IET Computer Vision

DOI: 10.1049/cvi2.12283

The Institution of Engineering and Technology WILEY

GUEST EDITORIAL

Guest Editorial: Advanced image restoration and enhancement in the wild

1 | INTRODUCTION

Image restoration and enhancement has always been a fundamental task in computer vision and is widely used in numerous applications, such as surveillance imaging, remote sensing, and medical imaging. In recent years, remarkable progress has been witnessed with deep learning techniques. Despite the promising performance achieved on synthetic data, compelling research challenges remain to be addressed in the wild. These include: (i) degradation models for low-quality images in the real world are complicated and unknown, (ii) paired low-quality and high-quality data are difficult to acquire in the real world, and a large quantity of real data are provided in an unpaired form, (iii) it is challenging to incorporate cross-modal information provided by advanced imaging techniques (e.g. RGB-D camera) for image restoration, (iv) real-time inference on edge devices is important for image restoration and enhancement methods, and (v) it is difficult to provide the confidence or performance bounds of a learning-based method on different images/regions. This special issue invites original contributions in datasets, innovative architectures, and training methods for image restoration and enhancement to address these and other challenges.

2 | PAPERS IN THE SPECIAL ISSUE

In this Special Issue, we have received 17 papers, of which 8 papers underwent the peer review process, while the rest were desk-rejected. Among these reviewed papers, 5 papers have been accepted and 3 papers have been rejected as they did not meet the criteria of IET Computer Vision. Thus, the overall submissions were of high quality, which marks the success of this Special Issue.

The five eventually accepted papers can be clustered into two categories, namely video reconstruction and image super-resolution. The first category of papers aims at reconstructing high-quality videos. The papers in this category are of Zhang et al., Gu et al., and Xu et al. The second category of papers studies the task of image super-resolution. The papers in this category are of Dou et al. and Yang et al. A brief presentation of each of the paper in this special issue is as follows.

Zhang et al. propose a point-image fusion network for event-based frame interpolation. Temporal information in event streams plays a critical role in this task as it provides temporal context cues complementary to images. Previous approaches commonly transform the unstructured event data to structured data formats through voxelisation and then employ advanced CNNs to extract temporal information. However, the voxelisation operation inevitably leads to information loss and introduces redundant computation. To address these limitations, the proposed method directly extracts temporal information from the events at the point level without relying on any voxelisation operation. Afterwards, a fusion module is adopted to aggregate complementary cues from both points and images for frame interpolation. Experiments on both synthetic and real-world datasets show that their method produces state-of-the-art accuracy with high efficiency.

Gu et al. develop a temporal shift reconstruction network for compressive video sensing. To exploit the temporal cues between adjacent frames during the reconstruction of videos, most previous approaches commonly preform alignment between initial reconstructions. However, the estimated motions are usually too coarse to provide accurate temporal information. To remedy this, the proposed network employs stacked temporal shift reconstruction blocks to enhance the initial reconstruction progressively. Within each block, an efficient temporal shift operation is used to capture temporal structures in addition to computational overheads. Then, a bidirectional alignment module is adopted to capture the temporal dependencies in a video sequence. Different from previous methods that only extract supplementary information from the key frames, the proposed alignment module can receive temporal information from the whole video sequence via bidirectional propagations. Experiments demonstrate the superior performance of the proposed method.

Qu et al. propose a lightweight video frame interpolation network with a three-scale encoding-decoding structure. Specifically, multi-scale motion information is first extracted from the input video. Then, recurrent convolutional layers are adopted to refine the resultant features. Afterwards, the

resultant features are aggregated to generate high-quality interpolated frames. Experimental results on the CelebA and Helen datasets show that the proposed method outperforms state-of-the-art methods while using fewer parameters.

Dou et al. introduce a decoder structure-guided CNN-Transformer network for face super-resolution. Most previous approaches follow a multi-task learning paradigm to perform landmark detection while super-resolving the low-resolution images. However, these methods require additional annotation cost, and the extracted facial prior structures are usually of low quality. To address these issues, the proposed network employs a global-local feature extraction unit to extract the global structure while capturing local texture details. In addition, a multi-state fusion module is incorporated to aggregate embeddings from different stages. Experiments show that the proposed method surpasses previous approaches by notable margins.

Yang et al. study the problem of blind super-resolution and propose a method to exploit degradation information through degradation representation learning. Specifically, a generative adversarial network is employed to model the degradation process from HR images to LR images and constrain the data distribution of the synthetic LR images. Then, the learnt representation is adopted to super-resolve the input low-resolution images using a transformer-based SR network. Experiments on both synthetic and real-world datasets demonstrate the effectiveness and superiority of the proposed method.

3 | GUEST EDITORS

3.1 | Longguang Wang

Longguang Wang received his BE and PhD degrees from Shandong University and National University of Defence Technology (NUDT) in 2015 and 2022, respectively. He is currently an assistant professor with Aviation University of Air Force. He authored more than 40 peer-reviewed journals and conference publications (including TPAMI, TIP, CVPR, ICCV, and ECCV). He has organised three workshops at CVPR 2022 and 2023. His research interests include low-level vision and 3D vision, particularly on image restoration, image enhancement, image generation, depth estimation, point cloud understanding, and network acceleration. He received the CSIG Excellent Doctoral Dissertation Nomination Award in 2022 (17 nationwide).

3.2 | Juncheng Li

Juncheng Li received the Ph.D. degree from the School of Computer Science and Technology, East China Normal University, in 2021. He also worked as a Postdoctoral Fellow at the Center for Mathematical Artificial Intelligence, The Chinese University of Hong Kong. He is currently an assistant professor with Shanghai University. His research interests include artificial intelligence and its applications to computer vision (e.g. image segmentation) and image processing (e.g. image super-resolution, image denoising, and image dehazing). He has published more than 25 papers in top journals and conferences, including TIP, TNNLS, TMM, ECCV, ICCV, AAAI, ACMMM, and IJCAI. He also received several premium awards, including the Shanghai Outstanding Ph.D. Graduates, CUHK Research Fellowship Scheme, and the winner of 2019 ICCV-AIM.

3.3 | Naoto Yokoya

Naoto Yokoya received the M.Eng. and Ph.D. degrees from the Department of Aeronautics and Astronautics, The University of Tokyo, Tokyo, Japan, in 2010 and 2013, respectively. From 2013 to 2017, he was an assistant professor with The University of Tokyo. From 2015 to 2017, he was an Alexander von Humboldt Fellow, working at the German Aerospace Center, Oberpfaffenhofen, Germany and at the Technical University of Munich, Munich, Germany. He is currently a lecturer with The University of Tokyo and a unit leader with the RIKEN Center for Advanced Intelligence Project, Tokyo, where he leads the Geoinformatics Unit. His research interests include image processing, data fusion, and machine learning for understanding remote sensing images with applications to disaster management. Dr. Yokoya received the First Place in the 2017 IEEE Geoscience and Remote Sensing Society (GRSS) Data Fusion Contest organised by the IEEE Image Analysis and Data Fusion Technical Committee (IADF TC). From 2019 to 2021, he was the Chair and the Co-Chair (2017–2019) of the IEEE GRSS IADF TC. Since 2018, he has been an associate editor of IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing (JSTARS).

3.4 | Radu Timofte

Radu Timofte received his Ph.D. degree in Electrical Engineering from the KU Leuven, Belgium, in 2013. Currently, he is a professor and holds the Chair for Computer Science IV (Computer Vision) at the University of Wurzburg, Germany. Also, he is a lecturer and a group leader at ETH Zurich, Switzerland. He is a member of the editorial board of top journals such as IEEE TPAMI, Elsevier's CVIU and NEUCOM, and SIAM's SIIMS. He regularly serves as an area chair and as a reviewer for top conferences such as CVPR, ICCV, IJCAI, and ECCV. His work received several awards. Radu Timofte is the 2022 awardee of the Alexander von Humboldt Professorship for Artificial Intelligence. He is a co-founder of Merantix and a co-organiser of NTIRE, CLIC, AIM, Mobile AI, and PIRM workshops and challenges. His current research interests include deep learning, mobile AI, visual tracking, computational photography, and image/video compression, restoration, enhancement, and manipulation.

3.5 | Yulan Guo

Yulan Guo received the B.E. and Ph.D. degrees from NUDT in 2008 and 2015, respectively. He has authored over 100 articles at highly referred journals and conferences. His current research interests focus on 3D vision, particularly on 3D feature learning, 3D modelling, 3D object recognition, and scene understanding. He served as an associate editor for IEEE Transactions on Image Processing, IET Computer Vision, IET Image Processing, and Computers & Graphics. He also served as an area chair for CVPR 2023/2021, ICCV 2021, and ACM Multimedia 2021. He organised several tutorials, workshops, and challenges in prestigious conferences, such as CVPR 2016, CVPR 2019, ICCV 2021, 3DV 2021, CVPR 2022, ICPR 2022, and ECCV 2022. He is a senior member of IEEE and ACM.

KEYWORDS
computer vision, image processing

Longguang Wang[1]
Juncheng Li[2]
Naoto Yokoya[3]
Radu Timofte[4]
Yulan Guo[5]

[1]*National University of Defense Technology, Changsha, China*
[2]*Shanghai University, Shanghai, China*
[3]*The University of Tokyo, Bunkyo-ku, China*
[4]*University of Würzburg, Würzburg, Germany*
[5]*National University of Defense Technology & Sun Yat-sen University, Shenzhen, China*

Correspondence
Longguang Wang.
Email: wanglongguang15@nudt.edu.cn

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data were created or analysed in this study.

ORCID

Longguang Wang https://orcid.org/0000-0003-0429-0263

AUTHOR BIOGRAPHIES

Longguang Wang received his B.E. and Ph.D. degrees from Shandong University and National University of Defense Technology in 2015 and 2022, respectively. He is currently an assistant professor with Aviation University of Air Force. He authored more than 60 peer reviewed journal and conference publications (including TPAMI, TIP, CVPR, ICCV and ECCV). He served as a reviewer for more than 10 international journals (including TPAMI and TIP) and conferences (including CVPR, ICCV and ECCV). He has organized workshops at CVPR 2022/2023/2024. His research interests include low-level vision and 3D vision, particularly on image restoration, image generation, point cloud understanding, and network acceleration. His received the CSIG Excellent Doctoral Dissertation Nomination Award in 2022 (17 nationalwide).

Juncheng Li received the Ph.D. degree from the School of Computer Science and Technology, East China Normal University, in 2021. He also worked as a Postdoctoral Fellow at the Center for Mathematical Artificial Intelligence, The Chinese University of Hong Kong. He is currently an assistant professor with Shanghai University. His research interests include artificial intelligence and its applications to computer vision (e.g. image segmentation) and image processing (e.g. image super-resolution, image denoising, and image dehazing). He has published more than 25 papers in top journals and conferences, including TIP, TNNLS, TMM, ECCV, ICCV, AAAI, ACMMM and IJCAI. He also received several premium awards, including the Shanghai Outstanding Ph.D. Graduates, CUHK Research Fellowship Scheme, the winner of 2019 ICCV-AIM, etc. Meanwhile, he served as a reviewer for more than 20 international journals and conferences.

Naoto Yokoya received the M.Eng. and Ph.D. degrees from the Department of Aeronautics and Astronautics, The University of Tokyo, Tokyo, Japan, in 2010 and 2013, respectively. From 2013 to 2017, he was an Assistant Professor with The University of Tokyo. From 2015 to 2017, he was an Alexander von Humboldt Fellow, working at the German Aerospace Center, Oberpfaffenhofen, Germany, and at the Technical University of Munich, Munich, Germany. He is currently a Lecturer with The University of Tokyo, and a Unit Leader with the RIKEN Center for Advanced Intelligence Project, Tokyo, where he leads the Geoinformatics Unit. His research interests include image processing, data fusion, and machine learning for understanding remote sensing images, with applications to disaster management. Dr. Yokoya received the First Place in the 2017 IEEE Geoscience and Remote Sensing Society (GRSS) Data Fusion Contest organized by the IEEE Image Analysis and Data Fusion Technical Committee (IADF TC). From 2019 to 2021, he was the Chair and the Co-Chair (2017–2019) of the IEEE GRSS IADF TC. Since 2018, he has been an Associate Editor of IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing (JSTARS).

Radu Timofte received his Ph.D. degree in Electrical Engineering from the KU Leuven, Belgium, in 2013. Currently, he is a professor and holds the Chair for Computer Science IV (Computer Vision) at theUniversity of Wurzburg, Germany. He is a member of the editorial board of top journals such as IEEE TPAMI, Elsevier's CVIU and NEUCOM, and SIAM's SIIMS. He regularly serves as an area chair and as a reviewer for top conferences such as CVPR, ICCV, IJCAI and ECCV. His work received several awards. Radu Timofte is the 2022 awardee of an Alexandervon Humboldt Professorship for Artificial Intelligence. He is a co-founder of Merantix and a co-organizer of NTIRE, CLIC, AIM, Mobile AI and PIRM workshops and challenges. His current research interests include deep learning, mobile AI, visual tracking, computational photography, image/video compression, restoration, enhancement and manipulation.

Yulan Guo received the B.E. and Ph.D. degrees from National University of Defense Technology (NUDT) in 2008 and 2015, respectively. He has authored over 100 articles at highly referred journals and conferences. His current research interests focus on 3D vision, particularly on 3D feature learning, 3D modeling, 3D object recognition, and scene understanding. He served as an associate editor for IEEE Transactions on Image Processing, IET Computer Vision, IET Image Processing, and Computers & Graphics. He also served as an area chair for CVPR 2023/2021, ICCV 2021, and ACM Multimedia 2021. He organized several tutorials, workshops, and challenges in prestigious conferences, such as CVPR 2016, CVPR 2019, ICCV 2021, 3DV 2021, CVPR 2022, ICPR 2022 and ECCV 2022. He is a Senior Member of IEEE and ACM.