Multimedia and Language: Bridging Multimedia and Natural Language with Deep Learning

Tao Mei, Microsoft Research, China

Jiebo Luo, University of Rochester, USA

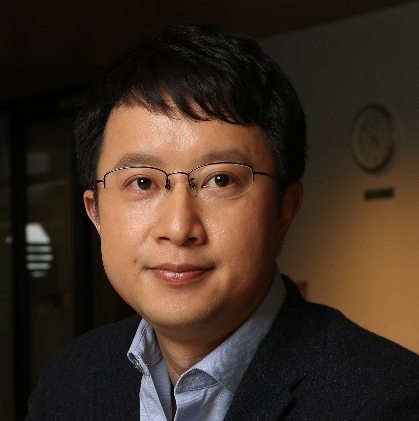
Abstract:

Recognition of visual content has been a fundamental challenge in computer vision and multimedia for decades, where previous research predominantly focused on understanding visual content using a predefined yet limited vocabulary. Thanks to the recent development of deep learning techniques, researchers in both computer vision and multimedia communities are now striving to bridge multimedia with natural language, which can be regarded as the ultimate goal of visual understanding. We will present recent advances in exploring the synergy of multimedia content understanding and language processing techniques, including multimedia-language alignment, visual captioning and commenting, visual emotion analysis, visual question answering, visual storytelling, and as well as open issues for this emerging research area.

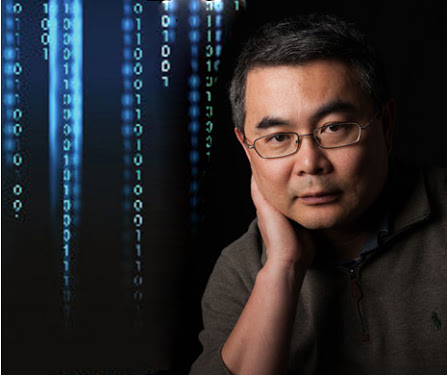
Bio:

**Tao Mei** is a Senior Researcher and Research Manager with Microsoft Research Asia. His current research interests include multimedia analysis and computer vision. He is leading a team working on image and video analysis, vision and language, and multimedia search. He has authored or co-authored over 150 papers with 11 best paper awards. He holds over 50 filed U.S. patents (with 20 granted) and has shipped a dozen inventions and technologies to Microsoft products and services. He is an Editorial Board Member of IEEE Trans. on Multimedia, ACM Trans. on Multimedia Computing, Communications, and Applications, and Pattern Recognition. He is the General Co-chair of IEEE ICME 2019, the Program Co-chair of ACM Multimedia 2018, IEEE ICME 2015, and IEEE MMSP 2015. Tao is as a Fellow of IAPR and a Distinguished Scientist of ACM.

**Jiebo Luo** joined the University of Rochester in Fall 2011 after over fifteen years at Kodak Research Laboratories, where he was a Senior Principal Scientist leading research and advanced development. He has been involved in numerous technical conferences, including serving as the program co-chair of ACM Multimedia 2010, IEEE CVPR 2012, and IEEE ICIP 2017. 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, Pattern Recognition, Machine Vision and Applications, and Journal of Electronic Imaging. He has authored over 300 technical papers and 90 US patents. Prof. Luo is a Fellow of the SPIE, IEEE, and IAPR.



Tao Mei



Jiebo Luo