IFTC 2023

2023

INTERNATIONAL ACADEMIC FORUM
OF DIGITAL MULTIMEDIA COMMUNICATION

CONFERENCE
GUIDE

2023 INTERNATIONAL ACADEMIC FORUM OF DIGITAL MULTIMEDIA COMMUNICATION

Conference Theme

Artificial Intelligence for Multimedia Communication

The 20th International Forum of Digital Multimedia Communication (IFTC 2023) is a summit forum in the field of digital multimedia communication. The Forum is co-hosted by SIGA and Springer CCIS (approval pending). It is orgnazied by State Key Laboratory of Media Convergence and Communication, School of Data Science and Intelligent Media of Communication University of China, and is co-organized by Key Laboratory of Media Audio & Video of Ministry of Education (Communication University of China), Guangdong South New Media Incorporated Company. The Forum is co-sponsored by Shanghai Jiao Tong University, Shanghai Institute for Advanced Communication and Data Science, Shanghai Telecom Company, Shanghai Key Laboratory of Digital Media Processing and Communication, and Shanghai Qing Assistant Technology Company Ltd.

The 20th IFTC serves as an international bridge for extensively exchanging the latest research advances of digital multimedia communication around the world as well as the relevant policies of industry authorities. The Forum is also to promote the technology, equipment and applications in the field of digital media by comparing the characteristics, frameworks, significant techniques and their maturity, analyzing the performance of various applications in terms of scalability, manageability and portability and discussing the interfaces among varieties of networks and platforms, and foster scientific exchange between researchers, practitioners, scientists, students and engineers in digital multimedia communication and its affiliated disciplines.

Conference Organization

- **HOST**

Shanghai Image & Graphics Association (SIGA)

Springer CCIS

- **SUPPORT**

Shanghai Association for Science & Technology

Association for Science & Technology of CUC

- **ORGANIZER**

State Key Laboratory of Media Convergence and Communication (CUC)

School of Data Science and Intelligent Media, CUC

- **CO-ORGANIZER**

Key Laboratory of Media Audio & Video (CUC), Ministry of Education

Guangdong South New Media Inc.

- **SPONSOR**

Shanghai Jiao Tong University

Shanghai Institute for Advanced Communication and Data Science

Shanghai Telecom Company

Shanghai Key Laboratory of Digital Media Processing and Communication

Shanghai Qing Assistant Technology Company Ltd.

Conference Committee

- ## General Chairs

  Xiaokang Yang (Shanghai Jiao Tong University, China)

  Ping An (Shanghai University, China)

  Long Ye (Communication University of China, China)

  Guangtao Zhai (Shanghai Jiao Tong University, China)

- ## Program Chairs

  Yue Lu (East China Normal University, China)

  Jun Zhou (Shanghai Jiao Tong University, China)

  Yuan Zhang (Communication University of China, China)

  Haonan Cheng (Communication University of China, China)

  Hua Yang (Shanghai Jiao Tong University, China)

- ## International Liaisons

  Weisi Lin (Nanyang Technological University, Singapore)

  Patrick Le Callet (Nantes University, France)

  Lu Zhang (INSA Rennes, France)

- ## Finance Chairs

  Yi Xu (Shanghai Jiao Tong University, China)

  Hao Liu (Donghua University China)

  Beibei Li (Shanghai Polytech. University, China)

  Xuefei Song (Shanghai Ninth People's Hospital, China)

- ## Tutorial Chairs

  Xiangyang Xue (Fudan University, China)

  Yuming Fang (Jiangxi University of Fin. & Eco., China)

  Jiantao Zhou (University of Macau, China)

Conference Committee

- ## Publications Chairs

  Hong Lu (Fudan University, China)

  Feiniu Yuan (Shanghai Normal University, China)

  Menghan Hu (East China Normal University, China)

  Liquan Shen (Shanghai University, China)

- ## Award Chairs

  Zhijun Fang (Donghua University, China)

  George Wang (Guangdong South New Media Inc., China)

  Xiaolin Huang (Shanghai Jiao Tong University, China)

  Hanli Wang (Tongji University, China)

  Yu Zhu (East China University Sci. & Tech., China)

- ## Publicity Chairs

  Bo Yan (Fudan University, China)

  Wei Zhong (Communication University of China, China)

  Juanjuan Cai (Communication University of China, China)

  Gang Hou (Central Res. Inst. of INESA, China)

- ## Industrial Program Chairs

  Yiyi Lu (China Telecom Shanghai Branch, China)

  Zhiye Chen (Guangdong South New Media Inc., China)

  Guozhong Wang (Shanghai University of Tech. Sci., China)

  Chen Yao (Third Res. Inst. of MPS, China)

  Yan Zhou (Renji Hospital)

- ## Arrangement Chairs

  Cheng Zhi (SIGA, China)

  Jia Wang (Shanghai Jiao Tong University, China)

Conference Schedule

Date: Dec. 21st, 2023 (Thursday)

Time	Program	Location
9:00-9:10	Opening Ceremony	Lecture Hall V01 Academic Center, CUC
9:10-9:50	Keynote Speech Jiebo Luo University of Rochester	
9:50-10:30	Keynote Speech Qin Zhang Communication University of China	
10:30-10:45	Photos & Coffee Break	Lecture Hall V10 Academic Center, CUC
10:45-12:15	Poster Presentation 1	
10:45-12:15	Oral Session 1	Lecture Hall V01 Academic Center, CUC
12:30-14:30	Lunch	International Convention Center, CUC
14:30-15:10	Keynote Speech Xiangyang Ji Tsinghua University	Lecture Hall V01 Academic Center, CUC
15:10-15:50	Keynote Speech Jingdong Wang Baidu	
15:50-16:10	Coffee Break	Lecture Hall V10 Academic Center, CUC
16:10-17:40	Poster Presentation 2	
16:10-17:40	Oral Session 2	Lecture Hall V01 Academic Center, CUC
18:00-20:00	Banque Best Paper & Best Poster Selection	International Convention Center, CUC

Conference Program

Date: Dec. 21st, 2023 (Thursday)

Time	Program		Location
9:00-9:10	Opening Ceremony	Opening Remarks: **Jianping Chai** Vice President Communication University of China **Guangtao Zhai** Chairman of Shanghai Image & Graphics Association Professor of Shanghai Jiao Tong University	Lecture Hall V01 Academic Center, CUC
9:10-9:50	Keynote Speech 1	**Jiebo Luo** University of Rochester Title: GPT-4V (vision) as A Social Media Analysis Engine	
9:50-10:30	Keynote Speech 2	**Qin Zhang** Communication University of China Title: Theory of Emotional Intelligence Model Based on AI Networks	
10:30-10:45	Photos & Coffee Break		Lecture Hall V10 Academic Center, CUC
10:45-12:15	Poster Presentation 1		
10:45-12:15	Oral Session 1		Lecture Hall V01 Academic Center, CUC
	10:45-11:00	**AUIQE: Attention-Based Underwater Image Quality Evaluator** Baochao Zhang, Chenghao Zhou, Runze Hu, Jingchao Cao, Yutao Liu	

Conference Program

Date: Dec. 21st, 2023 (Thursday)

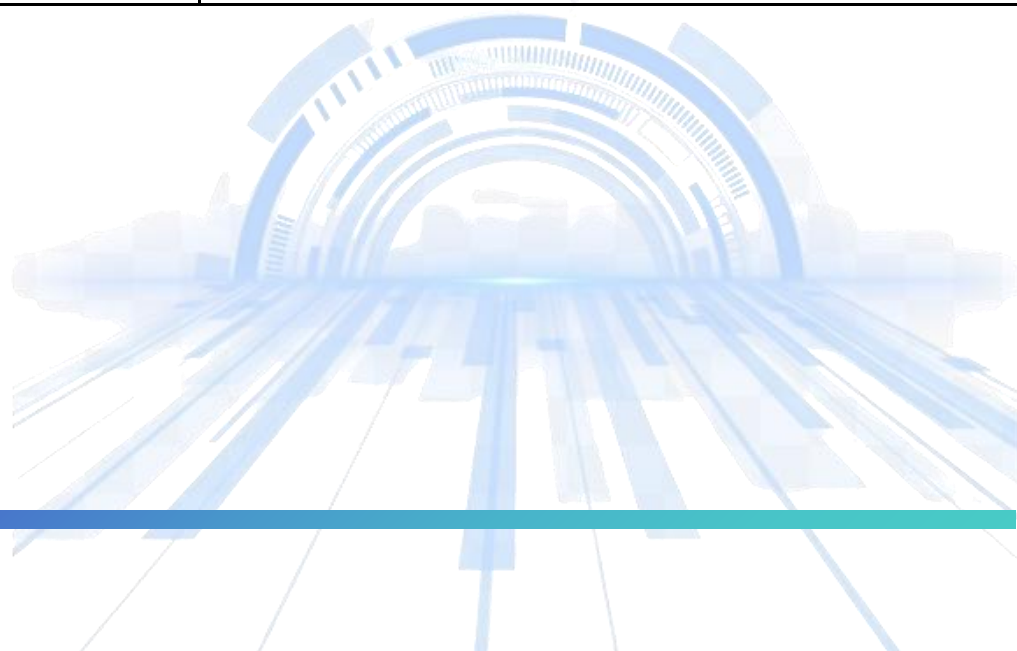| Time	Program		Location
10:45-12:15	Oral Session 1		Lecture Hall V01 Academic Center, CUC
	11:00-11:15	**Depth Map Super-Resolution via Asymmetrically Guided Feature Selection and Spatial Affine Transformation** Jintao Fan, Yi Xu	
	11:15-11:30	**Human-centered Financial Signal Processing: A Case Study on Stock Chart Analysis** Kaixun Zhang, Yuzhen Chen, Ji-Feng Luo, Menghan Hu, Guangtao Zhai, Xiao-Ping Zhang	
	11:30-11:45	**Unbiased Image Caption Generation based on Dynamic Counterfactual Inference** Jinfei Zhou	
	11:45-12:00	**A Multimodal Registration and Fusion Diagnostic System Based on Multi-scale Feature** Hongyi Jing, Jiannan Liu, Jing Han, Guangtao Zhai	
	12:00-12:15	**CCDaS: A Benchmark Dataset for Cartoon Character Detection in Application Scenarios** Zelu Qiu	
12:30-14:30	Lunch		International Convention Center, CUC

Conference Program

Date: Dec. 21st, 2023 (Thursday)

Time	Program			Location
14:30-15:10	Keynote Speech 3		**Xiangyang Ji** Tsinghua University Title: Vision-based object pose estimation	Lecture Hall V01 Academic Center, CUC
15:10-15:50	Keynote Speech 4		**Jingdong Wang** Baidu Title: Towards Large Vision Models: Self-Supervised Pretraining and Video Generation	
15:50-16:10	Coffee Break			Lecture Hall V10 Academic Center, CUC
16:10-17:40	Poster Presentation 2			
16:10-17:40	Oral Session 2			Lecture Hall V01 Academic Center, CUC
		16:10-16:25	**ULIC: Ultra Lightweight Image Coder on Wearable Devices** Muchen Dong, Hongwei Sha, Ming Lu, Zhan Ma	
		16:25-16:40	**Link-breakage Recovery Strategies for Tactical MANET Based on OLSR** Lianghui Ding, Hongyu Gao, Huan Lin, Feng Yang	
		16:40-16:55	**Temporal Dependency-Oriented Deep In-Loop Filter for VVC** Ziyi Zhuang, Li Li, Dong Liu	

Conference Program

Date: Dec. 21st, 2023 (Thursday)

Time	Program			Location
16:10-17:40	Oral Session 2			Lecture Hall V01 Academic Center, CUC
		16:55-17:10	**Unsupervised Event-to-Image Reconstruction Based on Domain Adaptation** Jupo Ma, Zhenqiang Zhao, Wen Yang	
		17:10-17:25	**ART-InvRec: Acquiring Rotation Invariance of 3D Object Reconstruction via Adversarial Rotation** Rui Yang, Fei Hu	
		17:25-17:40	**MRTV: A Novel Method for Facial Expression Recognition in Natural Scenes** Jihua Ye, Youcai Zou, Tiantian Wang, Chao Wang, Zhan Xu	
18:00-20:00	Banque Best Paper & Best Poster Selection			International Convention Center, CUC

Conference Program

Date: Dec. 21st, 2023 (Thursday)

Time	Program	Location
10:45-12:15	**Poster Presentation 1** ➢ **AquaSAM: Underwater Image Foreground Segmentation** Muduo Xu, Jianhao Su, Yutao Liu ➢ **RAUNE-Net: A Residual and Attention-Driven Underwater Image Enhancement Method** Wangzhen Peng, Chenghao Zhou, Runze Hu, Jingchao Cao, Yutao Liu ➢ **EHA3D: Expressive Head Avatar via Disentangled Latent Code** Jiayu Zhou, Xiaoqiang Zhu ➢ **Design of Distributed Synchronization Header for Short Burst Communication with Low Earth Orbit Satellites** Zehua Ding ➢ **PM2.5 Concentration Measurement Based On Natural Scene Statistics and Progressive Learning** Guangcheng Wang, Baojin Huang, Kezheng Sun, Lijuan Tang, Mengting Wei, Quan Shi ➢ **IIQD2023: A Novel Database for Assessing Infrared Imaging Quality Evaluation Algorithms** Chengxu Zhou, Xiaojie Fan, Ke Gu ➢ **Where to Forget: A New Attention Stability Metric for Continual Learning Evaluation** Haojie Wang, Qingbo Wu, Hongliang Li, Fanman Meng ➢ **Welding Defect Detection using X-ray Images based on Deep Segmentation Network** Yawen Fan, Zhengkai Hu, Xuefeng Fang, Junge Fang, Qinxin Li	Lecture Hall V10 Academic Center, CUC

Conference Program

Date: Dec. 21st, 2023 (Thursday)

Time	Program	Location
	Poster Presentation 1	
	➢ **Wide Activation Fourier Channel Attention Network for Super-Resolution** Xuan Wu, Ming Tan, Liang Chen, Yi Wu	
	➢ **Perceptual Blind Panoramic Image Quality Assessment Based on Super-Pixel** Shuyu Xiao, Yongfang Wang, Yihan Wang, Zhijun Fang	
	➢ **Motion-aware Topology Learning for Skeleton-based Action Recognition** Yanjun Chen, Hao Zhou, Yan Luo, Chongyang Zhang, Chuanping Hu	
10:45-12:15	➢ **Image Aesthetics Assessment Based on Visual Perception and Textual Semantic Understanding** Yun Liu, Zhipeng Wen, Sifan Li, Daoxin Fan, Guangtao Zhai	Lecture Hall V10 Academic Center, CUC
	➢ **LightNet+: Boosted Light-weighted Network for Smoke Semantic Segmentation** Kang Li, Chunmei Wang, Chunli Meng, Feiniu Yuan	
	➢ **Modeling and Analysis of Rumor Propagation Dynamics in Social Media** Shan Liu, Hanfei Zhao	
	➢ **MFNCA: Multi-level Fusion Network Based on Cross Attention for 3D Point Cloud Object Detection** Shuo Zhu, Yongfang Wang, Wei Chen, Zhijun Fang	
	➢ **Adjusting Exploitation and Exploration Rates of Differential Evolution: A Novel Mutation Strategy** Danting Duan, Yuhui Zhang, Wei Zhong, Long Ye, Qin Zhang	

Conference Program

Date: Dec. 21st, 2023 (Thursday)

Time	Program	Location
10:45-12:15	**Poster Presentation 1** ➢ **Revealing Real Face for Generalized Anti-Spoofing** Weiye Tang, Zhiyong Huang, Qiu Shen ➢ **Study on Sound Insulation Performance of Membrane-Type Acoustic Metamaterials with Pendulum Arm** Ke Wang, Lujin Xiong, Shouhao Wu ➢ **Local Dynamic Filter Network for Joint Low-light Enhancement and Deblurring** Nanxin Huang, Lifang Yang, Yirui Wang ➢ **An Omnidirectional Videos Quality Assessment Method using Salient Object Information** Zelu Qi, Kai Jia ➢ **An Anomaly Detection Framework for Propagation Networks Leveraging Deep Learning** Yuewei Wu, Zhenyu Yu, Zhiqiang Zhang, Junyi Chen, Fulian Yin ➢ **Bridging Recommendations Across Domains: An Overview of Cross-Domain Recommendation** Jing Liu, Lele Sun, Litao Shang ➢ **MABC-Net: Multimodal Mixed Attentional Network with Balanced Class for Temporal Forgery Localization** Haonan Cheng, Haixin Yu, Li Fang, Long Ye ➢ **Sequence Modeling Based Data Augmentation for Micro-Expression Recognition** Xin Lin, Shaojie Ai, Junying Gao, Jing He, Lei Yan, Jiaqi Zhang,Jing Liu	Lecture Hall V10 Academic Center, CUC

Conference Program

Date: Dec. 21st, 2023 (Thursday)

Time	Program	Location
16:10-17:40	**Poster Presentation 2** ➤ **Multi-view 3D Object Classification and Retrieval Based on Deep Models** Jing Liu, Xin Li ➤ **Resolution-Agnostic Neural Compression for High-Fidelity Portrait Video Conferencing via Implicit Radiance Fields** Yifei Li, Xiaohong Liu, Yicong Peng, Guangtao Zhai, Jun Zhou ➤ **Combined Particle Filter and Its Application on Human Pose Estimation** Xinyang Liu, Long Ye, Yinghao Yang ➤ **Exploiting Diffusion Model as Prompt Generator for Object Localization** Yuqi Jiang, Qiankun Liu, Yichen Li, Hao Jia, Ying Fu ➤ **Dual Transformer with Gated-Attention Fusion for News Disaster Image Captioning** Yinghua Li ➤ **Depression Recognition Based on Pre-Trained ResNet-18 Model and Brain Effective Connectivity Network** Xiaoying Zhao, Hailing Wang, Tingwei Jiang ➤ **A No-Reference Stereoscopic Image Quality Assessment Based on Cartoon Texture Decomposition and Human Visual System** Yun Liu, Yan Bai, Yaohui Wang, Minzhu Jin, Bo Liu ➤ **ChatASD: LLM-Based AI Therapist for ASD** Xiaoyu Ren, Yuanchen Bai, Huiyu Duan, Lei Fan, Erkang Fei, Geer Wu, Pradeep Ray, Menghan Hu, Guangtao Zhai	Lecture Hall V10 Academic Center, CUC

Conference Program

Date: Dec. 21st, 2023 (Thursday)

Time	Program	Location
16:10-17:40	**Poster Presentation 2** ➢ **Spatial-Angular Decoupling Interaction Networks for Light Field Angular Super-Resolution** Baoshuai Wang, Yilei Chen, Xinpeng Huang, Ping An ➢ **Billiards Hitting Assistance System** Jian Zhang, Menghan Hu, Hang Liu, Chaoyi Liu ➢ **Visual Detection System for Industrial Defects** Lei Wang, Aiming Xu, Zhiyong Huang, Qiu Shen ➢ **A Human Posture Estimation Method for Image Interaction System Based on ECA** Shuqi Wang ➢ **Fast QTMT Decision for H.266/VVC via Jointly Leveraging Neural Network and Machine Learning Models** Gongchun Ding, Xiujun Lin, Wenyu Wang, Dandan Ding ➢ **End-to-End Image Compression Through Machine Semantics** Jianran Liu, Chang Zhang, Wen Ji ➢ **PU-SSIM:A Perceptual Constraint for Point Cloud Up-Sampling** Ruijun Liu, Tiangang Huang, Xiaochuan Wang, Haisheng Li ➢ **A Low-Resolution Face Recognition Method Fusing Spatial Information and Multi-Scale Attention Features** Jihua Ye, Tiantian Wang, Youcai Zou, Chao Wang, Zhan Xu	Lecture Hall V10 Academic Center, CUC

Conference Program

Date: Dec. 21st, 2023 (Thursday)

Time	Program	Location
16:10-17:40	**Poster Presentation 2** ➢ **Quality Enhancement via Spatial-Angular Deformable Convolution for Compressed Light Field** Ping An, Yongjie Lu, Xinpeng Huang,Chao Yang ➢ **Coding Prior-Driven JPEG Image Artifact Removal** Dongliang Cui, Youwei Pan, Wuzhen Shi, Yang Wen, Zicheng Liu, Yutao Liu ➢ **User Preferences Based Preloading and ABR Algorithm For Short Video Streaming** Lanju Zhang, Yuan Zhang, Jinyao Yan ➢ **Constructing Personal Knowledge Graph from Conversation via Deep Reinforcement Learning** Fei Cai, Xiao Guo ➢ **Decoding the Flow Experience in Video Games: An Analysis of Physiological and Performance Metrics** Qi Wu, Di Zhang, Xinhui Huang, Long Ye, Boning Zhang ➢ **Driving Dynamics an In-depth Analysis of Attention Allocation Between Driver and Co-Driver in a Simulated Environment** Xinbin Chen, Kai Wang, Chudan Zhou, Wenshan Shi ➢ **Exploring the Efficacy of Interactive Digital Humans in Cultural Communication** Simin Chen, Wenshan Shi, Xinyue Ding	Lecture Hall V10 Academic Center, CUC

Keynote Speech

Presenter : Jiebo Luo

Jiebo Luo is the Albert Arendt Hopeman Professor of Engineering and Professor of Computer Science at the University of Rochester, which he joined in 2011 after a prolific career of fifteen years at Kodak Research Laboratories. He has authored over 600 technical papers and holds over 90 U.S. patents. His research interests include computer vision, NLP, machine learning, data mining, computational social science, and digital health. He has been involved in numerous technical conferences, including serving as program co-chair of ACM Multimedia 2010, IEEE CVPR 2012, ACM ICMR 2016, and IEEE ICIP 2017, and general co-chair of ACM Multimedia 2018 and IEEE ICME 2024. He has served on the editorial boards of the IEEE Transactions on Pattern Analysis and Machine Intelligence, IEEE Transactions on Multimedia, IEEE Transactions on Circuits and Systems for Video Technology, IEEE Transactions on Big Data, ACM Transactions on Intelligent Systems and Technology, Pattern Recognition, Knowledge and Information Systems, Machine Vision and Applications, and Intelligent Medicine. He was the Editor-in-Chief of the IEEE Transactions on Multimedia (2020-2022). Professor Luo is a Fellow of ACM, AAAI, IEEE, SPIE, and IAPR , as well as a Member of Academia Europaea and the US National Academy of Inventors. He received the ACM SIGMM Technical Achievement Award in 2021.

Time : 9:10-9:50 Location : Academic Center V01, CUC

Title : GPT-4V (vision) as A Social Media Analysis Engine

Abstract : Recent research has offered insights into the extraordinary capabilities of Large Multimodal Models (LMMs) in various general vision and language tasks. There is growing interest in how LMMs perform in more specialized domains. Social media content, inherently multimodal, blends text, images, videos, and sometimes audio. To effectively understand such content, models need to interpret the intricate interactions between these diverse communication modalities and their impact on the conveyed message. Understanding social multimedia content remains a challenging problem for contemporary machine learning frameworks. In this paper, we explore GPT-4V (Vision)'s capabilities for social multimedia analysis. We select five representative tasks, including sentiment analysis, hate speech detection, fake news identification, demographic inference, and political ideology detection, to evaluate GPT-4V. Our investigation begins with a preliminary quantitative analysis for each task using existing benchmark datasets, followed by a careful review of the results and a selection of qualitative samples that illustrate GPT-4V's potential in understanding multimodal social media content. GPT-4V demonstrates remarkable efficacy in these tasks, showcasing strengths such as joint understanding of image-text pairs, contextual and cultural awareness, and extensive commonsense knowledge. Despite the overall impressive capacity of GPT-4V in the social media domain, there remain notable challenges.GPT-4V struggles with tasks involving multilingual social multimedia comprehension and has difficulties in generalizing to the latest trends in social media. Additionally, it exhibits a tendency to generate erroneous information in the context of evolving celebrity and politician knowledge, reflecting the known hallucination problem. Our hope is that this preliminary study will provide insights into further research across disciplines, particularly in computational social science and social media-related studies. The insights gleaned from our findings underscore a promising future for LMMs in enhancing our comprehension of social media content and its users through the analysis of multimodal information.

Keynote Speech

Presenter : Qin Zhang

Qin Zhang is currently a Professor with the State Key Laboratory of Media Convergence and Communication, Communication University of China. His research interests include computer vision, audio & video signal processing and virtual reality. Since joining in the Communication University of China , he has led and completed more than 20 key and general projects funded by the National Natural Science Foundation of China, Ministry of Education, National Radio and Television Administration and so on, achieving multiple systematic and innovative achievements. In the field of audio technology, he has independently developed a 96-channel DMS sound field reconstruction system, breaking through Dolby and DTS in four main technical aspects of graded service, voice control, sound field synthesis and dynamic response. In the field of video technology, he has proposed a video content representation method based on primitives/textures and also developed corresponding parameterized models, filling the gap between pixel-level and object-level representations, with broad application prospects.

Time : 9:50-10:30   Location : Academic Center V01, CUC

Title : Theory of Emotional Intelligence Model Based on AI Networks

Abstract : One of the main scientific issues in human self-awareness is the coupling relationship and mechanism between emotions and intelligence under specific information inputs. At present, the research on the relationshipbetween emotions and intelligence is still in the qualitative research stage, lacking comprehensive theories and tools. This talk focuses on the perception and modeling of the coupling relationship between emotion and intelligence. Starting from the system structure, model framework and mathematical theory of the problem, a comprehensive emotion and intelligence model will be constructed. At the same time, the corresponding theoretical tools will be developed for the applications of emotion and intelligence models.
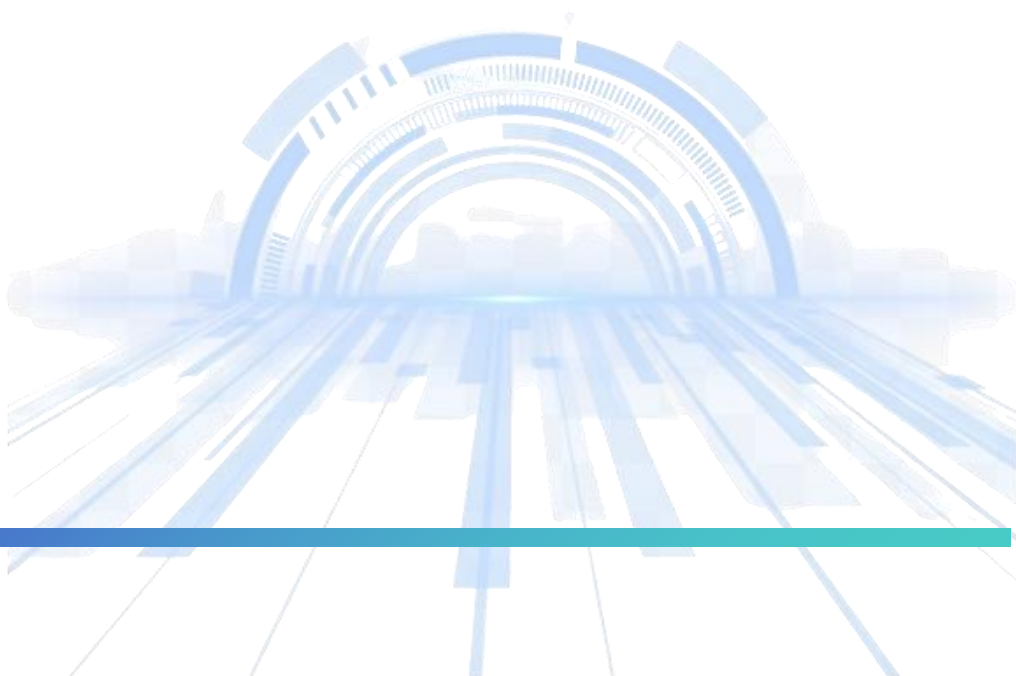
Keynote Speech

Presenter : Xiangyang Ji

Professor of the Department of Automation of Tsinghua University, the main research directions include machine learning, visual signal acquisition and processing. He is the director of the Intelligence Science Division of the Beijing National Research Center for Information and Technology, and the director of the Department of Brain Science and Cognitive Science of the Automation Department. The National Science Fund for Distinguished Young Scholars, China Youth Science and Technology Award and other academic honors are awarded. He serves as the chairman of the Youth Working Committee of the Chinese Institute of Electronics, and the director of the Deep Learning Special Committee of the Chinese Association for Artificial Intelligence. He has published more than 100 SCI/EI papers in TCI, TPAMI, IJCV, NIPS, ICML, CVPR, ICCV and was authorized more than 40 national invention patents, 10 international invention patents, and won the 70th Nuremberg International Invention Gold Award 2. In addition, he and his team won the championship MSCOCO visual instance segmentation, ECCV visual quality enhancement, ICCV/ECCV 6D pose estimation and other international competitions etc. He was awarded the second prize of the State Scientific and Technological Progress Award in 2019 (Rank 1st)
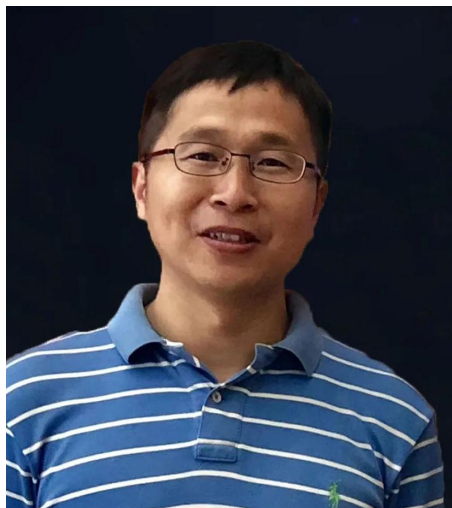
Time : 14:30-15:10   Location : Academic Center V01, CUC

Title : Vision-based object pose estimation

Abstract : One of the main scientific issues in human self-awareness is the coupling relationship and mechanism between emotions and intelligence under specific information inputs. At present, the research on the relationshipbetween emotions and intelligence is still in the qualitative research stage, lacking comprehensive theories and tools. This talk focuses on the perception and modeling of the coupling relationship between emotion and intelligence. Starting from the system structure, model framework and mathematical theory of the problem, a comprehensive emotion and intelligence model will be constructed. At the same time, the corresponding theoretical tools will be developed for the applications of emotion and intelligence models.

Keynote Speech

Presenter : Jingdong Wang

Jingdong Wang is a Chief Scientist for Computer Vision with Baidu. Before joining Baidu, he was a Senior Principal Research Manager with Microsoft Research Asia. His areas of interest are computer vision, deep learning, and multimedia search. His representative works include deep high-resolution network (HRNet), object-contextual representations for semantic segmentation (OCRNet), neighborhood graph search (SPTAG) for large scale vector search. He has been serving/served as an Associate Editor of IEEE TPAMI, IJCV, ACM TOMM, IEEE TMM, and IEEE TCSVT, and an area chair of leading conferences in vision, multimedia, and AI, such as CVPR, ICCV, NeurIPS, ECCV, ACM MM, IJCAI, and AAAI. He was elected as an ACM Distinguished Member, a Fellow of IAPR, and a Fellow of IEEE, for his contributions to visual content understanding and retrieval.

Time : 15:10-15:50   Location : Academic Center V01, CUC

Title : Towards Large Vision Models: Self-Supervised Pretraining and Video Generation

Abstract : In this talk, I introduce two works on large vision models. One is a masked image modeling approach, Context Autoencoder for self-supervised representation pretraining. The core ideas include transferring masked image modeling to masked representation modeling, making predictions in the latent representation space, and decoupling the pretraining task and the representation learning task, making representation learning much stronger. The other one is a text-to-video approach. The core ideas include leveraging a text-to-image to generate a high-quality reference image for visual fidelity and generating a video conditioned on both the text and the reference image for smooth dynamics. We demonstrate the two approaches in benchmark and real-world tasks.

State Key Laboratory of Media Convergence and Communication (CUC)

The State Key Laboratory of Media Convergence and Communication (Communication University of China) is a national key laboratory established with approval from the Ministry of Science and Technology on November 6, 2019. It operates under the supervision of the Ministry of Education and is supported by the Communication University of China for its establishment. The Communication University of China is a key university under the direct administration of the Ministry of Education and is a crucial part of the nation's "Project 211". It has officially been included in the elite group of universities receiving focused development under the "985 Project Innovation Platform".

In response to the national strategic needs of deepening the development of media integration and enhancing international communication capabilities, the goal for the establishment of the State Key Laboratory of Media Convergence and Communication is to create theories and service models for media convergence and communication, promote and innovate the entire chain of information computing models and technical systems including content production, information reach, cognitive decoding and efficiency evaluation, providing scientific theories, critical core technological supports and practical platforms for the construction of national strategic communication system with distinct Chinese characteristics. The ultimate vision is to establish a "laboratory within a laboratory" that integrates natural and social sciences, utilizing natural science methodologies to address significant social science issues. This institution is envisioned to become a significant national instrument serving state strategies in the field of information dissemination, a think tank for national security, a theoretical high ground for scientific research and innovation, and an important base for cultivating top-notch talents in media convergence.

School of Data Science and Intelligent Media

The School of Data Science and Intelligent Media was established in 2018 through the consolidation of high-caliber resources and units within the Communication University of China, which was evolved from the School of Science and belonged to the department of Information Science and Technology. Grounded in mathematics and physics, the school focuses on interdisciplinary directions with the majors of media big data, intelligent media, and mathematics as its core. In accordance with the requirements of the Ministry of Education for the construction of new engineering disciplines, it has developed an innovative talent development system which is dedicated to multi-tiered talent training, scientific research, industrial transformation, and international collaborations, thereby establishing itself as a renowned institution for teaching and research.

The faculty of the school is highly competent. Currently, it comprises 58 faculty members, including 44 dedicated teachers. Among these, 10 hold senior professional titles and 25 possess associate senior professional titles, accounting for 79% of the total number of faculty. The school currently has five departmental units, namely: Department of Data Science, Department of Intelligent Science, Department of Imaging Engineering, Department of Mathematics, and Experimental Teaching Center. Additionally, it has two supporting units, namely: Key Laboratory of Media Audio-Video Education Ministry and the National Radio and Television Administration Laboratory on Research and Application of Smart Media Microservices Technology.

The school offers an integrated and comprehensive training system for undergraduate, master's, and doctoral students. At the graduate level, there are currently five master's degree majors, namely: "Communication Data Science", "Visual Computing and Intelligent Imaging", "Advanced Computing", "Computational Mathematics and Intelligent Media" and "Applied Mathematics and Information Communication" direction. Additionally, it offers a professional master's direction in Big data technology and engineering, specifically

focusing on "Big data analysis and intelligent computing". At the undergraduate level, there are three majors, namely: data science and big data technology, intelligent science and technology (intelligent media technology), and intelligent science and technology (intelligent imaging technology). Furthermore, it has collaborated with the School of Advertising to jointly recruit undergraduate students in advertising since 2021 (Double Bachelor's Degree of compound talent training project of computational advertising). In 2022, the school has pioneered the recruitment of the first batch of undergraduate students in Intelligent science and technology (Sino-foreign cooperative education) at its Hainan campus, with a planned annual student enrollment of 100 individuals.

The school boasts a robust scientific research capacity, undertaking numerous projects funded by the National Natural Science Foundation, key R&D initiatives of the Ministry of Science and Technology, crucial technical support projects from the Ministry of Education, State Administration of Radio, Film and Television, as well as Beijing-based projects. In the past five years, it has published more than 50 SCI papers and 3 academic books.

Key Laboratory of Media Audio & Video (CUC), Ministry of Education

The Key Laboratory of Media Audio & Video (Communication University of China), Ministry of Education (hereinafter referred to as "the Laboratory"), is established by Communication University of China, relying on the advantages and characteristics of the broadcasting and television field, integrating three research units, namely, the New Creative Information Technology Research Institute, the Communication Acoustics Laboratory and the High Performance Computing Center, based on the research achievements of the applied basic discipline of media audio and video. The Laboratory was approved as a key laboratory of the Ministry of Education (MOE) in 2010 after being reviewed by the MOE (MOE Department of Science, Technology and Informatization [2011] No. 36), and officially opened for operation in January 2016 after passing the acceptance (MOE Department of Science, Technology and Informatization Letter [2016] No. 64). The Laboratory follows the new round of technological industrial revolution and digital economic development trend, focuses on immersive audiovisual content creation, explores the cross-research paradigm of "information technology" and "digital art", builds the theoretical system of immersive audiovisual computing, enhances artistic expression and cultural influence, breaks through the key technologies of three-dimensional, virtual-real fusion immersive audio and video, and demonstrates the development of immersive audiovisual industry. The Laboratory takes "immersive audiovisual information computing and artistic presentation for enhancing the influence of Chinese culture" as the scientific problem, and sets up four research directions: immersive computing and visual art presentation, sound culture and sound imitation computing, cloud-edge immersive content processing, immersive interaction and spatial art evaluation. Since its establishment, the Laboratory has undertaken 101 projects of various types, focusing on immersive audiovisual information computing and artistic presentation for enhancing the influence of Chinese

culture and the needs of industry and regional development, in the research directions or fields of immersive computing and visual art presentation, sound culture and sound imitation computing, cloud-edge immersive content processing, immersive interaction and spatial art evaluation, including 1 project and 6 sub-projects of national key R&D program, 3 key programs, 7 general program and 6 youth programs of the National Natural Science Foundation of China. It has published more than 100 papers in domestic and international journals and conferences, such as IEEE Journal on Selected Areas in Communications, IEEE Internet of Things Journal, IEEE Transactions on Multimedia, IEEE Transactions on Neural Networks and Learning Systems, IEEE Transactions on Information Forensics and Security, IEEE Transactions on Affective Computing, ACM SIGCOMM, IEEE INFOCOM, ACM MM, IJCAI, IEEE VR, etc.

Guangdong South New Media Co., Ltd.

Guangdong South New Media Co., Ltd. is one of the "Top 30 National Cultural Enterprises". Established in 2010, our company operates as a new media enterprise that seamlessly integrates modern audio-visual technology and new media platforms under Guangdong Broadcasting and Television Station. It holds the distinction of being the first broadcasting platform operating company to achieve independent IPO listing in China. The primary objective of our company is to construct an Internet-based audio-visual flagship platform that originates from Guangdong but extends its influence nationwide. Focusing on the scene of "smart home", it offers a wide range of diversified content for over 270 million users, including video, music, education, games, lifestyle services, and etc.