Guest Editors' Introduction

Urban Multimedia Computing: Emerging Methods in Multimedia Computing for Urban Data Analysis and Applications

Weiqing Min
Institute of Computing Technology, Chinese Academy of Sciences

Yonghong Tian
Peking University

Zi Huang
The University of Queensland

Wen-Huang Cheng
National Chiao Tung University

Abdulmotaleb El Saddik
University of Ottawa

URBAN COMPUTING IS a process of acquisition, integration, analysis, and understanding of the urban data generated from various sources, e.g., sensors, devices, and social media, for various applications, e.g., tackling the air pollution and monitoring urban activities. Multimedia computing plays a vital role in urban computing due to the huge presence of multimodal sensors, heterogeneous multimedia data (e.g., check-ins, location, images, videos, and text) from the social media, multimedia interaction between human and cities. Recently, benefiting from both the easy availability of urban multimedia big data and the rapid development of deep learning technologies, some novel methods in multimedia computing have emerged as a promising tool for solving new urban computing tasks, such as quantifying the urban perception and estimating the demographic makeup of neighborhoods. Meanwhile, these advances have led to a new research direction: urban multimedia computing at the intersection between multimedia computing and urban computing. In order to effectively utilize multimedia computing methods for modeling the urban multimedia data, we should consider relevant theories from other fields. Therefore, there's a growing demand for developing and designing new multimedia computing methods for urban data analysis and applications, especially in current deep learning and multimedia big data era. We thus believe it is necessary to bring together researchers and practitioners from all related areas to discuss this topic. Hopefully, some new models in multimedia computing for urban multimedia computing can be proposed by fusing the knowledge of experts from different fields.

Digital Object Identifier 10.1109/MMUL.2020.3017877
Date of current version 28 August 2020.

IN THIS ISSUE

These six articles in this special issue cover different techniques and applications for

1070-986X © 2020 IEEE Published by the IEEE Computer Society

urban multimedia computing. These include intelligent video analysis, large-scale visual geo-localization, urban scene understanding, and patient records management for urban telehealth service.

Feature coding has recently become an important part to facilitate intelligent video analysis for urban multimedia computing. Instead of raw videos, extracted features in the front-end are encoded and transmitted to the back-end for further processing. In the article "Key-Point Sequence Lossless Compression for Intelligent Video Analysis," W. Lin, X. He, W. Dai, J. See, T. Shinde, H. Xiong, and L. Duan present a lossless key-point sequence compression approach for efficient feature coding. In their method, multiple prediction modes with an adaptive mode selection method are proposed to handle key-point sequences with various structures and motion. The essence of this predict-and-encode strategy is to eliminate the spatial and temporal redundancies of key points in videos. Experimental results validate the effectiveness of the proposed scheme on four types of widely used key-point sequences in video analysis.

As one important application of intelligent video analysis, person reidentification has been gaining considerable attention. In the article, "PGAN: Part-Based Nondirect Coupling Embedded GAN for Person Reidentification," Y. Zhang, Y. Jin, J. Chen, S. Kan, Y. Cen, and Q. Cao present a deep neural network for person reidentification, which involves the interaction between humans and cities. In their method, a part-based nondirect coupling representation learning method is proposed to constrain features of different blocks for person feature diversity, and a part-based nondirect coupling embedded generative adversarial networks (GANs) method is proposed to extract more common features of different postures for one person. In this way, the data representation is robust for posture changes of a person, and the accuracy of cross-camera pedestrian rerecognition is improved. Experimental results on public datasets verify the efficacy of the proposed method.

Large-scale visual geo-localization (LSVGL) is increasingly important in urban multimedia computing, where the task is to accurately and efficiently recognize the geo-location of a given query image. In the article, "Learning Quintuplet Loss for Large-scale Visual Geo-Localization," Q. Zhai, R. Huang, H. Cheng, H. Zhan, J. Li, and Z. Liu design a new metric learning method to deal with the performance degradation caused by perspective difference in large-scale geo-localization. In their method, considering about the perspective deviation almost inevitably exists between training images and query images, a quintuplet loss is embedded into traditional metric learning methods through embedding all potential positive samples to primitive metric learning loss. Then, several classical metric learning losses are compared to provide the first reference for scholars in related fields. Comparison experiments are designed to prove that the new quintuplet loss can improve the performance of all previous learning losses. Extensive experiments have been conducted to verify the effectiveness of the proposed approach and the results demonstrate that the proposed new metric learning loss enhances various LSVGL methods.

Urban scene understanding is another important topic for urban intelligent transportation systems, such as self-driving. In the article, "Domain Adaptation with Foreground/Background Cues and Gated Discriminators," Y.-X. Lin, D. S. Tan, Y.-Y. Chen, K.-L. Hua, and C.-C. Huang propose an adaptation method that uses foreground and background cues and adapts them separately for urban scene segmentation. In their method, they propose a mask-aware gated discriminator to learn soft masks from the input foreground and background masks instead of simple binary masking. The evaluation on two different datasets demonstrates that the proposed method outperforms several state-of-the-art baselines.

Style transfer of urban environment is used for evaluating the driving behavior of unmanned vehicles, which can also benefit urban intelligent transportation systems. In "Style Transfer of Urban Road Images using Generative Adversarial Networks with Structural Details," Y. Li, X. Wu, D. Lu, L. Li, Y. Liu, and L. Zhu address the evaluation of the driving behavior of unmanned vehicles. The testing of driving

algorithms using urban environmental data is necessary. In their method, a framework using GANs with structural information is proposed for image style transfer: StructureGAN and Gradient-GAN. Different types of urban image transfers are generated using the proposed framework, such as day to night, sunny to foggy, and summer to winter transfers. The proposed method can well maintain the integrity of foreground objects and the image structural information. The experiments indicate the effectiveness of the proposed framework, which can produce transferred images with high quality.

Finally, in "Joint Watermarking-Encryption-ECC for Patient Record Security in Wavelet Domain," A. Anand and A. K. Singh develop a joint watermarking-encryption error correcting code (ECC) technique for securing patient records, which is the primary role for the urban telehealth services. The proposed method consists of three steps, including 1) watermarks generation and scrambling, 2) encryption of cover and generated watermark, and 3) imperceptibly embedding and robustly recovery/extraction of the watermarks. Experimental results show that the proposed method offers high robustness and fulfills all watermarking needs for telehealth applications in terms of imperceptibility, robustness, capacity, and security.

ACKNOWLEDGMENTS

The authors would like to thank all the authors and reviewers for their great efforts on this successful special issue. They would also like to thank EIC S.-C. Chen and AEIC M. Kankanhalli for giving them the opportunity to organize this special issue.

Weiqing Min is currently an Associate Professor with the Key Laboratory of Intelligent Information Processing, Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China. His current research interests include geo-multimedia analysis and food computing. He has authored or coauthored more than 40 academic papers in leading international conferences and journals, including *ACM Computing Surveys*, IEEE Transactions on Image Processing, IEEE Transactions on Multimedia, IEEE Multimedia Magazine, *ACM Transactions on Multimedia Computing, Communications, and Applications*, *ACM Multimedia (*ACM MM), Association for the Advancement of Artificial Intelligence (AAAI), International Joint Conferences on Artificial Intelligence (IJCAI), etc. He is a Guest Editor for several journals, such as *Neurocomputing*, *Multimedia Tools and Applications*. He served as a TPC Member of many academic conferences, including ACM MM, AAAI, IJCAI, etc. He is the recipient of 2016 ACM TOMM Nicolas D. Georganas Best Paper Award and 2017 IEEE Multimedia Magazine Best Paper Award. Contact him at minweiqing@ict.ac.cn.

Yonghong Tian is currently a Boya Distinguished Professor with the School of EECS, Peking University, Beijing, China, and is also the Deputy Director of the Artificial Intelligence Research Center, PengCheng Laboratory, Shenzhen, China. His research interests include computer vision, multimedia big data, and brain-inspired computation. He is the author or coauthor of more than 200 technical articles in refereed journals such as IEEE Transactions on Pattern Analysis and Machine Intelligence/IEEE Transactions on Neural Networks and Learning Systems /IEEE Transactions on Image Processing/IEEE Transactions on Multimedia (TMM)/IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems (TCSVT)/IEEE Transactions on Knowledge and Data Engineering/IEEE Transactions on Parallel and Distributed Systems, *ACM Computing Surveys*/*ACM Transactions on Information Systems*/*ACM Transactions on Multimedia Computing, Communications, and Applications*, and conferences such as NeurIPS/CVPR/ICCV/AAAI/ACMMM/WWW. Prof. Tian was/is an Associate Editor of IEEE TCSVT (2018.1-), IEEE TMM (2014.8-2018.8), IEEE Multimedia Magazine (2018.1-), and IEEE Access (2017.1-). He co-initiated IEEE International Conference on Multimedia Big Data (BigMM) and served as the TPC Co-chair of BigMM 2015, and also as the Technical Program Co-chair of IEEE ICME 2015, IEEE ISM 2015, and IEEE MIPR 2018/2019, and General Co-chair of IEEE MIPR 2020. He is the steering member of IEEE ICME (2018-) and IEEE BigMM (2015-), and is a TPC Member of more than ten conferences such as CVPR, ICCV, ACM KDD, AAAI, ACM MM, and ECCV. He was the recipient of the Chinese National Science Foundation for Distinguished Young Scholars in 2018, two National Science and Technology Awards, and three ministerial-level awards in China, and obtained the 2015 EURASIP Best Paper Award for the *Journal on Image and Video Processing*, and the best paper award of IEEE BigMM 2018. He is a senior member of IEEE, CIE and CCF, a member of ACM. Contact him at yhtian@pku.edu.cn.

Zi Huang is currently an Associate Professor and ARC Future Fellow with the School of ITEE, The University of Queensland, Brisbane, QLD, Australia. Her research interests mainly include multimedia indexing and search, social data analysis and knowledge discovery. She has authored and coauthored 100+ papers in prestigious venues, and is currently an Associate Editor of *The VLDB Journal*. He has received 2016 Chris Wallace Award from Computing Research and Education (CORE) Australasia for a notable breakthrough or a contribution of particular significance in Computer Science, and Women in Technology (WiT) Infotech Research Award 2014, Queensland. She was also a recipient of the Excellence in Higher Degree by Research Supervision Award, University of Queensland, 2018. She is the Data Science Discipline Leader, UQ. She received the B.Sc. degree from the Department of Computer Science, Tsinghua University, Beijing, China, and the Ph.D. degree in computer science from the School of ITEE, The University of Queensland in 2001 and 2007, respectively. Contact her at huang@itee.uq.edu.au.

Wen-Huang Cheng is currently a Professor with the Institute of Electronics, National Chiao Tung University (NCTU), Hsinchu, Taiwan, where he is the Founding Director with the Artificial Intelligence and Multimedia Laboratory (AIMMLab). Before joining NCTU, he led the Multimedia Computing Research Group with the Research Center for Information Technology Innovation (CITI), Academia Sinica, Taipei, Taiwan, from 2010 to 2018. His current research interests include multimedia, artificial intelligence, computer vision, machine learning, social media, and financial technology. He has actively participated in international events and played important leading roles in prestigious journals and conferences and professional organizations, like an Associate Editor for the IEEE TRANSACTIONS ON MULTIMEDIA, General co-chair for ACM ICMR (2021), TPC co-chair for ICME (2020), Chair-Elect for IEEE MSA-TC, governing board member for IAPR. He has received numerous research and service awards, including the 2018 MSRA Collaborative Research Award, the 2017 Ta-Yu Wu Memorial Award from Taiwan's Ministry of Science and Technology (the highest national research honor for young Taiwanese researchers under age 42), the Top 10% Paper Award from the 2015 IEEE MMSP, the K. T. Li Young Researcher Award from the ACM Taipei/Taiwan Chapter in 2014, the 2017 Significant Research Achievements of Academia Sinica, the 2016 Y. Z. Hsu Scientific Paper Award, the Outstanding Youth Electrical Engineer Award from the Chinese Institute of Electrical Engineering in 2015, and the Outstanding Reviewer Award of 2018 IEEE ICME. He is APSIPA Distinguished Lecturer. Contact him at whcheng@nctu.edu.tw.

Abdulmotaleb El Saddik is a Distinguished University Professor and the University Research Chair in the School of Electrical Engineering and Computer Science, University of Ottawa, Ottawa, ON, Canada. He has authored or coauthored 10 books and more than 550 publications and chaired more than 50 conferences and workshops, and has received research grants and contracts totalling over $20M. His research focuses on multimodal interactions with sensory information in smart cities. Prof. El Saddik is a Senior Associate Editor of the *ACM Transactions on Multimedia Computing, Communications and Applications and IEEE Multimedia*, and a Guest Editor for several IEEE transactions and journals. He has supervised more than 120 researchers and received several international awards, among which, are the ACM Distinguished Scientist, the A.G.L. McNaughton Gold Medal, and IEEE Instrumentation and Measurement Technical Achievement Award. He is a Fellow of the Engineering Institute of Canada, Fellow of the Canadian Academy of Engineers and Fellow of the IEEE. Contact him at elsaddik@uottawa.ca.