

# **ACM** Journals

# **CALL FOR PAPERS**

## ACM Transactions on Multimedia Computing, Communications, and Applications

Special Issue on Deep Multimodal Generation and Retrieval

#### **Guest Editors:**

- Hao Fei, National University of Singapore, Singapore, haofei37@nus.edu.sg
- Wei Ji, National University of Singapore, Singapore, jiwei@nus.edu.sg
- Yinwei Wei, Monash University, Australia, weiyinwei@hotmail.com
- Zhedong Zheng, University of Macau, Macau, China <u>zhedongzheng@um.edu.mo</u>
- Jerry Jialie Shen, City, University of London, UK, jialie.shen.2@city.ac.uk
- Alan Hanjalic, Delft University of Technology, Netherlands, <u>A.Hanjalic@tudelft.nl</u>
- Roger Zimmermann, National University of Singapore, Singapore, dcsrz@nus.edu.sg

Recent advancements in Artificial Intelligence Generated Content (AIGC) have spotlighted two main strategies, information generation (IG) and information retrieval (IR), which respectively synthesize new content or search existing data to answer user queries. Despite successes in textual data, the full potential of IG and IR is limited by the underutilization of diverse data sources across various modalities. The push for deep multimodal learning aims to harness text, images, audio, and video for richer IG and IR applications, as demonstrated by technologies like DALL-E, GPT-4V, and Sora. However, challenges remain in aligning and fusing semantic features across modalities without redundancy, and in developing robust systems capable of handling real-world multimodal inputs. Addressing these issues, alongside exploring large-scale multimodal language models and structured metadata extraction, presents significant opportunities for enhancing multimodal IG and IR.

This special issue of ACM ToMM focuses on 'Deep Multimodal Generation and Retrieval' with the objective of advancing the field by uniting researchers and practitioners, fostering collaboration. It aims to promote the exchange of ideas and best practices between academia and industry, narrowing the gap in multimodal information generation and retrieval standards.

#### Topics

The main topics covered in this special issue (but not limited to) are shown as follows:

- Vision-Language Alignment Learning; Multimodal Fusion and Embeddings
- Commonsense-aware Vision-Language Learning
- Semantic-aware Vision-Language Discovery
- Multimodal Large Language Models (MLLMs); Large-scale Vision-Language Pre-training; Visually Grounded Interaction of Language Modeling
- Text-free/conditioned Image/Video Synthesis; Temporal Coherence in Video Generation; Image/Video Editing/Inpainting; LLM-empowered Multimodal Generation
- Multimodal Dialogue Response Generation; Image/Video Dialogue
- Image/Video-Text Compositional Retrieval; Video Moment Retrieval; Multimodal Retrieval with MLLMs
- Image/Video Captioning; Image/Video Question Answering
- Image/Video Relation Detection; Multimodal Event/Situation Recognition
- Hybrid Synthesis with Retrieval and Generation
- Explainable Multimodal Retrieval; Relieving Hallucination of LLMs; Adversarial Attack and Defense;

Efficient Learning of MLLMs

- New Benchmark; New Evaluation Metrics
- Multimodal-based Reasoning; Multimodal Instruction Tuning

## **Important Dates**

- Open for submissions: January 10, 2024
- Submission deadline: June 15, 2024
- First-round review decisions: August 15, 2024
- Deadline for revision submissions: September 30, 2024
- Notification of final decisions: October 30, 2024
- Tentative publication: December 30, 2024

#### **Submission Information**

Prospective authors are invited to submit their manuscripts electronically adhering to the ACM TOMM journal guidelines (see https://tomm.acm.org/authors.cfm). The manuscript will not be entertained if guidelines are not followed. The manuscriptshould be within the scope of ACM TOMM. Please submit your papers through the online system (https://mc.manuscriptcentral.com/tomm) and be sure to select the special issue. Manuscripts should not be published or currently submitted for publication elsewhere.

For questions and further information, please contact <u>haofei37@nus.edu.sg</u> and <u>jiwei@nus.edu.sg</u> and <u>weiyinwei@hotmail.com</u> and <u>zdzheng12@gmail.com</u>