: Wenhui Wu (邬文慧)
10:10-10:30	Coffee / Tea Break
10:30-12:00	Special Discussion @多功能 2 厅 Topic: Will AI Develop Consciousness? Host: Yun Zhang / Ran Wang (张云/王冉)
12:00-14:00	Lunch Buffet @美食汇西餐厅
14:00-17:00	Group Discussion (闭门会议)

July 28, 2025

9:00-12:00	Closing Ceremony @多功能 2 厅 Session 1: MLMC 2025 Awarding Ceremony Session 2: Summary Report for MLMC 2025 Session 3: Plan Discussion for MLMC 2026

Floor Plan





Detail of Sessions

Oral Session 1 @多功能 3 厅（14:00-15:40, July 26, 2025）

Machine Learning / Session Chair: Yu Tian (田钰)

14:00-14:25	ID 10: Partial Label Learning Tailored Graph Construction Authors: Fuchao Yang, Yongqiang Dong, Yuheng Jia
14:25-14:50	ID 38: Long-Tailed Multi-Label Learning: A Benchmark of Evaluation Metrics Authors: Baoxuan Wang, Jiayi Lu, Xinlei Zhou, Yuxuan Luo, Jun Li, Ran Wang
14:50-15:15	ID 41: Leveraging Pretrained ECG-PPG Model for Continuous Blood Pressure Estimation Authors: Weihao Zhuang, Zengding Liu, Weixiu Qiu, Fen Miao
15:15-15:40	ID 57: Contrastive Swin-Transformer for Fine-Grained Visual Classification Authors: Pingping Zhang, Xizhao Wang, Xinlei Zhou

Oral Session 2 @董事会厅（14:00-15:40, July 26, 2025）

Video Coding / Session Chair: Haifeng Guo (郭海峰)

14:00-14:25	ID 11: Eye Fixation and Saliency Field based Greedy Method for Volumetric Video Streaming Authors: Zhiye Tang, Kai Wang, Qiudan Zhang, Haowen Mo, Yelang Gao, Lei Zhang, Jun Li, Xu Wang

14:25-14:50	ID 16: Scene-aware Graph-enhanced Multimodal Collaboration for Video-grounded Dialogue Authors: Shanshan Du, Hanli Wang
14:50-15:15	ID 21: DM-360VFI: 360° Spherical Video Interpolation Using Diffusion Model Authors: Xing Li, Yang Zhou, Haibing Yin
15:15-15:40	ID 54: Softmax-based Intra Luminance Prediction for H.266/VVC Authors: Tuo Li, Chaofei Li, Wenrui Lv, Shiqi Jiang, Hui Yuan

Oral Session 3 @多功能 3 厅（16:00-17:40, July 26, 2025）

Compression / Session Chair: Xinlei Zhou (周欣蕾)

16:00-16:25	ID 15: Multi-frame Context Generation for Neural Video Compression Authors: Qinhao Huang, Yun Zhang, Junle Liu
16:25-16:50	ID 27: Adaptive Weight Lightweight Image Compression Authors: Lingjie Luo, Zhicheng Zhang
16:50-17:15	ID 32: OR-JRD: Observation Region-wise Just Recognition Distortion Guided Whole Slide Image Compression for Histopathology Classification Authors: Wenlong Shi, Na Li, Xiuli Li, Guanghui Yue, Yun Zhang
17:15-17:40	ID 36: FDIC: An Adaptive End-to-End Compression Framework for Medical lmages Authors: Weilin Yang, Guifu Wu

Oral Session 4 @董事会厅（16:00-17:40, July 26, 2025）

Computer Vision / Session Chair: Chao Zeng (曾超)

16:00-16:25	ID 20: Pseudo-depth Guided Multi-object Tracking Authors: Zijie Zhuang, Xiaokai Yi, Shuaixiong Hui, Jing Yang, Hanli Wang
16:25-16:50	ID 29: RetUIE: Retention-based Underwater Image Enhancement Authors: Meisheng Guan, Haiyong Xu, Yeyao Chen, Ting Luo, Yang Song, Huaping Wang
16:50-17:15	ID 37: Fourier Domain Diffusion Priors for Robust Underwater Image Enhancement Authors: Yuanlin Zhang, Jieyu Yuan, Xiongxin Tang, Xiao Chen, Chunle Guo, Chongyi Li
17:15-17:40	ID 48: Towards Unified Face Verification Against Quality Variations and Deepfake Authors: Fu-Zhao Ou, Haifeng Guo, Sam Kwong

Poster Session @多功能 2 厅外侧（14:00-17:40, July 26, 2025）

Session Chair: Meng Hu (胡猛)

14:00-17:40	ID: 13 A Survey on Compression and Quality Assessment Techniques for 3D Gaussian Splatting Authors: Xinju Wu, Xiangrui Liu, Meng Wang, Shiqi Wang, Sam Kwong
	ID 14: Domain-Specific Question Answering System of Large Language Model Enhanced by Knowledge Graph Authors: Sipu Liu, Bo Chen, Junwei Li, Xiaoli Sun, Jianliang Tang, Weiqiang Zhang
	ID 17: Collaborative Human-machine Perception for Cross-modal Image Compression Authors: Zhisen Tang, Xiaokai Yi, Hanli Wang
	ID 18: Image-specific Bit Allocation Optimization for Multiscale Feature Coding for Machines Authors: Junle Liu, Yun Zhang, Qinhao Huang, Long Xu
	ID 23: Deep Transfer Learning Based Cuff-less Continuous Blood Pressure Measurement for Arrhythmia Patients Authors: Bo Wen, Weixiu Qiu, Zengding Liu, Min Tang, Fen Miao
	ID 24: Referenced Frame Generation Network for Learned Video Compression Authors: Qinhao Huang, Yun Zhang
	ID 26: Deep Feature Compression on Edge Hardware Authors: Xingyu Yan, Jingjing Deng, Long Xu, Zhuo Chen
	ID 31: Gastric Cancer Staging of CT Images Based on Deep Learning Authors: Lida Yin, Liang Chen, Hongrui Chen
	ID 42: Optimization of Steel Coil Stowage on Ships Using a Hybrid Heuristic Algorithm Authors: Chenyang Dong, Jingjing Cao
	ID 45: MAFE-Net: Mixed Attention and Filter Enhancement Network Authors: simin Gao, Hua Li
	ID 47: A Joint Framework for Underwater Image Enhancement and Super-Resolution Reconstruction Authors: Yutong Shi, Yuemiao Wang, Wenhui Wu, Jun Li
	ID 49: TensoNRV: Tensorial Neural Representation for Videos Authors: Yuanjie Cao, Na Li, Yun Zhang
	ID 51: Dense-RVCNN: A Deep Learning-based Inter-frame Luminance Quality

	Enhancement Method Authors: Shiqi Jiang, Yanhan Chu, Wenrui Lv, Chaofei Li, Hui Yuan
	ID 52: Multi-table based Incremental Hashing for Cross Modal Retrieval Authors: Dezhong Zhu, Chen Luo, Qihua Li, Xing Tian, Wing W. Y. Ng
	ID 53: High-Precision Centroid Extraction for Space Targets by GNC and Image Processing Integration Authors: Yue Dou, Meng Xie, Weitao Liu, Feng Chen, Yongqiang Jin, Jing Guo
	ID 56: Contrastive Adversarial Domain Generalization for Face Anti-Spoofing Authors: Siyuan Wang, Yang Chen, Guopu Zhu

Oral Presenter Guidelines

1. In-Person Presentation
   a. Presenters are required to arrive at the Presentation Room 5 minutes before the start of the Session.
   b. A computer will be provided in each Presentation Room. Bring a USB drive with your presentation file.
2. Each paper is allocated a total of **25 minutes**, consisting of a presentation and Q&A session.
3. Contact the host or Chair of the session if there is any problem.

Poster Presenter Guidelines

1. In-Person Presentation
   a. Presenters are required to arrive at the Presentation Room 5 minutes before the start of the Session.
   b. An exhibition board will be provided for each poster paper. Bring your well-printed poster with the maximum size allowed for printing: **0.841m×1.189m (A0) width/height**.
2. A poster paper can be exhibited throughout the session.
3. Contact the host or Chair of the session if there is any problem.

Introduction to Keynote Speech A

Title: E-CARGO/RBC: Enabling Research Innovations in the Era of AI
Speaker: Haibin Zhu, Fellow, IEEE, Professor, Nipissing University, Canada

ABSTRACT

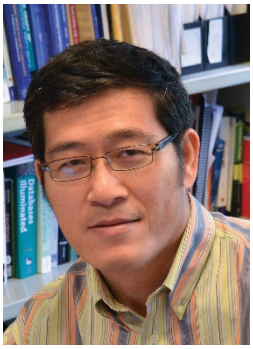
In the AI (Artificial Intelligence) time, many AI tools, such as LLMs (Large Language Models), can help people accomplish many low-level intelligent tasks, such as coding and reporting. Many low-level routine jobs have high potential to be replaced by such LLMs. Traditional programmers need to master powerful high-level modelling tools to meet these new challenges. E-CARGO/RBC (Environments - Classes, Agents, Roles, Groups, and Objects /Role-Based Collaboration) is a modelling methodology, which helps people deal with complex problems by designing systematic strategies other than using low level programming skills.

RBC is a computational methodology that uses roles as the primary underlying mechanism to facilitate collaboration activities. It consists of a set of concepts, principles, models, processes, and algorithms. RBC and its E-CARGO model have been developed to a powerful tool for investigating collaboration and complex systems. Related research has brought and will bring in exciting improvements to the development, evaluation, and management of systems including collaboration, services, clouds, productions, and administration systems. RBC and E-CARGO grow gradually into a strong fundamental methodology and model for exploring solutions to problems of complex systems including Collective Intelligence, Sensor Networking, Scheduling, Smart Cities, Internet of Things, Cyber-Physical Systems, and Social Simulation Systems.

E-CARGO assists scientists and engineering to formalize abstract problems, which originally are taken as complex problems, and finally points out solutions to such problems including programming. The E-CARGO model possesses all the preferred properties of a computational model. It has been verified by formalizing and solving significant problems in collaboration and complex systems, e.g., Group Role Assignment (GRA). With the help of E-CARGO, the methodology of RBC can be applied to solve various real-world problems. E-CARGO itself can be extended to formalize abstract problems as innovative investigations in research. On the other hand, the details of E-CARGO components are still open for renovations for specific fields to make the model easily applied. For example, in programming, we need to specify the primitive elements for each component of E-CARGO. When these primitive elements are well-specified, a new type of modelling/programming language can be developed and applied to solve general problems with software design and implementations.

In this talk, the speaker examines the requirement of research on collaboration systems and technologies, discusses RBC and its model E-CARGO; reviews the related research achievements on RBC and E-CARGO in the past years; illustrates those problems that have not yet been solved satisfactorily; presents the fundamental methods to conduct research related to RBC and E-CRAGO and discover related problems; and analyzes their connections with other cutting-edge fields. This talk aims to inform the audience that E-CARGO is a well-developed model and has been investigated and applied in many ways. The speaker welcomes queries, reviews, studies, applications, and criticisms.

As case studies of E-CARGO, GRA and its related problem models are inspired by delving into the details of the E-CARGO components and the RBC process. GRA can help solve related collaboration problems with the help of programming and optimization platforms. All the related Java codes can be downloaded by GitHub: https://github.com/haibinnipissing/E-CARGO-Codes. The speaker welcomes interested researchers and practitioners to use these codes in their research and practice and contact the speaker if there are any questions about them.



Dr. Haibin Zhu is a Full Professor and the Coordinator of the Computer Science Program, the Founding Director of the Collaborative Systems Laboratory, a member of Arts and Science Executive Committee, Nipissing University, Canada. He is an affiliate professor of Concordia Univ. and an adjunct professor of Laurentian Univ., Canada. He received his PhD degree in computer science from the National Univ. of Defense Tech. (NUDT), China. He was the chair of the Department of Computer Science and Mathematics, Nipissing University, Canada (2019-2021), a visiting professor and special lecturer in the College of Computing Sciences, New Jersey Institute of Technology, USA (1999-2002) and a lecturer, an associate professor and a full professor at NUDT (1988-2000). He has accomplished (published or in press) over 300+ research works including 60+ IEEE Transactions articles, six books, five book chapters, four journal issues, and four conference proceedings. He is a Fellow of *IEEE, AAIA (Asia-Pacific Artificial Intelligence Association)* and *I2CICC (International Institute of Cognitive Informatics and Cognitive Computing)*, a Senior Member of *ACM*, a Full Member of *Sigma Xi*, and a Life Member of *CAST-USA* (Chinese Association of Science and Technology, USA).

He is serving as Vice President, Systems Science and Engineering (SSE) (2023-), a member-at-large of the Board of Governors (2022-), and a co-chair (2006-) of the technical committee of *Distributed Intelligent Systems* of *IEEE Systems, Man and Cybernetics (SMC) Society (SMCS)*, SMCS Primary Representative, *IEEE Systems Council,* Editor-in-Chief of *IEEE SMC Magazine* (2022), Associate Editor (AE) *of IEEE Systems Journal* (2024-), *IEEE Transactions on SMC: Systems* (2018-), *IEEE Transactions on Computational Social Systems*(2018-), *Frontiers of Computer Science* (2021-), and IEEE Canada Review (2017-). He was AE of *IEEE SMC Magazine* (2015-2021), Associate Vice President (AVP), SSE (2021), IEEE SMCS, a Conference (Co-)Chair and Program (Co-)Chair for many international conferences, and a PC member for 150+ academic conferences.

He is the founding researcher of *Role-Based Collaboration* and the creator of *the E-CARGO model.* His research monograph *E-CARGO and Role-Based Collaboration* can be found https://www.wiley-vch.de/en/areas-interest/engineering/electrical-electronics-engineering-10ee/systems-engineering-management-10ee9/e-cargo-and-role-based-collaboration-978-1-119-69306-2. The accompanying codes can be downloaded from GitHub: https://github.com/haibinnipissing/E-CARGO-Codes. He has offered 38 keynote and plenary speeches for international conferences and 93 invited talks internationally. He has been granted more than $1M CAD from SSHRC, NSERC, IBM, DNDC, DRDC, and OPIC.

He was listed as "Most Influential Robotics Trailblazers, Making Wave in The Industry – 2024", InsightsSuccess Magazine. He was the recipient of the best paper award from the 28th Int'l conf. on CSCWD, Compiegne, France, 2025 and the best paper award for international collaboration from the 25th Int'l conf. on CSCWD, Hangzhou, China, 2022, the meritorious service award from IEEE SMC Society (2018), the chancellor's award for excellence in research (2011) and two research achievement awards from Nipissing University (2006, 2012), the IBM Eclipse Innovation Grant Awards (2004, 2005), the Best Paper Award from the 11th ISPE Int'l Conf. on Concurrent Engineering (ISPE/CE2004), the Educator's Fellowship of OOPSLA'03, a 2nd class National Award for Education Achievement (1997), and three 1st Class Ministerial Research Achievement Awards from China (1997, 1994, and 1991).

His research interests include Collaboration/Complex Systems, Human-Machine Systems, Computational Social Systems, Collective Intelligence, Multi-Agent Systems, Software Engineering, and Distributed Intelligent Systems.

Introduction to Keynote Speech B

Title: 深伪鉴伪与网络认知安全
Speaker: 罗向阳教授、信息工程大学教授、博导、河南省网络态势感知重点实验室主任

ABSTRACT

阐述网络认知安全的概念和内涵要义，介绍以生成式人工智能为代表的信息内容深度伪造和鉴伪典型方法，阐述其在网络认知安全中的应用场景和重要意义，报告团队相关研究进展。

罗向阳，信息工程大学教授、博导，河南省网络态势感知重点实验室主任，中原学者、国防卓青、军队学科拔尖人才。主要从事网络安全与多媒体安全领域研究，先后主持国自然基金6项（其中重点3项），主持国家重点研发计划、军队和省部级重点项目30余项，在TPAMI、TIFS、TNET、CCS、《中国科学》等重要期刊会议发表论文200余篇，获省部级自然科学、技术发明和科技进步一等奖4项、二等奖4项，获国家教学成果一等奖、省级教学成果特等奖和一等奖各1项。

Introduction to Keynote Speech C

Title: 复杂三维模型的处理、理解与生成技术及应用
Speaker: 范晓鹏教授，哈工大计算学部二级教授、博导、智能接口与人机交互研究中心主任

ABSTRACT

随着数字城市、元宇宙等应用需求的增长，以及计算机处理能力的提升，以点云、Mesh 等为代表的三维模型相关技术获得了更多的关注。点云、Mesh 等三维信号相较于文本、图像等数据具有结构复杂、数据量大、形式不规则等特点。如何深入挖掘其特性，提升三维模型的处理能力，是目前图形图像领域热点问题。本报告将从信号特性入手，介绍团队在三维模型的去噪、修补、建模、生成、搜索、渲染等方面的研究成果，以及这些技术在数字媒体应用落地方面的探索。

范晓鹏，哈工大计算学部二级教授、博导、智能接口与人机交互研究中心主任、苏州研究院数字孪生与具身智能团队负责人、鹏城国家实验室高级研究员。入选长江学者（2021）、教育部新世纪优秀人才（2011年）等。在虚拟现实方面牵头制定全景式交互化网络视听行业标准体系并立项工信部行业标准10余项。在视频编码方面作为副主编制定了IEEE1857国际标准，作为主要贡献者之一制定了AVS/AVS3等国家/行业标准。获发明专利20余项。2013年获IEEE标准杰出贡献奖。2023年获电子学会创新团队奖。在图形图像、无线通信及机器人控制等方向发表论文240余篇，其中包括CCF A类或IEEE汇刊论文90余篇。牵头主持国家重点研发计划项目、国家自然科学基金重点项目等国家级项目6项。培养的学生获国家自然科学基金委首批博士生项目资助。研发的多项技术在航天科工集团、河南电视台、华为公司、神华集团、字节跳动等落地应用，覆盖用户5000万以上。2017年作为程序主席主办CCF推荐会议PCM2017。担任中国人工智能学会（CAAI）教育工作委员会副主任、黑龙江省计算机学会学术工作委员会主任、脑机接口技术应用应急管理部重点实验室学术委员会主任等。

Introduction to Keynote Speech D

Title: Generative Visual Data Compression and Quality Assessment
Speaker: Shiqi Wang, Professor, City University of Hong Kong

ABSTRACT

This talk introduces the latest advancements in quality assessment and multimedia compression techniques driven by generative models. It begins by presenting an overview of generative models, in particular the large multimodal foundation models and their transformative role in various vision tasks. The shift from traditional learning paradigms to prompt- and instruction-based reasoning is further emphasized. The talk then delves into the methodologies of leveraging these models for visual quality assessment (VQA). Subsequently, we explore the multimedia compression techniques enhanced by AIGC methodologies, covering lossless and lossy compression strategies for images and videos. State-of-the-art performance benchmarks and case studies demonstrate how these advanced methods outperform traditional approaches significantly.

Shiqi Wang received the PhD degree in computer application technology from Peking University in 2014. He is a professor with the Department of Computer Science at the City University of Hong Kong, and serves as the Associate Dean in the College of Computing, CityUHK. He has proposed more than 70 technical proposals to ISO/MPEG, ITU-T, and AVS standards. He authored or coauthored more than 300 refereed journal articles/conference papers, including more than 150 IEEE Transactions. These papers have received more than 18000 Citations (Google Scholar). His research interests include semantic and visual communication, AI-generated content management, machine learning, information forensics and security, and image/video quality assessment. He received the Best Paper Award from IEEE VCIP 2019, ICME 2019, IEEE Multimedia 2018, and PCM 2017. His co-authored article received the Best Student Paper Award at the IEEE ICIP 2018. He was the TPC co-chair of ICME 2024. He served or serves as an associate editor for IEEE Transactions on Circuits and Systems for Video Technology, IEEE Transactions on Multimedia, IEEE Transactions on Image Processing, and IEEE Transactions on Cybernetics.

Introduction to Keynote Speech E

Title: Implicit Neural Representation for Video Coding
Speaker: Xifeng Zhang, Associate Professor, University of Chinese Academy of Sciences

ABSTRACT

With the development of technology, video coding is facing performance bottlenecks. The emergence of deep learning technology has made it possible to break through the performance of video coding. This talk will introduce the main problems in previous video coding frameworks, and share our research progress in video implicit representation coding, which is also a new video coding framework.

Dr. Zhang Xinfeng is an associate professor at the University of Chinese Academy of Sciences. His research interests include video coding, quality assessment and video enhancement processing. He has published over 200 academic papers in international conferences and journals, with 5 best paper awards in journals and conferences. At present, he serves as an Associate Editor member for IEEE TIP and TCSVT.