



MIPR 2026

The 9th IEEE International Conference on
**Multimedia Information Processing
and Retrieval**

CONFERENCE HANDBOOK

August 9-11, 2026 | Bangkok, Thailand





Guideline

Welcome all leaders and researchers to **the 9th IEEE International Conference on Multimedia Information Processing and Retrieval (IEEE MIPR 2026)**. For your convenience during the conference, please pay attention to the following concerns:

Participation

Venue: Chatrium Grand Bangkok Hotel, Thailand

Address: 728 Phetchaburi Road, Khwaeng Thanon Phetchaburi, Ratchathewi, Bangkok 10400, Thailand.

Direction

LOCATION	KM.
Suvarnabhumi International Airport	28
Don Mueang Airport	23
Siam Paragon	0.35
BTS Skytrain Siam Station	0.6
Siam Square	0.65
Siam Center	0.7
Siam Discovery	0.8
Bangkok Art and Culture Center	0.85
Central World	0.85
MBK Center	1.3
Gayson Plaza	1.4
King Power	2
Central Chidlom	2.4



MIPR 2026

The 9th IEEE International Conference on
Multimedia Information Processing and Retrieval

Guideline

Reception and Banquet

Meet, connect, and enjoy two special social events during IEEE MIPR 2026: the Welcome Reception at the conference hotel and the Conference Banquet at Jim Thompson Heritage Quarter.



Welcome Reception

Foyer outside the ballroom



Conference Banquet

Jim Thompson Heritage Quarter



📍 Foyer — outside the ballroom

Welcome Reception

Begin the conference in a relaxed and welcoming setting at Chatrium Grand Bangkok. The Welcome Reception offers attendees an opportunity to meet fellow researchers, reconnect with colleagues, and start new conversations before the main technical program. The reception will take place in the **foyer immediately outside the ballroom**, making it easy to join after arriving at the conference venue.



Conference Banquet

📍 VENUE

Jim Thompson, A Thai Restaurant

6 Kasem San 2 Alley, Wang Mai, Pathum Wan,
Bangkok 10330

🕒 DAILY HOURS

11:30–16:30

Dinner: 18:00–23:00

📞 CONTACT

+66 2 612 3601

jtrestaurant@jimthompson.com

MIPR 2026

9th International Conference on
Multimedia Information Processing and Retrieval

— Multimedia Intelligence for Sustainable Futures —

How to Get to the Banquet Venue

— ♦ From Chatrium Grand Bangkok Hotel to Jim Thompson, A Thai Restaurant ♦ —

Approximate travel options for MIPR 2026 attendees

1 Taxi / Tuk-Tuk / Ride-hailing



⌚ Approx. 30–50 min

- Convenient and direct door-to-door service.
- Evening traffic can be very heavy.
- This is often the slowest option around banquet time.
- Best for small groups with luggage.

2 BTS Skytrain (RECOMMENDED)



⌚ Approx. 15–25 min

- Walk to Siam Station.
- Take BTS Silom Line to National Stadium (1 stop).
- Short walk to the restaurant.
- Fast, reliable, and avoids traffic.

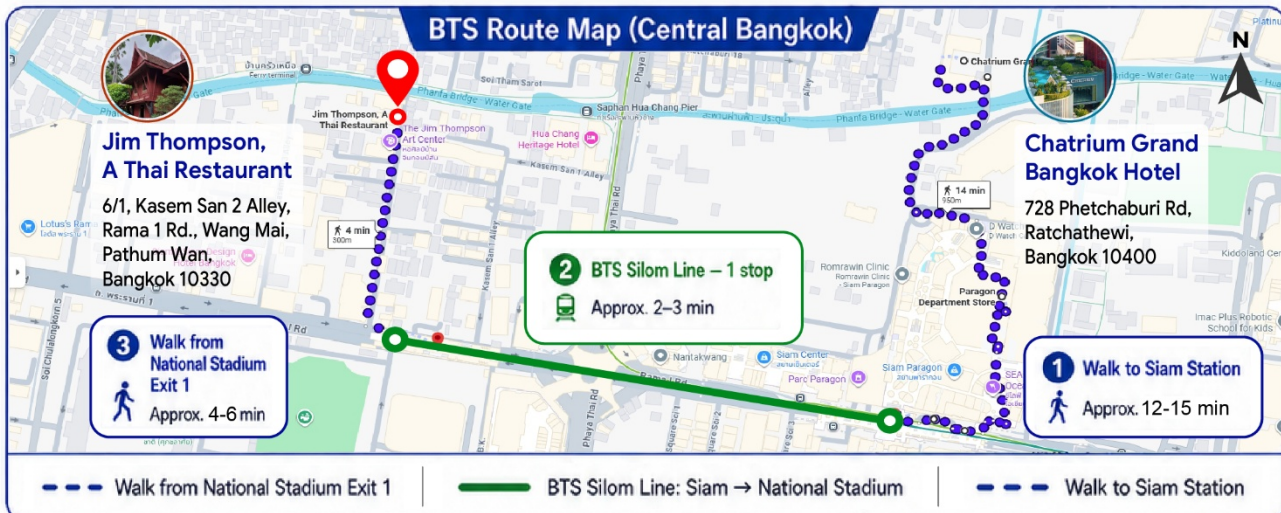
3 Walk



⌚ Approx. 20–30 min

- Possible in good weather.
- Suitable for attendees who enjoy walking.
- Use a navigation app for the latest route.

BTS Route Map (Central Bangkok)



**Recommended: During evening rush hour,
BTS is the fastest and most reliable option.**

Walk from the hotel to Siam Station, take BTS to National Stadium, then walk to Jim Thompson, A Thai Restaurant.



Travel times are approximate and may vary depending on traffic, weather, and walking pace.





MIPR 2026

The 9th IEEE International Conference on
Multimedia Information Processing and Retrieval

About MIPR 2026

With vast amounts of multimedia data now widely accessible, understanding spatial and/or temporal phenomena has become essential for numerous applications, driving the need for advanced techniques in processing, analyzing, searching, mining, and managing such data.

The 9th IEEE International Conference on Multimedia Information Processing and Retrieval (IEEE MIPR 2026) will be held in Bangkok, Thailand, 9–11 August 2026. It offers a premier forum for presenting original research and practical advances in the design, implementation, and application of multimedia information processing and retrieval. The congress convenes university researchers, scientists, industry professionals, software engineers, and graduate students for a four-day program that blends a flagship technical track with keynote speeches, workshops, demonstrations, posters, tutorials, and panels.

The Organizing Committee members of the conference look forward to welcoming you in person in Bangkok in August 2026.

General Co-Chairs:



Paisarn Muneesawang
Mahidol University,
Thailand



Supavadee Aramvith
Chulalongkorn University,
Thailand



Jianquan Liu
NEC Corporation,
Japan



Iole Moccagatta
Intel Corporation,
USA



Ivan Lee
Adelaide University,
Australia

Honorary Co-Chairs:



Shu-Ching Chen
University of Missouri–
Kansas City, USA



Ling Guan
Toronto Metropolitan
University, Canada



Kiyoharu Aizawa
The University of Tokyo,
Japan



Conference Committee

Honorary Co-Chairs

- Shu-Ching Chen, University of Missouri–Kansas City, USA
- Ling Guan, Toronto Metropolitan University, Canada
- Kiyoharu Aizawa, The University of Tokyo, Japan

General Co-Chairs

- Paisarn Muneesawang, Mahidol University, Thailand
- Supavadee Aramvith, Chulalongkorn University, Thailand
- Jianquan Liu, NEC Corporation, Japan
- Iole Moccagatta, Intel Corporation, USA
- Ivan Lee, Adelaide University, Australia

Program Co-Chairs

- Lin Li, Wuhan University of Technology, China
- Toshihiko Yamasaki, The University of Tokyo, Japan
- Yap Kim Hui, Nanyang Technological University, Singapore
- Albert Ali Salah, Utrecht University, the Netherlands
- Pradeep K. Atrey, University at Albany, SUNY, USA

Operation Co-Chair

- Wenyan Hu, AiScholar, China

Award Co-Chairs

- Mei-Ling Shyu, University of Missouri-Kansas City, USA
- Ichiro Ide, Nagoya University, Japan



Conference Committee

Industry Co-Chairs

- Yukiko Yanagawa, OMRON, Japan
- Vishy Swaminathan, Adobe, USA
- Jing Li, Alibaba, China
- Makoto Takamoto, NEC Laboratories Europe, Germany
- Mostafa El-Khamy, Samsung Semiconductor, USA

Local Arrangement Co-Chairs

- Konlakorn Wongpatikaseree, Mahidol University, Thailand
- Mingmanas Sivaraksa, Mahidol University, Thailand

Invited Paper Co-Chair

- Chih-Fan Hsu, Taiwan

Special Session Co-Chairs

- Lei Gao, Wilfrid Laurier University, Canada
- Lai-Kuan Wong, Multimedia University, Malaysia
- Patrick Le Callet, University of Nantes, France
- Jiaying Liu, Peking University, China
- Frode Eika Sandnes, Oslo Metropolitan University, Norway

Workshop Co-Chairs

- Luntian Mou, Beijing University of Technology, China
- Daniela D'Auria, University Mercatorum, Italy
- Jianwei Zhang, Iwate University, Japan
- Chengcui Zhang, University of Alabama at Birmingham, USA



Conference Committee

Tutorial Co-Chairs

- Giuseppe Valenzise, Université Paris-Saclay, France
- Min Chen, University of Washington Bothell, USA
- Wei Hu, Peking University, China
- Matthew Kyan, York University, Canada

Demo Co-Chairs

- Hieu Hanh Le, Ochanomizu University, Japan
- Fabio Persia, University of L'Aquila, Italy
- Dan Zeng, Shanghai University, China
- Yajun Yang, Tianjin University, China
- Mustafa Sert, Başkent University, Turkey

Poster Co-Chairs

- Terumi Umematsu, NEC Corporation, Japan
- Zheng Wang, Wuhan University, China
- Takahiro Komamizu, Nagoya University, Japan

Sponsorship Co-Chair

- Narit Hnoohom, Mahidol University, Thailand

Publicity Co-Chairs

- Guan-Ming Su, Dolby Laboratories, USA
- Baoning Niu, Taiyuan University of Technology, China
- Keiji Yanai, The University of Electro-Communications, Japan
- Ru Wang, Shandong Normal University, China
- Peipei Wang, Qilu University of Technology, China



MIPR 2026

**The 9th IEEE International Conference on
Multimedia Information Processing and Retrieval**

Conference Committee

Publication Co-Chair

- Supatta Viriyavisuthisakul, King Mongkut's University of Technology Thonburi, Thailand

Registration Co-Chair

- Shanshan Yao, Shanxi University, China

Finance Co-Chairs

- Konglit Hunchangsith, Mahidol University, Thailand
- Wei-Bang Chen, Virginia State University, USA

Web & SNS Co-Chairs

- Chakkrit Termritthikun, SGtech, Naresuan University, Thailand
- Kaixi Hu, Wuhan Textile University, China



MIPR 2026

The 9th IEEE International Conference on
Multimedia Information Processing and Retrieval

Organization



Mahidol University

Wisdom of the Land



**IEEE
COMPUTER
SOCIETY**



IEEE

Sponsors

PLATINUM SPONSOR



SILVER SPONSOR



ELSEVIER

BRONZE SPONSOR

Bentley®



Conference Program Overview

Day 1: Sunday, 9 August 2026

Time	Ballroom	Komes 1	Komes 2
09:00-12:30	Workshop 1	Tutorial (10:30-12:30) 90 mins + Q&A/buffer	Workshop 2
12:30-13:30	Lunch Break		
13:30-15:15	Special Session 2	Special Session 1 & 3 (Room sharing)	Demo Session
15:15-15:45	Coffee Break		
15:45-16:15	Opening		
16:15-18:00	Industry Talk Session (3 Talks)		Demo Session
18:00-18:45	Distinguished Plenary Talk		
19:00-21:00	Welcome Reception (Foyer: outside the Ballroom)		

Day 2: Monday, 10 August 2026

Time	Ballroom	Komes 1	Komes 2
09:00-11:00	Regular Oral Session RO-1		Workshop 3 (09:00-11:05)
11:00-11:30	Coffee Break		
11:30-12:30	Keynote Session 1		
12:30-13:30	Lunch Break		
14:00-15:00	Keynote Session 2		
15:00-15:30	Coffee Break		
15:30-17:45	Regular Oral Session RO-2	Regular Poster Session RP-1	Workshop 4
18:30-20:00	Banquet (Jim Thompson, A Thai Restaurant)		



Conference Program Overview

Day 3: Tuesday, 11 August 2026

Time	Ballroom	Komes 1	Komes 2
09:00-10:00	Invited Paper Session IP-1	Invited Paper Session IP-2	Invited Paper Session IP-3
10:00-10:30	Coffee Break		
10:30-11:30	Keynote Session 3		
11:30-12:15	Invited Paper Session IP-1	Invited Paper Session IP-2	Invited Paper Session IP-3
12:30-13:00	Closing & Award Ceremony		
13:00-14:00	Lunch Break		
14:00-18:00	City Tour		



Program Details

Day 1: Sunday, 9 August 2026

Track	Chair	Venue	Time	Note
Workshop 1	Gaurav Bhatnagar	Ballroom	09:00-10:00	4 Papers
Workshop 2	Wanwan Bu	Komes 2	09:00-12:30	IPMML 2026
Tutorial	Iole Moccagatta	Komes 1	10:30-12:30	
Special Session 1	Song Wang	Komes 1	13:30-14:00	2 Papers
Special Session 2	Teng-Wen Chang	Ballroom	13:30-15:15	7 Papers
Special Session 3	Lei Gao	Komes 1	14:00-15:00	4 Papers
Opening	Iole Moccagatta	Ballroom	15:45-16:15	
Industry Talk Session	Yukiko Yanagawa	Ballroom	16:20-17:50	Chenyang Lyu Chiraphat Boonnag Shiv Kumar Saini
Distinguished Plenary Talk	Paisarn Muneesawang	Ballroom	18:00-18:45	
Demo Session	Kim-Hui Yap	Komes 2	13:30-15:15/16:15-18:00	



Program Details

Workshop Session WS-1: Healthcare Through Multimedia Analytics

Chair: Luntian Mou | **Venue:** Ballroom | **Time:** 09:00-10:00, Sunday, 9 August 2026

Time	Slot	Paper ID	Paper Title	Authors
09:00-09:15	WS-1-1	135	Explainable Fundus-Based AI Glaucoma Screening: Design and Evaluation of a Healthcare System	Allan Pascual Erasmo (Yoobee Colleges); Jaya Tiwari (Waikato Hospital); Samaneh Chaharmiri Dokhaharani (Mae Fah Luang University); Mohammad Norouzifard (The University of Auckland)*
09:15-09:30	WS-1-2	162	RadLens: Radiology Report Generation Using a Vision Transformer and Retrieval-Augmented Generation	Philip Niron Nithianandan (Yoobee College of Creative Innovation); Jaya Tiwari (Waikato Hospital); Seyedhashem Davarpanah (Agentic AI Hub); Mohammad Norouzifard (The University of Auckland)*
09:30-09:45	WS-1-3	171	Ethical Implications of AI-Assisted Digital Phenotyping in Clinical Psychiatry	Deepa Shukla (IITJ)*; Nisha Daga (IITJ); George Kodimattam Joseph (IITJ)
09:45-10:00	WS-1-4	172	Neuro-Adaptive Virtual Reality for Spider Phobia: Integrating EEG and Heart Rate Biofeedback in Closed-Loop Exposure Therapy	Rahul Kumar Ray (FLAME University); Rudra Krishna (FLAME University); Chiranjoy Chattopadhyay (FLAME University)*



Program Details

Workshop Session WS-2: 2026 3rd International Conference on Image Processing, Multimedia Technology and Machine Learning

Chair: Wanwan Bu | **Venue:** Komes 2 | **Time:** 09:00-11:00, Sunday, 9 August 2026

Time	Slot	Paper Title	Authors
09:00-09:15	WS-2-1	Key Image Processing Technologies for Intelligent Recruitment Document Parsing and Supply-Demand Prediction in Digital Economy	Ruiqian Su, Xiamen Institute of Technology
09:15-09:30	WS-2-2	Multimodal Feature Fusion with Adaptive Cross-Attention for Attention Economy Propagation Effect Prediction	Yan Zhuang, Xiamen Institute of Technology
09:30-09:45	WS-2-3	Analyzing Facial Expression Dynamics from Naturalistic Video Stimuli for Autism-Oriented Behavioral Assessment	Abigail Copiaco, University of Dubai
09:45-10:00	WS-2-4	Dual-Encoder Model with Interactive Time Horizon for Vehicle Trajectory Prediction in Dense Traffic	Shuaipeng Liu, Southeast University
10:00-10:15	WS-2-5	A Consensus-Based Symbolic-Regression Framework for Cross-Topology SI/PI Prediction of EV Suspension Oil Pump Control Board	ChangSoo Moon, GYEONGBUK RESEARCH INSTITUTE OF VEHICLE EMBEDDED TECHNOLOGY
10:15-10:30	WS-2-6	Infrared and Visible Image Fusion Based on Joint Gradient Collaborative Constraint	Kaixin Gong, Huaiyin Normal University
10:30-10:45	WS-2-7	A signal processing method based on resampling for quasi-constant-speed rotor signals	Wei Zhang, Panzhihua College
10:45-11:00	WS-2-8	MedEvidence-RAG for Joint Evidence Retrieval and Generator Alignment in Medical Visual Question Answering and Report Generation	Xingyu Qu, Harbin University of Science and Technology



Program Details

Tutorial Session

Time	Title	Speaker	Venue
10:30-12:30, Sun. 9 Aug.	Design and Optimization for Interactive Spatial-Temporal Multi-Modal Gaussian Splatting	Dr. Guan-Ming Su (Dolby Laboratories, USA)	Komes 1

Special Session SS-1: Large Models for Embodied Multimedia Processing

Chair: Song Wang | **Venue:** Komes 1 | **Time:** 13:30-14:00, Sunday, 9 August 2026

Time	Slot	Paper ID	Paper Title	Authors
13:30-13:45	SS-1-1	124	Arbiter: Unifying Executable Verification and Semantic Assessment	Yuhang Lu (Zhengzhou university)*; Yong Huang (Zhengzhou university); Ying Yu (The PLA Information Engineering University); Lei Li (The PLA Information Engineering University); Tao Yang (Zhengzhou university); Yun Tie (Zhengzhou university)
13:45-14:00	SS-1-2	129	A Diffusion-Based Data Augmentation Framework for Semantic Segmentation with Limited Data	Zhiyan Wang (Henan University of Technology)*; Xinyu Li (Zhengzhou University); Chunhua Zhu (Henan University of Technology); Lin Qi (Zhengzhou University)



Program Details

Special Session SS-2: Multimodal AI-Mediated Interaction for Cross-Disciplinary Design Collaboration

Chair: Teng-Wen Chang | **Venue:** Ballroom | **Time:** 13:30-15:15, Sunday, 9 August 2026

Time	Slot	Paper ID	Paper Title	Authors
13:30-13:45	SS-2-1	131	The Domain-Specific Fidelity Gap: Trajectory-Based T2I Alignment in AI-Mediated Cross-Disciplinary Design Collaboration	Chen-Syuan Lin (National Yunlin University of Science and Technology); Teng-Wen Chang (National Yunlin University of Science and Technology)*
13:45-14:00	SS-2-2	134	How Do Ideas Evolve Across HumanAI Co-Ideation Cycles? A Trajectory Analysis of Conversational AI in a Semester-Long Interactive-Installation Design Studio	Chor-Kheng Lim (Yuan Ze University)*
14:00-14:15	SS-2-3	128	Surfacing Stress Subtypes through Cognitive-Variation-Conditioned Agentic LLM Co-Ideation: A Walkthrough Probe of Behavioral Trajectories	Siao-Rou Jheng (National Yunlin University of Science and Technology); Ya-Chen Chang (National Yunlin University of Science and Technology); Teng-Wen Chang (National Yunlin University of Science and Technology)*
14:15-14:30	SS-2-4	125	Beyond Semantic Distance: Fusing Physiological and Linguistic Signals for Multimodal Coordination Assessment in Cross-Disciplinary Design	Teng-Wen Chang (National Yunlin University of Science and Technology)*; Yi-Sheng (Ezra) Wu (National Yunlin University of Science and Technology); Ya-Chen Chang (National Yunlin University of Science and Technology); Jia-Rong (Lotus) Li (National Yunlin University of Science and Technology); Siao-Rou Jheng (National Yunlin University of Science and Technology); Wu-Haw Jue (National Formosa University)



Program Details

Time	Slot	Paper ID	Paper Title	Authors
14:30-14:45	SS-2-5	122	Toward AI-Mediated Attention Regulation: A Stability-Oriented Closed-Loop Architecture Using Consumer VR IMU as Wearable Behavioral Sensor	Chia Hui Nico Lo (YuanZe University)*
14:45-15:00	SS-2-6	127	Structured Literary Knowledge as AI-Mediated Generative Mediation: From Cultural Interpretation to Immersive Scene Construction	Tse Wei Hsu (YunTech)*
15:00-15:15	SS-2-7	123	Creatura Era: Multimodal Generative AI and Tangible AR for Cross-Disciplinary Design Collaboration	Wingly Shih (NYUST)*; Kathryn Su (NYUST)



Program Details

Special Session SS-3: From Macro & Micro Expression Analysis to Multimodal Affective Computing: Use Cases, Ethics, and Sector Applications

Chair: Lei Gao | **Venue:** Komes 1 | **Time:** 14:00-15:00, Sunday, 9 August 2026

Time	Slot	Paper ID	Paper Title	Authors
14:00-14:15	SS-3-1	113	Capturing Dynamics: The 4D Facial Expression Intensity Dataset	Zesheng Wang (University of Nantes)*; Alexandre Bruckert (University of Nantes); Pierre Lebreton (University of Nantes); Patrick Le Callet (University of Nantes); Yante Li (University of Oulu); Guoying Zhao (University of Oulu)
14:15-14:30	SS-3-2	170	EmotionBench: Benchmarking Facial Emotion Understanding in Multimodal Large Language Models	Linhan Xu (Shanghai Jiao Tong University); Baiyi Chen (Shanghai Jiao Tong University); Yixuan Gao (Shanghai Jiao Tong University)*; Huiyu Duan (Shanghai Jiao Tong University); Yuqin Cao (Shanghai Jiao Tong University); Sijing Wu (Shanghai Jiao Tong University); Yunhao Li (Shanghai Jiao Tong University); Xiongkuo Min (Shanghai Jiao Tong University); Guangtao Zhai (Shanghai Jiao Tong University)"
14:30-14:45	SS-3-3	140	Towards Emotion-Aware Wearables: Distilling PPG Foundation Models for Real-Time Affective and Physiological Monitoring	Duc Hoang Long Nguyen (University of Engineering and Technology, Vietnam National Universit)*; Patrick Le Callet (Nantes Université, Ecole Centrale Nantes); Thanh Ha Le (University of Engineering and Technology, Vietnam National Universit); Thi Duyen Ngo (University of Engineering and Technology, Vietnam National Universit)
14:45-15:00	SS-3-4	159	Involuntary Leakage, Epistemic Harm, and Regulatory Arbitrage: A Cross-Jurisdictional Analysis of Micro-Expression Governance	Jenjira Jaimunk (Department of Computer Engineering, Faculty of Engineering, Chiang Mai University)*; Patrick Le Callet (Équipe Image, Perception et Interaction, Laboratoire des Sciences du Numérique de Nantes, Université de Nantes); Pradorn Sureephong (College of Arts, Media and Technology, Chiang Mai University)



Program Details

Distinguished Plenary Talk

Time	Chair	Title	Speaker	Venue
18:00-18:45	Paisarn Muneesawang	Learning beyond Deep Learning: Mathematics-Inspired Models for Multi-View Analysis	Prof. Ling Guan (Toronto Metropolitan University, Canada)	Ballroom

Demo Session

Chair: Kim-Hui Yap | **Venue:** Komes 2 | **Time:** 13:30-15:15/16:15-18:00, Sunday, 9 August 2026

Slot	Paper ID	Paper Title	Authors
DS-1-1	110	Multi Agent LLMs for Diversified Large Scale Spatial Consistent 3D Scene Data Synthesis	Sandeep Bangalore (Dolby Labs)*; Lakshmi Priya Muraleedharan (Dolby Labs); Arpitha Nagaraj Hegde (Dolby Labs); Guan-Ming Su (Dolby Labs)
DS-1-2	111	AuthentiQ: Real-Time On-Device Deepfake Detection for the Internet of Multimedia Things	Raghu Iyengar (pi-labs)*; Kapardi Mallampalli (pi-labs); Hitesh Sanghvi (Directorate of Forensic Sciences)
DS-1-3	148	Real-Time Clinical Decision Support in the Trauma Room: An Edge-Native Two-Agent System for Structured Information Processing	Niklas Beck (Fraunhofer IAIS)*; Johann Jasper Schulze Buschhoff (Fraunhofer IAIS); Emmy Lai (Fraunhofer IAIS); Najmeh Mousavi (Fraunhofer IAIS); Julian Maluck (T-Systems International); Teh-Hai Julian Zheng (T-Systems International); Stefan Swoboda (T-Systems International); Benny Jörg Stein (Fraunhofer IAIS)
DS-1-4	151	Personalized Federated Fine-Tuning for Distributed Image-Text Retrieval	Yifan Zhou (shanghai university); Mengxin Yang (shanghai university); Hongmei Shi (shanghai university); Yixiao Gu (shanghai university)*
DS-1-5	152	BETA-Cap: An Interactive System for Bilingual English-Turkish Audio Captioning	Melek Tanyeri Yener (Baskent University)*; Mustafa Sert (Baskent University)
DS-1-6	169	Harness Engineering for Multi-Agent Data Visualization with Small Language Models	Xinxiao Li (Shonan Institute of Technology); Peilin Wang (Shonan Institute of Technology); Eisei Nakahara (Nihon University)*
DS-1-7	178	Designing an Intelligent Multimedia Vending Interface for Ethical-Choice Interaction	Wing Yan Jasman PANG (Technological and Higher Education Institute of Hong Kong)*; Yan Chi Wan (Technological and Higher Education Institute of Hong Kong)



Program Details

Day 2: Monday, 10 August 2026

Keynote Sessions

Time	Chair	Title	Speaker	Venue
11:30-12:30	Albert Ali Salah	Towards Trustworthy Multimodal Models	Prof. Mohan Kankanhalli (National University of Singapore, Singapore)	Ballroom
14:00-15:00	Supavadee Aramvith	Stable Components in High-Dimensional Data: Joint Decomposition Across Datasets	Prof. Tülay Adalı (University of Maryland, Baltimore County, USA)	Ballroom

Regular Sessions

Session	Chair	Venue	Time	Note
RO-1	Toshihiko Yamasaki	Ballroom	09:00-10:30	6 Papers
RO-2	Albert Ali Salah	Ballroom	15:30-17:15	7 Papers
RP-1	Shanshan Yao	Komes 1	15:30-17:45	24 Papers



Program Details

Regular Oral Session RO-1: Multimodal Perception, Reasoning, and Intelligent Agents

Chair: Toshihiko Yamasaki | **Venue:** Ballroom | **Time:** 09:00-10:30, Monday, 10 August 2026

Time	Slot	Paper ID	Paper Title	Authors
09:00-09:15	RO-1-1	17	Single-View Biomechanical State Estimation for Maize Architecture Recovery	Yue Zhuo (Cornell University); Fengqi You (Cornell University)*
09:15-09:30	RO-1-2	18	3D Object Recovery for Data Tampering Attack on In-vehicle Networks	Gangtao Han (Zhengzhou University); Yurui Chen (Zhengzhou University); Song Wang (Zhengzhou University)*; Enqing Chen (Zhengzhou University); Lingling Li (Zhengzhou University of Aeronautics); Gaofeng Pan (Beijing Institute of Technology)
09:30-09:45	RO-1-3	45	Boosting Vision-Language Model for Cross-Device GUI Agent Grounding and Navigation 🏆 Best Student Paper Candidate	Bo-Yu Chen (National Chung Cheng University); Liang-Han Shih (National Chung Cheng University); Chen-Kuo Chiang (National Chung Cheng University)*
09:45-10:00	RO-1-4	47	SpatialAgent: Spatial Question Answering with LLM Agent and Perception Models 🏆 Best Student Paper Candidate	Hsiang-Wei Huang (University of Washington)*; Junbin Lu (University of Washington); Jianxu Shanguan (University of Washington); Jen-Hao Cheng (University of Washington); Kuang-Ming Chen (University of Washington); Cheng-Yen Yang (University of Washington); Bahaa Alattar (University of Washington); Yi-Ru Lin (University of Washington); Pyongkun Kim (Electronics and Telecommunications Research Institute); Sangwon Kim (Electronics and Telecommunications Research Institute); Kwangju Kim (Electronics and Telecommunications Research Institute); Chung-I Huang (National Chung Hsing University); Jenq-Neng Hwang (University of Washington)
10:00-10:15	RO-1-5	63	Document-Grounded Coaching Agent for Video Skill Assessment 🏆 Best Paper Candidate	Masaya Fujiwaka (NEC)*; Takuro Kawada (Hosei University); Xiaotong Ji (NEC); Jianquan Liu (NEC)
10:15-10:30	RO-1-6	84	Trajectory-Centered Reasoning for Counterfactual Driving Question Answering 🏆 Best Paper Candidate	Xiaotong Ji (NEC Corporation)*; Masaya Fujiwaka (NEC Corporation); Jianquan Liu (NEC Corporation)



Program Details

Regular Oral Session RO-2: Multimedia Intelligence and Learning for Domain-Specific Applications

Chair: Albert Ali Salah | **Venue:** Ballroom | **Time:** 15:30-17:15, Monday, 10 August 2026

Time	Slot	Paper ID	Paper Title	Authors
15:30-15:45	RO-2-1	3	Bridging Design and Quality: A Technical Overview of FMs in Manufacturing	Chih-Fan Hsu (Inventec Corp.)*; Yung-Huei Chiou (Inventec Corp.); Po-Han Huang (Inventec Corp.); Yu-chi Liao (Inventec Corp.); Shu Huei Yang (Inventec Corp.); Jyun-Ze Tang (Inventec Corp.); Yu-Lun Chang (Inventec Corp.); Ming-Ching Chang (University at Albany)
15:45-16:00	RO-2-2	24	QuranRTT: Real-Time Domain-Specific Translation of Classical Arabic Quranic Recitation	Nagendhra Madishetti (University at Albany, State University of New York); Md Monjurul Ahsan Prodhan (University at Albany, State University of New York); Md Nour Hossain (University at Albany, State University of New York)*
16:00-16:15	RO-2-3	36	Hard-Decision Bottleneck in Belief Propagation for Speaker Diarization: A Factor Graph Analysis	Snit Sanghlao (Faculty of Information and Communication Technology, Mahidol University)*; Jantarat Saenghlao (Department of Computer Engineering, KMUTT); Threrawee Sanglaow (Department of Physics, KMUTT)
16:15-16:30	RO-2-4	53	Continual Learning with Support Boundary Experience Blending 🏆 Best Paper Candidate	Chih-Fan Hsu*; Yung-Huei Chiou (Inventec Corp.); Ming-Ching Chang (University at Albany, State University of New York); Wei-Chao Chen (Inventec Corp.)
16:30-16:45	RO-2-5	62	Modeling Video Ads with Structured Semantics for CTR Prediction 🏆 Best Student Paper Candidate	Yoshiaki Tanabe (The university of Tokyo)*; Shuntaro Masuda (The university of Tokyo); Gakumatsu Ryu (Septeni Japan, Inc.); Ryoichi Katsuya (Septeni Japan, Inc.); Toshihiko Yamasaki (The university of Tokyo)
16:45-17:00	RO-2-6	100	VCE-DINO: Robust Video Capsule Endoscopy Anatomy Recognition via DINO Pretraining and Imbalance-Aware Learning	Chayapon Sasnarukkit (Chulalongkorn university); Natawut Nupairoj (Chulalongkorn university); Peerapon Vateekul (Chulalongkorn university)*
16:45-17:15	RO-2-7	40	Bi-projection Fusion for 360 Depth Estimation Based on Coordinate Attention	Lei Li (Information Engineering University);Chenyang He (Zhengzhou University);Ying Yu (Information Engineering University);Tao Yang (Zhengzhou University);Yuhang Lu (Zhengzhou University);Yun Tie (Zhengzhou University)*



Program Details

Regular Poster Session

Chair: Shanshan Yao | **Venue:** Komes 1 | **Time:** 15:30-17:45, Monday, 10 August 2026

Slot	Paper ID	Paper Title	Authors
RP-1-1	5	Optimized Smart Home Energy Management Systems Based on Hybrid Method	Alexander N. Ndife (Chalmers University of Technology)*; Arinze Ijeoma (Institut für Kernphysik University of Münster); Paisarn Muneesawang (Department of Computer Engineering Mahidol University)
RP-1-2	16	A Rotation-Guided and Feature-Refined Framework for Robust Nuclei Segmentation in Histopathology	Arbab Sufyan Wadood (Multimedia University)*; Mohammad Faizal Ahmad Fauzi (KPJU Healthcare University); Lai Kuan Wong (Multimedia University); Md Serajun Nabi (KPJU Healthcare University)
RP-1-3	20	Hierarchical Federated Class-Incremental Learning with Client-Local Replay for Industrial IoT Intrusion Detection	Waiteng CHING (University of Aizu)*; Xiuheng LIAO (University of Aizu); Jingcheng SONG (Linyi University); Weiyu Wang (Hosei University); Zhiwei Si (University of Aizu); Chunhua SU (University of Aizu)
RP-1-4	25	Multi Local Modeling: Toward Structured Symbolic Music Generation Via Local Learning	Van Hoang Bao Bach (University of Central Florida)*; Sansiri Tarnpradab (King Mongkut's University of Technology Thonburi); Kien Hua (University of Central Florida)
RP-1-5	28	DeepVisionSafety: An intelligent CCTV Monitoring System for Multi-class Anomaly Recognition	Napasrapee Satittham (Department of Computer Engineering, Faculty of Engineering, Mahidol University, Nakhon Pathom, Thailand.)*; Purin Pongpanich (Department of Computer Engineering, Faculty of Engineering, Mahidol University, Nakhon Pathom, Thailand.); Kunanont Vutiprasertwong (Department of Computer Engineering, Faculty of Engineering, Mahidol University, Nakhon Pathom, Thailand.); Narit Hnoohom (Department of Computer Engineering, Faculty of Engineering, Mahidol University, Nakhon Pathom, Thailand.)
RP-1-6	32	A Distributed Learning Approach for Controlled Grammar Transfer using Structure-Timbre Disentangled Audio Synthesis	Sridhar Varma Kalidindi (National Institute of Technology Andhra Pradesh); Karthick S (National Institute of Technology Tiruchirappalli)*



Program Details

Regular Poster Session

Chair: Chair: Shanshan Yao | **Venue:** Komes 1 | **Time:** 15:30-17:45, Monday, 10 August 2026

Slot	Paper ID	Paper Title	Authors
RP-1-7	33	Attention-Enhanced EfficientNet with Anatomy-Aware Test-Time Augmentation for Robust Dental Caries Detection	Phubase Sangliamthong (Faculty of Information and Communication Technology, Mahidol University); Aroonrat Choochue (Faculty of Information and Communication Technology, Mahidol University); Siriwan Suebnukarn (Faculty of Dentistry, Thammasat University); Kanoksak Wattanachote (Mahidol University)*
RP-1-8	41	Language-Conditioned Geographic Localization of Indic Speech	Raghu Iyengar (pi-labs)*; Hitesh Sanghvi (Directorate of Forensic Science)
RP-1-9	46	Detection-based Spatial Token Selection for Long Video Understanding with MLLMs	Yuta Goto (NEC Corporation)*; Kengo Ishida (NEC Corporation); Satoshi Yamazaki (NEC Corporation); Junichi Funada (NEC Corporation); Jianquan Liu (NEC Corporation)
RP-1-10	57	Extreme Learning Machine Ensemble for Deep Learning-based List Decoding in Error-prone Video Transmission Systems	Anton Louise De Ocampo (Batangas State University)*; Francois-Xavier Coudoux (Institut d'Électronique de Microélectronique et de Nanotechnologie (IEMN)); Stéphane Coulombe (École de technologie supérieure (Université du Québec)); Patrick Corlay (Université Polytechnique Hauts-de-France)
RP-1-11	61	Convolution and Transformer Integration for Glaucoma Classification on Enhanced Fundus Images	Chayawat Nilsumrit (Mahidol University); Narit Hnoohom (Mahidol University)
RP-1-12	66	An Empirical Study of Model-Specific Preprocessing and Focal Loss for Thai Depression Text Classification	Rawipas Samoonddee (Chiang Mai University)*; Dussadee Praserttitipong (Chiang Mai University)



Program Details

Regular Poster Session

Chair: Shanshan Yao | **Venue:** Komes 1 | **Time:** 15:30-17:45, Monday, 10 August 2026

Slot	Paper ID	Paper Title	Authors
RP-1-13	70	Real-Time Video Anomaly Detection Using YOLO Pose Estimation and CLIP-Based Semantic Scoring	Vanodhya Warnasooriya (Chulalongkorn University); Amir Hajian (Chulalongkorn University); Watchara Ruangsang (King Mongkut's University of Technology Thonburi); Supavadee Aramvith (Chulalongkorn University)*
RP-1-14	71	CARRE: Comprehensive Automated Refinement for Real Estate Using Large Language Models	Yuta Takatsuji (The University of Tokyo)*; Tatsuya Iwanari (estie, inc.); Ryosuke Katsuta (estie, inc.); Shuntaro Masuda (The University of Tokyo); Shengzhou Yi (The University of Tokyo); Toshihiko Yamasaki (The University of Tokyo)
RP-1-15	73	Discriminative Correlation Perceptron-Style Neural Network for Multi-View Glaucoma Image Classification	Chakkrit Termritthikun (Naresuan University)*; Yirga Belay Muna (Naresuan University); Paisarn Muneesawang (Mahidol University); Patchara Suwanbordin (Mahidol University); Lei Gao (Wilfrid Laurier University); Kai Liu (Toronto Metropolitan University); Ling Guan (Toronto Metropolitan University)
RP-1-16	77	SAM3 Adaptation for Root CT Image Segmentation	Hurriyatul Fitriyah (Adelaide University)*; Huafei Huang (University of South Australia); Ivan Lee (University of South Australia)
RP-1-17	82	Beyond Global and Local: A Stacked Ensemble Approach to Symbolic Music Emotion Recognition	Hema Kanoujia (University of Central Florida)*; Hemant Kanoujia (University of Central Florida); Kien Hua (University of Central Florida)
RP-1-18	83	Depth-Aware Pothole Detection Using YOLO and RT-DETR at the Edge	Md Monjurul Prodhan (University at Albany, State University of New York); Md Nour Hossain (University at Albany, State University of New York)*



Program Details

Regular Poster Session

Chair: Chair: Shanshan Yao | **Venue:** Komes 1 | **Time:** 15:30-17:45, Monday, 10 August 2026

Slot	Paper ID	Paper Title	Authors
RP-1-19	87	Evaluating Diffusion Models for Single-Image 3D Gaussian Splatting Scene Reconstruction	Max Eskandari (York University); Xingbang Tang (York University)*; Matthew J. Kyan (York University)
RP-1-20	88	A Harmless Owner-Specific Verification Scheme for DNN Copyright Protection	Jiangnan Zhu (Kyushu University)*; Yuzhang Chen (Kyushu University); Yujie Gu (Kyushu University); Minoru Kuribayashi (Tohoku University); Kouichi Sakurai (Kyushu University)
RP-1-21	98	Exploring LMM-as-a-Judge for Image Harmonization Evaluation	Jeonghun Baek (The University of Tokyo)*; Eunchung Noh (Samsung Electronics)
RP-1-22	101	Evaluating Symptom Fidelity in LLM-Driven Simulated Patients Using PHQ-9 and GAD-7	Pathanin Opachalearn (Mahidol University)*; Konlakorn Wongpatikaseree (Mahidol University); Natthawut Kertkeidkachorn (Japan Advanced Institute of Science and Technology)
RP-1-23	105	Characterizing Spatial Interface Metaphors for Exploratory 360° Video Retrieval in VR	Mike Knoop (Utrecht University); Wolfgang Hürst (Utrecht University)*
RP-1-24	107	Improving GPT-Based Subjective Floor Plan Evaluation with Structural Cue Correction	Taro Narahara (New Jersey Institute of Technology)*; Toshihiko Yamasaki (The University of Tokyo)



Program Details

Workshop Sessions

Track	Chair	Venue	Time	Note
Workshop 3	Luntian Mou and Zichun Guo	Komes 2	09:00-11:05, Mon. 10 Aug.	3 Keynotes and 5 Papers

Workshop Session WS-3: AI for Culture, Art, and Design

Chair: Luntian Mou | **Venue:** Komes 2 | **Time:**09:00-11:05, Monday, 10 August 2026

Time	Slot	Paper ID	Paper Title	Authors
09:00-09:25	WS-3-1	Keynote 1	Provenance Computation of Historical and Cultural Data	Xiaoru Yuan (Peking University)
09:25-09:35	WS-3-2	144	A Value-Aligned Multimodal Agent for Culturally Grounded Textbook Illustration Generation and Evaluation	Chenrui Qin (East China University of Science and Technology); Yumiao Chen (East China University of Science and Technology)*; Yifan Yan (East China University of Science and Technology)
09:35-09:45	WS-3-3	145	SensiPillow: An Embodied Multimodal Agent for Bedtime Smartphone Intervention	Qi Fang (Southern University of Science and Technology); Qianwen Tan (Southern University of Science and Technology); Kai Lin (Southern University of Science and Technology); Ziqian Bai (Southern University of Science and Technology)*
09:45-10:10	WS-3-4	Keynote 2	Music Large Model Shao 1.0	Jiafeng Liu (Central Conservatory of Music)
10:10-10:20	WS-3-5	175	Phyigital Symbiosis and Innovation Dynamics of Location-Based Entertainment in China	Jie Pang (Peking University)*; Jiaxin Li (Peking University)
10:20-10:30	WS-3-6	176	EAC2-Curation:Embodied Attention Capture and Curatorial Reconstruction	Jiaxin Li(Peking University)*; Jie Pang(Peking university); Xinyi Wang(Zhejiang University)
10:30-10:55	WS-3-7	Keynote 3	Jewelry and Theater: Design and Artistic Practice Spanning Physical and Digital Realities	Ting Song (Ting Museum)
10:55-11:05	WS-3-8	180	From Viewing to Experiencing: An Embodied Narrative Study of Dunhuang Mural Content	Xianning Meng (Daegu University)*; Zichun Guo (Beijing University of Chemical Technology); Qiang Hao (Tsinghua University); Chang Pan (Daegu University)



Program Details

Day 3: Tuesday, 11 August 2026

Keynote Sessions

Time	Chair	Title	Speaker	Venue
10:30-11:30	Iole Moccagatta	Journey to the Next Era: AI-driven Multi-view Analysis and the Future of Spatial Intelligence	Prof. Hideo Saito (Keio University, Japan)	Ballroom

Invited Paper Sessions

Session	Chair	Venue	Time	Note
IP-1	Mei-Chen Yeh	Ballroom	09:00–10:00 / 11:30–12:15	4/3 papers
IP-2	Guan-Ming Su	Komes 1	09:00–10:00 / 11:30–12:15	4/3 papers
IP-3	Chih-Fan Hsu	Komes 2	09:00–10:00 / 11:30–12:15	4/3 Papers

Invited Paper Session IP-1: Applications

Chair: Mei-Chen Yeh | **Venue:** Ballroom | **Time:** 09:00-10:00, Tuesday, 11 August 2026

Time	Slot	Paper ID	Paper Title	Authors
09:00-09:15	IP-1-1	158	Designing Storyboards for Immersive Video Retrieval: A Comparative Study of 360- Degree Video Visualizations	Marloes van der Haagen (Utrecht University); Wolfgang Hürst (Utrecht University)*
09:15-09:30	IP-1-2	109	Self-Supervised Representation Learning Using Object-Aware Data Augmentation	Yong-Hong Lin (National Taiwan Normal University); Mei-Chen Yeh (National Taiwan Normal University)*
09:30-09:45	IP-1-3	173	Force-Responsive Virtual Makeup	Ning Xu (Adeia Inc.)*; Aldis Sipolins (Adeia Inc.); Mareeta Mathai (Adeia Inc.); Mathew Adams (Adeia Inc.); Serhad Doken (Adeia Inc.)
09:45-10:00	IP-1-4	117	Transformer-Based Local Path Planning for Quadruped Robots Using Multi-View Semantic Perception	Ju Wang (Virginia State University)*; Wei-Bang Chen (Virginia State University); Qian Qu (Virginia State University); Xuanchen Xiang (Virginia State University); Ahmed Mohammed (Virginia State University)



Program Details

Invited Paper Session IP-1: Applications

Chair: Mei-Chen Yeh | **Venue:** Ballroom | **Time:** 11:30-12:15, Tuesday, 11 August 2026

Time	Slot	Paper ID	Paper Title	Authors
11:30-11:45	IP-1-5	133	Super-Resolution Preprocessing for Small Object Detection: An Empirical Study on VisDrone-Det at $\times 2$ and $\times 4$ Scales	Musa Yahaya (Chulalongkorn University)*; Amir Hajian (Chulalongkorn University); Kanapoj Ngambenjavichaikul (Chulalongkorn University); Watchara Ruangsang (King Mongkut's University of Technology Thonburi); Supavadee Aramvith (Chulalongkorn University)
11:45-12:00	IP-1-6	155	A Lightweight Speaker Recognition Method with Multi-Layer Self-Distillation	Shanshan Yao (Shanxi University); Jing You (Shanxi University); Yaozhang Zhu (Shanxi Lu'an Mining (Group) Co., Ltd.); Baoning Niu (Taiyuan University of Technology)*
12:00-12:15	IP-1-7	149	Physics-Guided Cascade Classification for Multi-Granularity Liquid Recognition in Dual-Energy X-ray Security Inspection	Hongping Zhi (Tsinghua University)*; Lifeng Sun (Tsinghua University); Xiaoyun Su (Suzhou ZMore Technology Co., Ltd.); Weiwei Zheng (Suzhou ZMore Technology Co., Ltd.)



Program Details

Invited Paper Session IP-2: Content Understanding

Chair: Guan-Ming Su | **Venue:** Komes 1 | **Time:** 09:00-10:00, Tuesday, 11 August 2026

Time	Slot	Paper ID	Paper Title	Authors
09:00-09:15	IP-2-1	86	Global Attribution Meets Local Representation: A Hybrid Pipeline for Binary Composer Classification	Hemant Kanoujia (University of Central Florida)*; Hema Kanoujia (University of Central Florida); Kien Hua (University of Central Florida)
09:15-09:30	IP-2-2	116	Anchor Tracking and Prediction with Triplane Channel-to-Spatial Mapping for Dynamic Gaussian Splatting Compression	Birendra Kathariya (Dolby)*; Tsung-Wei Huang (Dolby Laboratories Inc.); Dae Yeol Lee (Dolby Laboratories Inc.); Fangjun Pu (Dolby Laboratories Inc.); Guan-Ming Su (Dolby Laboratories Inc.); Peng Yin (Dolby Laboratories Inc.); Gary J. Sullivan (Dolby Laboratories Inc.)
09:30-09:45	IP-2-3	160	Capture-Stage Multi-Camera Placement Recommendation System for Improved 3D Scene Coverage	Dae Yeol Lee (Dolby Laboratories)*; Sandeep Bangalore Venkatesh (Dolby Laboratories); Anshul Rai (Dolby Laboratories); Tsung-Wei Huang (Dolby Laboratories); Guan-Ming Su (Dolby Laboratories)
09:45-10:00	IP-2-4	174	Dynamic Range Reshaping for Gaussian Splatting	Tsung-Wei Huang (Dolby Laboratories)*; Dae Yeol Lee (Dolby Laboratories); Birendra Kathariya (Dolby Laboratories); Guan-Ming Su (Dolby Laboratories); Peng Yin (Dolby Laboratories)

Chair: Guan-Ming Su | **Venue:** Komes 1 | **Time:** 11:30-12:15, Tuesday, 11 August 2026

Time	Slot	Paper ID	Paper Title	Authors
11:30-11:45	IP-2-5	163	Origin–Destination Proportion Analysis Using Machine Learning on Non-Overlapping Camera Videos	Sajan Kumar (Purdue University); Stanley Chien (Purdue University)*; Shail Desai (Purdue University); Yaobin Chen (Purdue University); Nathan Shellhamer (Indiana Department of Transportation)
11:45-12:00	IP-2-6	164	Open-set Visual Template Discovery in Large-Scale Propaganda Image Collections	Minh Tu Nguyen (National Chung Cheng University); Yu-Yao Tseng (National Chung Cheng University); Ming-Hung Wang (National Chung Cheng University)*
12:00-12:15	IP-2-7	154	Emotional Plants: When Artificial Agents Report Their Internal Affective States	Vito Vekic (Utrecht University); Albert Ali Salah (Utrecht University)*; Almila Akdag (Utrecht University)



Program Details

Invited Paper Session IP-3: Machine Learning, Deep Learning, Foundation Model, and Trustworthy AI

Chair: Chih-Fan Hsu | **Venue:** Komes 2 | **Time:** 09:00-10:00, Tuesday, 11 August 2026

Time	Slot	Paper ID	Paper Title	Authors
09:00-09:15	IP-3-1	91	Lesion-DDPM: Lesion-Enhanced 3D Diffusion for MS MRI Synthesis	Weidong Zhang (Northeastern University); Yongchan Jung (Fairleigh Dickinson University); Shafayat Mowla Anik (University of Colorado - Colorado Springs); Furen Xiao (National Taiwan University); Vasudevan Janarthanan (Fairleigh Dickinson University); Enkhzaya Chuluunbaatar (University of British Columbia); Marco Ho (British Columbia Institute of Technology); Byeong Kil Lee (University of Colorado - Colorado Springs); Jeeho Ryoo (Fairleigh Dickinson University)*
09:15-09:30	IP-3-2	156	DESNet: A Multimodal Collaborative Network for Joint Driver Emotion and Stress Recognition	Luntian Mou (Beijing University of Technology)*; Wenyan Zhao (Beijing University of Technology); Shasha Mao (Xidian University); Nan Ma (Beijing University of Technology)
09:30-09:45	IP-3-3	167	Edge Computing with Green Super-Resolution	Li-Heng Wang (USC)*; Yixing Wu (USC); Wei Wang (USC); C.-C. Jay Kuo (USC)
09:45-10:00	IP-3-4	76	A Discriminant Multi-view Correlation Analysis Platform for Multimedia Computing	Lei Gao (Wilfrid Laurier University)*; Kai Liu (Toronto Metropolitan University); Ling Guan (Toronto Metropolitan University)



Program Details

Invited Paper Session IP-3: Machine Learning, Deep Learning, Foundation Model, and Trustworthy AI

Chair: Chih-Fan Hsu | Venue: Komes 2 | Time: 09:00-10:00, Tuesday, 11 August 2026

Time	Slot	Paper ID	Paper Title	Authors
11:30-11:45	IP-3-5	161	A Multi-Source Ultrasound Benchmark Revealing the Limits of Contemporary Self-Supervised Anomaly Detection Methods	Marco Riedenauer (University of Augsburg)*; Daniel Kienzle (University of Augsburg); Pratik Mayekar (University of Augsburg); Rainer Lienhart (University of Augsburg)
11:45-12:00	IP-3-6	139	Comparison of multimodal LLM embedding models in the task of cross- modal recipe retrieval	Keisuke Gomi (The University of Electro- Communications)*; Keiji Yanai (The University of Electro-Communications)
12:00-12:15	IP-3-7	115	Differentiable Proxy Learning for Adaptive Quantization Control in H.264 Video Coding	Qihan Xu (Simon Fraser University)*; Ivan Bajic (Simon Fraser University)



Distinguished Plenary Talk



Prof. Ling Guan

Toronto Metropolitan University, Canada

IEEE Life Fellow

CAE Fellow

EIC Fellow

IEEE Canada C.C.
Gotlieb Medal

Title: Learning beyond Deep Learning: Mathematics-Inspired Models for Multi-View Analysis

Abstract:

Acknowledging the tremendous contributions deep learning (DL) made to a broad range of image, video and multimedia processing tasks, further optimizing quality of the DL-based features has consistently presented a challenge, especially when working with multi-view data, arguably the most natural data format. Though DL models have been able to extract high quality features, jointly processing these features is largely confined to late fusion (score level and decision level) when information has been condensed into a few numbers, thus limiting quality of performance. In this talk, we present recent development from a significantly different perspective, a mathematics-inspired model; an approach belonging to the cooperative learning paradigm. Particularly, we call upon a natural multi-view processing architecture, discriminant correlation analysis, setting the stage for the development of an innovative platform. Due to its power to handle information with multiple views, the platform is termed as discriminant multiple correlation (DMC) analysis. Depending on the nature of the data sources to work with, DMC features two distinct designs, one with a perceptron-style NN (PNN) to handle the nonlinear processing part, and the other with a convolution-style NN (CNN). Statistics collected from experiments of numerous multi-view analysis and recognition benchmarks evidently show that the MI models generate impressive (sometimes unprecedented) performance accuracies.



Distinguished Plenary Talk

Biography:

Dr. Ling Guan received his Ph.D. Degree from University of British Columbia, Canada. From 1992 to 2001, he was on the Faculty of Engineering, University of Sydney, Australia. Dr. Guan served as a Tier I Canada Research Chair in Multimedia and Computer Technology and a Professor of Electrical and Computer Engineering at Ryerson University/Toronto Metropolitan University, Toronto, Canada, as well as the founding director of Ryerson/TMU Multimedia Research Lab. Dr. Guan has been working on image, video and multimedia signal processing, human-computer interaction, pattern recognition and machine intelligence, and published extensively in the field.

Recently, Dr. C.-C. Jay Kuo of University of Southern California and Dr. Guan jointly initiated Learning beyond Deep Learning to pave the way for new machine learning research which has drawn the attention of many colleagues in academia and industry. Dr. Guan served on half a dozen editorial boards of IEEE Transactions and Magazines and chaired the 2006 IEEE International Conference on Multimedia and Expo in Toronto. He played the leading role in the inauguration of IEEE Pacific-Rim Conference on Multimedia in 2000 and served as the Founding General Chair. Dr. Guan is a Life Fellow of the IEEE, a Fellow of the Canadian Academy of Engineering and the Engineering Institute of Canada. He is an IEEE Circuits and System Society Distinguished Lecturer and a recipient of the 2005 IEEE Transactions on Circuits and Systems for Video Technology Best Paper Award. Dr. Guan is the awardee of the 2014 IEEE Canada C.C. Gotlieb Medel for his contributions to computer science and engineering.



Keynote Speech I



Prof. Hideo Saito

Keio University, Japan

IEICE Fellow

VRSJ Fellow

President of ITE

Former President
of IEICE ISS

Title: Journey to the Next Era: AI-driven Multi-view Analysis and the Future of Spatial Intelligence

Abstract:

Since 1997, our exploration has been fueled by the ambition to reconstruct 3D worlds from large-scale multi-view camera arrays, giving birth to “free-viewpoint video.” For over two decades, we have refined the foundational pillars of multi-view geometry, such as camera calibration and SLAM. However, we are now entering a transformative “Next Era” where the synergy between these traditional methodologies and modern AI is redefining the boundaries of what is possible.

In this talk, I will present our journey as it evolves from early 50-camera setups to the cutting edge of AI-driven 3D vision. I will discuss how recent breakthroughs in neural rendering and generative video intelligence are unleashing the true potential of multi-view systems—moving beyond mere geometric reconstruction toward deep “Spatial Intelligence.”

While these technologies have found impactful applications in fields like surgical scene analysis and Medical AR, their potential extends far beyond, promising a future where AI can perceive, understand, and interact with complex 3D environments in real-time. I will introduce our latest research projects and share my excitement for this new chapter in computer vision, where multi-view analysis serves as the cornerstone for the next generation of intelligent systems.



Keynote Speech I

Biography:

Hideo Saito received his Ph.D. in Electrical Engineering from Keio University, Japan, in 1992. He has been a faculty member at Keio University since then and has served as a Full Professor since 2006. From 1997 to 1999, he was a Visiting Researcher at the Robotics Institute, Carnegie Mellon University, contributing to the pioneering "Virtualized Reality" project.

He currently holds key academic leadership positions at Keio University, serving as the Chair of the School of Informatics, Management, and Human Sciences in the Graduate School of Science and Technology, as well as a Group Leader at the AI Computing and Research Center within the Keio University Global Research Institute (KGRI).

His research expertise lies in 3D computer vision and scene analysis, with a current focus on the synergy between multi-view geometry and modern AI technologies. His work spans a wide range of high-impact applications, including Augmented Reality (AR), Mixed Reality (MR), and advanced surgical scene analysis.

Dr. Saito has a distinguished record of leadership in the global academic community, having served as Program Chair for ACCV 2014, General Chair for ISMAR 2015, Program Chair for ISMAR 2016, and Scientific Program Chair for Euro VR 2020. He is a Fellow of the Institute of Electronics, Information and Communication Engineers (IEICE) and the Virtual Reality Society of Japan (VRSJ).

In addition to his university leadership, he currently serves as the President of the Institute of Image Information and Television Engineers (ITE) and is the former President of the Information and Systems Society of IEICE. His extensive contributions have been recognized with numerous accolades, including the prestigious Best Journal Paper Award at ISMAR 2023.



MIPR 2026

The 9th IEEE International Conference on
Multimedia Information Processing and Retrieval

Keynote Speech II



Prof. Mohan Kankanhalli

National University of Singapore, Singapore

ACM Fellow

IEEE Fellow

IAPR Fellow

SNAS Fellow

Deputy Executive Chairman of AI Singapore

Director of NUS AI Institute

Title: Towards Trustworthy Multimodal Models

Abstract:

Large Language Models and Vision Language Models have seen dramatic improvements in capability over the last three years. There are multiple ongoing efforts to further enhance the capability of these models. Despite their continually improving performance on a variety of benchmarks, there are still some fundamental trust issues that hinder the wide-scale real-world deployment of these powerful multimodal models, especially in safety critical applications such as healthcare and finance. This talk will present some of the recent work done in our group at NUS that highlight these issues. We will cover the areas of bias, hallucinations and robustness. We will end the talk by detailing some of the challenges and the open research questions in the field.



Keynote Speech II

Biography:

Mohan Kankanhalli is Provost's Chair Professor of Computer Science at the National University of Singapore (NUS), where he is the Founding Director of the NUS AI Institute. He is also the Deputy Executive Chairman (Talent) of AI Singapore, which is Singapore's National AI R&D Program. He has held senior leadership roles as the Dean of NUS School of Computing and Vice Provost of Graduate Education for NUS. He obtained his BTech from IIT Kharagpur and MS & PhD from the Rensselaer Polytechnic Institute. Mohan's research interests are in Multimodal Computing, Computer Vision, and Trustworthy AI. He has made fundamental contributions in the areas of image and video understanding, visual saliency and trustworthy AI. He is the Associate Editor-in-Chief of IEEE Multimedia magazine. He is on the editorial boards of Springer Multimedia Systems and Springer Journal on Big Data. Mohan was a member of World Economic Forum's 2023-2024 Global Future Council on Artificial Intelligence. He is currently a member of ACM's Global Technology Policy Council. He is also on the EY Global AI Advisory Council. He is a Fellow of IEEE, SNAS, IAPR and ACM.



Keynote Speech III



Prof. Tülay Adalı

University of Maryland, Baltimore County (UMBC), USA

IEEE Fellow

AIMBE Fellow

AAIA Fellow

Fulbright Scholar

EiC of IEEE SPM

Title: Stable Components in High-Dimensional Data: Joint Decomposition Across Datasets

Abstract:

In many fields today, such as neuroscience, remote sensing, computational social science, and the physical sciences, multiple datasets are readily available. Matrix and tensor factorizations provide a principled framework for joint analysis, enabling the fusion of such datasets so that they can interact and inform each other while making relatively modest assumptions about their underlying relationships. A central motivation for these methods is the recovery of components that are directly interpretable and, in favorable settings, comparable across datasets, subjects, or measurement sessions. This naturally raises questions of reproducibility and stability of the estimated components, particularly in relation to more flexible representation learning approaches.

This talk presents an overview of matrix and tensor factorization models, with emphasis on independent component analysis and its multi-dataset extension, independent vector analysis. We discuss how these approaches identify a directly interpretable latent structure, and how resulting components can be analyzed across modalities and observations. Examples from multi-subject neuroimaging data illustrate how recurring components can be identified across individuals and used in applications such as subgroup discovery for precision medicine. We also briefly contrast these properties with those of modern deep learning methods in terms of interpretability and identifiability, and highlight remaining challenges related to robustness, model selection, and practical reproducibility.



Keynote Speech III

Biography:

Tülay Adali received the Ph.D. degree in Electrical Engineering from North Carolina State University, Raleigh, NC, USA, in 1992 and joined the faculty at the University of Maryland Baltimore County (UMBC), Baltimore, MD, the same year. She is currently a Distinguished University Professor in the Department of Computer Science and Electrical Engineering at UMBC and is the director of the Machine Learning for Signal Processing Lab (MLSP Lab).

Prof. Adali assisted in the organization of a number of international conferences and workshops including the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), the IEEE International Workshop on Neural Networks for Signal Processing (NNSP), and the IEEE International Workshop on Machine Learning for Signal Processing (MLSP). She was the General Co-Chair, NNSP (2001–2003), MLSP 2023, Technical Chair, MLSP (2004–2008); Program Chair, MLSP (2008, 2009, and 2014), International Conference on Independent Component Analysis and Source Separation (2009); Publicity Chair, ICASSP (2000 and 2005); and Publications Chair, ICASSP 2008. She was the Technical Program Co-Chair for ICASSP 2017 and Special Sessions Chair for ICASSP 2018, ICASSP 2024, and EUSIPCO 2020. She is the Technical Program and Workshops Co-Chair for ICASSP 2026.

Prof. Adali is currently the editor-in-chief of the Signal Processing Magazine, and has served as the Chair of the IEEE Brain Technical Community, and Vice President-Technical Directions of the IEEE Signal Processing Society (SPS) for the period 2019–2022. She chaired the IEEE SPS MLSP Technical Committee (2003–2005, 2011–2013), and served or currently serving on a number of committees/boards including the SPS Conference Board (1998–2006, 2017–Present), IEEE SPS Fellow Reference Committee (2011–2014), IEEE SPS Signal Processing Theory and Methods (2010–2015) Technical Committee, and the IEEE SPS Bio Imaging and Signal Processing Technical Committee (2004–2007, 2018–Present). She was an Associate Editor for IEEE Transactions on Signal Processing (2003–2006), IEEE Transactions on Biomedical Engineering (2007–2013), IEEE Journal of Selected Areas in Signal Processing (2010–2013), and Elsevier Signal Processing Journal (2007–2010), and served on the Editorial Board of the Proceedings of the IEEE (2013–2019). She is currently serving on the Senior Editorial Board of IEEE Journal of Selected Topics in Signal Processing and the Editorial Board of Journal of Signal Processing Systems for Signal, Image, and Video Technology.

Prof. Adali is a Fellow of the IEEE, AIMBE, AAIA, a Fulbright Scholar, and an IEEE Signal Processing Society Distinguished Lecturer. She is the recipient of a 2020 Humboldt Research Award, 2013 University System of Maryland Regents' Award for Research, an NSF CAREER Award, and a number of paper awards including the 2010 IEEE Signal Processing Society Best Paper Award. Her current research interests are in the areas of statistical signal processing, machine learning, and applications in medical image analysis and fusion.



