VSCC'2017: Visual Analysis for Smart and Connected Communities

Xiaobai Liu
Department of Computer Science
San Diego State University (SDSU)
San Diego, California
xiaobai.liu@mail.sdsu.edu

Yadong Mu
Institute of Computer Science and Technology
Peking University
Beijing, China
myd@pku.edu.cn

Yu-Gang Jiang
School of Computer Science
Fudan University
Shanghai, China
ygj@fudan.edu.cn

Jiebo Luo
Department of Computer Science
University of Rochester
Rochester, New York
jluo@cs.rochester.edu

ABSTRACT

This paper presents a brief summary of the first workshop on Visual Analysis for Smart and Connected Communities (VSCC'2017), which is held in conjunction with the ACM Conference on Multimedia, 2017. VSCC'2017 is the first workshop on visual analysis for smart and connected communities, and is aimed at the creation of a multi-discipline community with the concentration on visual data analysis. The topics addressed in VSCC'2017 cover all aspects of smart communities, including safety, security, retrieval, transportation, information technologies, Internet of Things, etc. The focus is on the development of fundamental theories, algorithms, models towards the parsing of visual data generated by various digital devices (e.g., smart phones, vehicles) in smart communities.

Keywords Smart Communities; Visual Data; Video Parsing; Image Parsing; Machine Learning; Connected Vehicles; Commonsense; traffic sign recognition; 3D scene reconstruction; object detection

CCS CONCEPTS

• **Information systems → Data mining**; • **Computing methodologies → Computer vision**; **Machine learning**;

1 INTRODUCTION

Advances in smart technologies [7] are transforming cities and communities around the world through connecting their inhabitants and the surrounding built and natural environments, leading to new opportunities for innovations, improved services, and enhanced quality of life. In this new ear of transformational change,

most developed communities are already equipped with visual sensors overseeing critical infrastructures [5] (e.g. buildings, roadways, parking-lots etc.), which provides an intrusive way to extracting community-wide data (e.g., traffic conditions, emergency events) and thus supporting decision making in all aspects of smart communities. However, the popularity of visual data imposes new challenges due to their volume, arrival scale, variances, and complexity.

The mission of this workshop is to bring together researchers and practitioners from diverse areas of artificial intelligence âĂŞ vision, multimedia, learning, reasoning, planning, etc., to explore both established and novel applications of visual analytic techniques. The developed techniques would potentially address problems related to all aspects of smart communities, including safety [4], security [3], transportation [2], information technologies [6], retrieval [8], Internet of Things [1] etc.

In particular, the three goals of this event are to (a) identify the key challenges in visual analysis for smart and connected communities; (b) recognize the promising solutions to these challenges; and (c) remark the future of research in this inter-discipline domain.

2 TOPICS AND SCOPES

The topics of this event include, but not being limited to,

(1) Visual Commonsense reasoning
(2) 3D Scene parsing and Reconstruction
(3) Multi-view object tracking with overlapping or non-overlapping cameras
(4) Image classification and segmentation
(5) Video segmentation
(6) Traffic conditions assessment
(7) Emergency event recognitions
(8) Visual perception in autonomous systems
(9) Large-scale image/video annotation, indexing and search
(10) Visual odometry
(11) Visual landmark search
(12) Vision-based localization and navigation
(13) Object (pedestrian/vehicle/road sign etc.) detection

3 WORKSHOP PROGRAM FORMAT

The workshop accepted Six full research papers, which cover various topics of visual analysis, including video-based localization, autonomous driving systems, video retrieval, semi-supervised distance learning, and human skeleton detection etc. These papers will be presented before the main conference of ACM Conference on Multimedia, 2017.

4 ORGANIZATION TEAM

Dr. Xiaobai Liu is an Assistant Professor of Computer Science at the San Diego State University (SDSU), San Diego, California, U.S.A. He is also affiliated with the Center for Vision, Cognition, Learning and Autonomy (VCLA) at the University of California, Los Angeles (UCLA). His researches focus on the fundamental problems of scene parsing with a variety of topics, e.g. joint inference for recognition and reconstruction, commonsense reasoning, etc. He has published more than 35 peer-reviewed articles in top conferences (e.g. IEEE Conference on Computer Vision and Pattern Recognition) and leading journals (e.g. IEEE Transactions on Pattern Analysis and Machine Intelligence). He received a number of awards for his academic contributions, including the 2013 outstanding thesis award by CCF (China Computer Federation).

Dr. Yadong Mu is an Assistant Professor of Computer Science at the Peking University, Peking, China. Dr. Mu leads the Machine Intelligence Lab at Institute of Computer Science and Technology, Peking Univeristy. Before joining Peking University, he has ever worked as research fellow at National University of Singapore, research scientist at the DVMM lab of Columbia University, researcher at the data mining team of Huawei Noah's Ark Lab in Hong Kong, and senior scientist at Multimedia Department of AT&T Labs, New Jersey, U.S.A. He obtained both the B.S. and Ph.D. degrees from Peking University. His research interests cover the broad research topics in computer vision, data mining and machine learning, particularly large-scale image and video computing (search, indexing, event detection etc), telecom data mining, autonomous driving techniques, and distributed / approximate large-scale machine learning.

Dr. Yu-Gang Jiang is a Professor in the School of Computer Science at Fudan University, China and serves as the Vice Director of Shanghai Engineering Research Center for Video Technology and System. He is leading the Lab of Big Video Data Analytics to conduct research on all aspects of extracting high-level information from big video data, such as video event recognition, object/scene recognition and large-scale visual search. He is the lead architect of a few best-performing video analytic systems in worldwide competitions such as the annual U.S. NIST TRECVID evaluation. His visual concept detector library (VIREO-374) and video datasets (e.g., CCV, FCVID and THUMOS) are widely used resources in the research community. His work has led to many awards, including "emerging leader in multimedia" award from IBM T.J. Watson Research in 2009, early career faculty award from Intel and China Computer Federation, the inaugural ACM China Rising Star Award, the 2015 ACM SIGMM Rising Star Award, and the research award for outstanding young researchers from NSF China. He holds a PhD in Computer Science from City University of Hong Kong and spent three years working at Columbia University before joining Fudan in 2011.

Dr. Jiebo Luo is a Professor of Computer Science at the University of Rochester, Rochester, NY, U.S.A. Dr. Luo jointed the university in Fall 2011 after over fifteen prolific 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, ACM Transactions on Intelligent Systems and Technology, Pattern Recognition, Machine Vision and Applications, and Journal of Electronic Imaging. Dr. Luo is a Fellow of the SPIE, IEEE, and IAPR, and a member of ACM, AAAI, and AAAS.

ACKNOWLEDGMENTS

REFERENCES

[1] Luigi Atzori, Antonio Iera, and Giacomo Morabito. The internet of things: A survey. *Computer networks*, 54(15):2787–2805, 2010.

[2] Benjamin Coifman, David Beymer, Philip McLauchlan, and Jitendra Malik. A real-time computer vision system for vehicle tracking and traffic surveillance. *Transportation Research Part C: Emerging Technologies*, 6(4):271–288, 1998.

[3] Robert T Collins, Alan J Lipton, Takeo Kanade, Hironobu Fujiyoshi, David Duggins, Yanghai Tsin, David Tolliver, Nobuyoshi Enomoto, Osamu Hasegawa, Peter Burt, et al. A system for video surveillance and monitoring. *VSAM final report*, pages 1–68, 2000.

[4] Xiaobai Liu, Liang Lin, Shuicheng Yan, Hai Jin, and Wenbin Jiang. Adaptive object tracking by learning hybrid template online. *IEEE Transactions on Circuits and Systems for Video Technology*, 21(11):1588–1599, 2011.

[5] Xiaobai Liu, Liang Lin, Shuicheng Yan, Hai Jin, and Wenbing Tao. Integrating spatio-temporal context with multiview representation for object recognition in visual surveillance. *IEEE Transactions on Circuits and Systems for Video Technology*, 21(4):393–407, 2011.

[6] Debashis Saha and Amitava Mukherjee. Pervasive computing: a paradigm for the 21st century. *Computer*, 36(3):25–31, 2003.

[7] Hans Schaffers, Nicos Komninos, Marc Pallot, Brigitte Trousse, Michael Nilsson, and Alvaro Oliveira. Smart cities and the future internet: Towards cooperation frameworks for open innovation. *The future internet*, pages 431–446, 2011.

[8] Ke Xia, Yuqing Ma, Xianglong Liu, Yadong Mu, and Li Liu. Temporal binary coding for large-scale video search. In *ACM Multimedia*, 2017.