



**Association for
Computing Machinery**

Advancing Computing as a Science & Profession

NEWS RELEASE

Contact: Jim Ormond
212-626-0505
ormond@acm.org

ACM MULTIMEDIA CONFERENCE SHOWCASES BROAD RANGE OF ADVANCES IN MULTISENSORY SYSTEMS

New York, NY, October 17, 2018 –The Association for Computing Machinery’s Special Interest Group on Multimedia (SIGMM) will host its annual [ACM Multimedia Conference \(ACM MM\)](#) in Seoul, Korea, from October 22-26, 2018. Now in its second quarter century, ACM MM is the premiere conference for multimedia experts and practitioners across academia and industry, who gather to present advanced innovations in mobile and wearable technologies, virtual/augmented/mixed reality, multisensory research and design, and much more.

“Although we have established a rich ecosystem for visual and auditory design, there is a lot that remains to be explored in the design of software, tools and techniques that will similarly engage our senses of touch, taste and smell,” notes Chang Wen Chen, ACM MM 2018 Program Co-Chair and Dean of the School of Science and Engineering at The Chinese University of Hong Kong, Shenzhen. “Combined with advances in machine learning, the Internet of Things and other computing subfields, our ability to design comprehensive, experience-centered technology is poised to dramatically impact modern ways of life.”

ACM MM 2018 HIGHLIGHTS

Keynote Speakers (partial list)

“Don’t Just Look–Smell, Taste, and Feel the Interaction”

Marianna Obrist, University of Sussex, Brighton, UK

While our understanding of the sensory modalities for human-computer interaction (HCI) is advancing, there is still a huge gap in our understanding of “how” to best integrate different modalities into the interaction with technology. Without engagement of all the senses we are missing out on the opportunity to exploit the power of all sensory modalities, their strong link to emotions and memory, and their ability to facilitate recall and recognition in information processing and decision making. In this keynote, Obrist presents an overview of scientific and technological developments in multisensory research and design with an emphasis on an experience-centered design approach that bridges and integrates knowledge on human sensory perception and advances in computing technology.

“Challenges and Practices of Large-Scale Visual Intelligence in the Real World”

Xian-Sheng Hua, Alibaba Group

Visual intelligence serves as an integral cornerstone in the development of many artificial intelligence systems, and considerable progress has been made in this rapidly advancing field in recent years. At the same time, however, how exactly to incubate the right technologies and ensure they subsequently produce real-world business value remains a challenge in many regards. In this talk, Hua will analyze current challenges facing the integration of visual intelligence, and identify key points that will help guide the future development and application of visual intelligence to solve real-world problems.

“Living with Artificial Intelligence Technology in Connected Devices around Us”

Gary Gunbae Lee, Samsung

As the Internet of Things and machine learning continue to advance simultaneously, we find the environments in which we live to be increasingly populated with intelligence and connected devices. In this talk Lee will discuss Samsung’s AI vision as a device company, and how the uses and applications of their devices have evolved along strategic lines over time.

“Transforming Retailing Experiences with Artificial Intelligence”

Bowen Zhou, JD.com

With its ability to combine knowledge of consumers, products and scenarios simultaneously, AI has already had a dramatic effect on the retail industry. Zhou will show how computer vision techniques can lead to a better understanding of consumers, how natural language processing can be used to support customer services through emotion computing, and how AI is building the very fundamental technology infrastructure for RaaS.

Best Papers (partial list of the finalists)

“GestureGan for Hand Gesture-to-Gesture Translation in the Wild”

Hao Tang, Nicu Sebe, University of Trento; Wei Wang, University of Trento, EPFL; Dan Xu, University of Trento, University of Oxford; Yan Yan, Texas State University

Hand gesture-to-gesture translation in the wild is a challenging task since hand gestures can have arbitrary poses, sizes, locations and self-occlusions. Therefore, this task requires a high-level understanding of the mapping between the input source gesture and the output target gesture. To tackle this problem, the authors propose a novel hand Gesture Generative Adversarial Network (GestureGAN).

“Understanding Humans in Crowded Scenes: Deep Nested Adversarial Learning and a New Benchmark for Multi-Human Parsing”

Jian Zhao, Yu Cheng, Jianshu Li, Li Zhou, Terence Sim, Yan Shuicheng, Jiashi Feng, National University of Singapore

Despite the noticeable progress in perceptual tasks like detection, instance segmentation and human parsing, computers still perform unsatisfactorily on visually understanding humans in crowded scenes, such as group behavior analysis, person re-identification and autonomous driving, etc. The authors present a new large-scale database Multi-Human Parsing (MHP) for algorithm development and evaluation, which advances the state of the art in understanding humans in crowded scenes.

“Beyond Narrative Description: Generating Poetry from Images by Multi-Adversarial Training”

Bei Liu, Makoto P Kato, Masatoshi Yoshikawa, Kyoto University; Jianlong Fu, Microsoft Research

Automatic generation of natural language from images has attracted extensive attention. The authors go one step further to investigate a method of generating poetic language (with multiple lines) from an image.

Additional Conference Highlights Include:

AltMM 2018, the 3rd International Workshop on Multimedia Alternate Realities

Multimedia experiences allow us to access other worlds, to live other people's stories, to communicate with or experience alternate realities. Different spaces, times or situations can be entered thanks to multimedia contents and systems, which coexist with our current reality, and are sometimes so vivid and engaging that we feel we are living in them. Advances in multimedia are making it possible to create immersive experiences that may involve the user in a different or augmented world, as an alternate reality. This workshop aims at exploring how the synergy between multimedia technologies can foster the creation of alternate realities and make their access an enriching and valuable experience. The program will contain a combination of oral and invited keynote presentations, and poster, demo and discussion sessions, altogether enabling interactive scientific sharing and discussion between practitioners and researchers.

Multimedia Grand Challenge

The Multimedia Grand Challenge presents a set of problems and issues from industry leaders and top academic institutions, geared to engage the multimedia research community in solving relevant, interesting and challenging questions for multimedia on a three-to-five-year vision. The three challenges are: Content-based Video Relevance Prediction Challenge; Perfect Corp. Challenge 2018: Half Million Beauty Product Image Recognition; and Social Media Headline Prediction.

Brave New Ideas (partial list)

Fluid Annotation: a Human-Machine Collaboration Interface for Full Image Annotation

Misha Andriluka, Google; Jasper Uijlings, Vittorio Ferrari, Google Research

The authors introduce Fluid Annotation, an intuitive human-machine collaboration interface for annotating the class label and outline of every object and background region in an image.

Harnessing AI for Speech Reconstruction Using Multi-view Silent Video Feed

Yaman Kumar, Mayank Aggarwal, Pratham Nawal, Nataji Subhas Institute of Technology; Shin'ichi Satoh, National Institute of Informatics; Rajiv Ratn Shah, Indraprastha Institute of Information Technology; Roger Zimmermann, National University of Singapore

Speechreading or lipreading is the technique of understanding and getting phonetic features from a speaker's visual features such as movement of lips, face, teeth and tongue. It has a wide range of multimedia applications such as in surveillance, internet telephony, and as an aid to a person with hearing impairments. In this paper, the authors present the world's first ever multi-view speech reading and reconstruction system.

Cross-Modal Health State Estimation

Nitish Nag, Vaibhav Pandey, Preston Putzel, Hari Bhimaraju, Ramesh Jain, University of California, Irvine; Srikanth Krishnan, University of California, Los Angeles

Individuals create and consume more diverse data about themselves today than any time in history. Sources of this data include wearable devices, images, social media, geospatial information and more. A tremendous opportunity rests within cross-modal data analysis that leverages existing domain knowledge methods to understand and guide human health. In this work we fuse multiple user-created and open source data streams along with established biomedical domain knowledge to give two types

of quantitative state estimates of cardiovascular health.

About ACM SIGMM

[SIGMM](#) is ACM's Special Interest Group on Multimedia—the community of researchers and practitioners dedicated to building next-generation technologies and applications around multimedia. SIGMM hosts several vibrant premiere conferences including ACM Multimedia (with 600+ participants annually), ICMR on multimedia retrieval, and MMSys on multimedia systems. The community also takes pride in its publications including the flagship journal *ACM TOMCCAP* and the affiliated *Springer Multimedia Systems Journal (MMSJ)*.

About ACM

[ACM, the Association for Computing Machinery](#), is the world's largest educational and scientific computing society, uniting computing educators, researchers and professionals to inspire dialogue, share resources and address the field's challenges. ACM strengthens the computing profession's collective voice through strong leadership, promotion of the highest standards, and recognition of technical excellence. ACM supports the professional growth of its members by providing opportunities for life-long learning, career development, and professional networking.

###