# Pushing the boundary of multimedia big data: an overview of IEEE MIPR

**Guan-Ming Su, Chengcui Zhang, Haohong Wang, Min Chen, Mohan Kankanhalli, Rainer Lienhart**

**Angaben zur Veröffentlichung / Publication details:**

Su, Guan-Ming, Chengcui Zhang, Haohong Wang, Min Chen, Mohan Kankanhalli, and Rainer Lienhart. 2019. "Pushing the boundary of multimedia big data: an overview of IEEE MIPR." IEEE MultiMedia 26 (2): 87–91. https://doi.org/10.1109/mmul.2019.2916957.



# Pushing the Boundary of Multimedia Big Data: An Overview of IEEE MIPR

**Guan-Ming Su**
Imaging Applied Research, Dolby Labs

**Chengcui Zhang**
Department of Computer Science, University of Alabama

**Haohong Wang**
TCL Research America

**Min Chen**
Computing and Software Systems, School of STEM, University of Washington Bothell

**Mohan Kankanhalli**
School of Computing, National University, Singapore

**Rainer Lienhart**
Multimedia Computing, Computer Vision Laboratory, University of Augsburg

***Abstract*—New forms of multimedia data (such as text, numbers, tags, networking, signals, geo-tagged information, graphs/relationships, three-dimensional/VR/AR and sensor data, etc.) have emerged in many applications in addition to traditional multimedia data (image, video, audio). Multimedia has become the biggest of big data as the foundation of today's data-driven discoveries. Almost all disciplines of science and engineering, as well as social sciences, involve multimedia data in some forms, such as recording experiments, driverless cars, unmanned aerial vehicles, smart communities, biomedical instruments, security surveillance. Some recent events demonstrate the power of real-time broadcast of unfolding events on social networks. Multimedia data is not just big in volume, but also multi-modal and mostly unstructured. Storing, indexing, searching, integrating, and recognizing from the vast amounts of data create unprecedented challenges. Even though significant progress has been made processing multimedia data, today's solutions are inadequate in handling data from millions of sources simultaneously.**

**The IEEE International Conference on Multimedia Information Processing and Retrieval (IEEE-MIPR) aims to provide a forum for original research contributions and practical system design, implementation, and applications of multimedia information processing and retrieval for single modality or multiple modalities. The target audiences are university researchers, scientists, industry practitioners, software engineers, and graduate students who want to become acquainted with technologies for big data analytics, machine intelligence, information fusion in multimedia information processing and retrieval.**

**Following a successful debut in Miami, Florida in 2018, the second MIPR, held in San Jose, California, attracted over 140 submissions and more than 100 attendees—a big growth from its inaugural event and an indication of its growing visibility and popularity in the multimedia research society. In addition to four keynote talks, MIPR'19 also featured six innovation multimedia forums organized by both academic and industrial researchers, with topics ranging from computer vision, self-driving, Natural Language Processing, next generation video display, to cutting edge video/audio coding techniques.**

## KEYNOTE

**The keynote by** John Apostolopoulos (Cisco) highlighted three examples of problems that multimedia experts can tackle by leveraging the new capabilities of networked multimedia systems as a new source of data, owing to the recent advances in both machine learning (ML) and networking (see Figure 1). The first example is about how to use Intent-Based Networking (a modern architecture for designing and operating a network) and ML to increase visibility, diagnose problems and identify associated remedies, and provide assurance on application performance for multimedia traffic such as video conferencing or wireless interactive AR/VR. The second example showed how distributed ML across Edge and Cloud can be a promising approach to improving scalability, reducing latency, and improving privacy for multimedia applications. Finally, Dr. Apostolopoulos presented the challenges of malware sneaking in encrypted flows and how ML can be used to detect hidden malware in the encrypted flow while at the same time to preserve privacy and improve security, without incurring higher complexity.


**Figure 1.** John Apostolopoulos gives a keynote at MIPR'19.

Patrick Griffis, SMPTE President and Technology Vice President in the Dolby CTO Office, presented a keynote speech on the theory and resulting practice leading to the successful mainstream deployment of high dynamic range (HDR) in the market today in television, mobile devices, and computers (see Figure 2). HDR uses a new electro-optic transfer function defined by SMPTE's standard ST-2084, which is based on the human visual model called perceptual quantizer.

Traditionally, standard dynamic range content is typically color-graded at 100 nits and then viewed by consumers at a higher brightness of 300 nits or more. In contrast, HDR investigates two essential questions, i.e., "how black is black" and "how bright is white," and uses 0–10 000 nits dynamic range as a design goal to better cover preferences and provide future headroom. Color volume mapping is then performed, which considers both tone mapping (intensity) and Gamut mapping (color), to map content into different displays including TV's, tablets, mobile phones, etc. Currently, HDR based on ST-2084 is being deployed in various products and services as evidenced from the recent consumer electronics show.


**Figure 2.** Patrick Griffis gives a keynote at MIPR'19.


**Figure 4.** Ruzena Bajcsy gives a keynote at MIPR'19.

Danny Lange from Unity Technology (see Figure 3) gave an interesting talk that walked the audience through the role of machine intelligence in biological evolution and learning and explored the relationship between machine intelligence and the senses. Specifically, he demonstrated a game engine with a spatial (three-dimensional) environment in conjunction with a physics engine which is considered by him a perfect example of a self-sufficient, controlled system that mimics the natural environment. He concludes by sharing new developments in reinforcement learning and the positive effect they will have on a variety of industries.

Ruzena Bajcsy from University of California Berkeley (see Figure 4) presented her recent work of multimodality information integration and utilization in studying the interplay of acoustic, visual, motion, and the driver's attention information, collected from both outside and inside environments during driving. Specifically, she investigated how the multimodal environment input, i.e., visual, acoustic, and motion data, affects the driver's attention/state while driving. For example, how do the other occupants' behaviors affect the driver's state? She then pointed out that the detection of the driver's state, measured by pressure sensors on the seat to indicate the restless level of the driver, often faces challenges posed by the mix of the motion data coming from both the movement of the driver and the roughness of the road.

## PROGRAM

MIPR 2019 received submissions from author spanning 21 countries. The acceptance rate of the full regular research papers is 19.5%. In addition, the conference also features high-quality short papers, industrial papers, demo papers, and poster papers (see Figure 5). The best paper award went to Jean-Baptiste Boin, Andre Araujo, and Bernd Girod for their work presented in "Recurrent

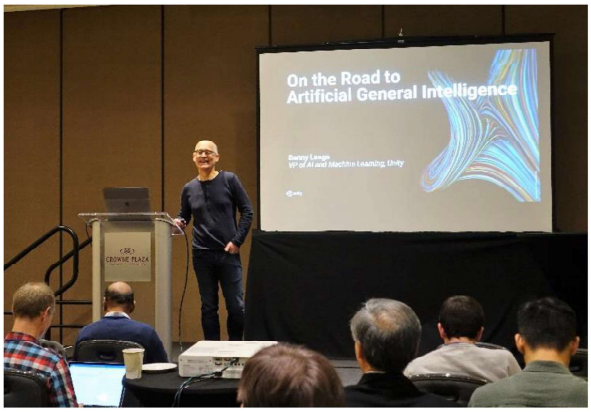


**Figure 3.** Danny Lange gives a keynote at MIPR'19.

**Figure 5.** Poster and demo session at MIPR'19.

Neural Networks for Person Re-Identification Revisited," which derives an approximation of RNN that leads to a much simpler architecture without sacrificing accuracy. Their new architecture, further combined with a new training process, has improved accuracy by up to 5% when compared with the other RNN-based frameworks.

Three workshops were held together with MIPR 2019 and covered several emerging multimedia research topics including (1) Fake MM 2019: The 2nd IEEE Workshop on Fake Multimedia, (2) MMPrag 2019: The 2nd IEEE Workshop on A Cross-Fertilization of Content-Based Retrieval and Natural Language Processing, and (3) AIArt 2019: the 1st IEEE Workshop on Artificial Intelligence for Art Creation.

Besides the regular sessions, MIPR 2019 features six innovation forums organized by renowned scientists in the multimedia area. The forum of the Latest Advances in Computer Vision and its Applications covered topics such as deep learning based medical image analysis, agriculture supervision, smart city, AI painting and beyond. In the Future of NLP/Audio Technology forum, the panel reviewed the latest research in NLP related to conversational speech recognition and performance with deep learning, audio event detection including isolating live spoken words from words spoken from devices to prevent false triggers or mitigate threat, emotion and behavior analysis using unimodal and multimodal (both camera and microphone) signals, speech synthesis and voice conversion, speech rate modeling and augmentation, and intelligent agents for natural language conversation.

The Autonomous Driving panel presented and discussed various recent auto-driving assisting technologies such as ambient intelligence enabled smart environment that delivers more complete pictures to self-driving cars to enable better negotiation, the utilization of LIDAR techniques in self-driving and the associated challenges, high definition map construction, and the advances in self-driving technology to enable large scale autonomous commercial fleets.

The Next Generation Video & Display Technology panel covered topics ranging from HDR television, audio–visual quality assessment, to perception-centered AR/VR displays. The Video Compression panel presented the recent advanced in utilizing deep learning for video compression and standardization, the next joint video coding standard called Versatile Video Coding, immersive media coding, and the future of video compression. The panel of Video for VR/AR discussed some of the challenges and tradeoffs for building AR/VR media systems that balance between high resolution, wide field-of-view, high frame rate, degree-of-freedom in user perspectives, etc., from a practical engineering point of view.

## CONCLUDING REMARKS

MIPR'19 has grown significantly from its first event, with more than 140 submissions and over 100 participants. The participants thoroughly enjoyed the program which is a mix of keynotes, research presentations, and innovation forums, and felt that they particularly benefited from the keynote speakers' and the panelists' vision and were inspired by their talks. ML and AI have increasingly penetrated into various areas in multimedia research, including auto-driving, biomedical research and healthcare, AR/VR/MR, video/audio coding and compression, and multimedia networking, etc. In turn, this has triggered unprecedented opportunities for new discoveries and applications that will keep pushing the front of multimedia research. We are confident that MIPR'19 will continue to play an important role in this golden age.

**Guan-Ming Su** is a Senior Manager of Imaging Applied Research at Dolby Labs, San Francisco, CA, USA. Contact him at guanmingsu@ieee.org.

**Chengcui Zhang** is a Full Professor with the Department of Computer Science, the University of Alabama at Birmingham, Birmingham, AL, USA. Contact her at czhang02@uab.edu.

**Haohong Wang** is the General Manager of TCL Research America, in charge of the overall product and technology innovations for TCL Corporation at North America. Contact him at haohong.wang@tcl.com.

**Min Chen** is an Associate Professor with the Computing and Software Systems, School of STEM, the University of Washington Bothell, Bothell, WA, USA. Contact her at minchen2@uw.edu.

**Mohan Kankanhalli** a Full Professor and the Dean with the School of Computing, the National University of Singapore. Contact him at mohan@comp.nus.edu.sg.

**Rainer Lienhart** is a Full Professor and Head with the Multimedia Computing and Computer Vision Laboratory, the University of Augsburg, Augsburg, Germany. Contact him at lienhart@informatik.uni-augsburg.de.