www.ipmv.org | www.icesp.org

CONFERENCE PROGRAM

IPMV 2023

2023 5th International Conference on
IMAGE PROCESSING AND MACHINE VISION

ICESP 2023

2023 4th International Conference on
Electronics and Signal Processing

January 13-15, 2023 | Virtual Conference (UTC+8)

Co-sponsored by

Supported by

Published by

Table of Content

MACAU CONFERENCE 2023 (Virtual)

January 13 (Fri.)-14 (Sat.)

Welcome Message

On behalf of the organizing committee, we warmly welcome you to the joint conference of 2023 5th International Conference on Image Processing and Machine Vision (IPMV) and 2023 4th International Conference on Electronics and Signal Processing (ICESP). The conference is co-sponsored by University of Macau, supported by Wuhan University of Technology, Daegu University, Mahanakorn University of Technology, and etc.

We would like to express our gratitude to the distinguished speakers, Prof. James Tin-Yau Kwok (HKUST, Hong Kong), Prof. Chi Man Pun (UM, Macau), Prof. Hong Lin (UHD, USA), and Prof. LEUNG Andrew C S (CityU, Hong Kong) for giving a speech in the conference. Apart from that, we'd like to extend our thanks to all the authors for your contribution as well as the technical program committee members and external reviewers. Their high competence, enthusiasm, valuable time and expertise knowledge, enabled us to prepare the high-quality final program and helped to make the conference become a successful event. Lastly, we would like to thank all of the conference participants for their contributions which are the foundation of this conference.

We truly hope this conference will provide each one of you with a good platform for networking opportunities and interactions with other delegates from both the academics and industry. At last, we appreciate your participation and support.

With Warmest Regards,

Conference Organizing Committee

IPMV 2023 & ICEPS 2023

Conference Committee

CONFERENCE GENERAL CHAIR

James Tin-Yau Kwok, *The Hong Kong University of Science and Technology, Hong Kong, China*

LOCAL ORGANIZING CHAIR

Chi Man Pun, *University of Macau, Macau, China*

CONFERENCE GENERAL CO-CHAIR

LEUNG Andrew C S, *City Univesity of Hong Kong, Hong Kong, China*

CONFERENCE PROGRAM CHAIR

Kenneth K.Y. Wong, *University of Hong Kong, Hong Kong, China*

Hong Lin, University of Houston-Downtown, USA

CONFERENCE PROGRAM CO-CHAIR

Xiao Zhou, *Wuhan University of Technology, China*

PUBLICITY CHAIR

Seokwon Yeom, *Daegu University Gyeongsan, South Korea*

TECHNICAL PROGRAM COMMITTEE

Aldo André Díaz Salazar, *Federal University of Goiás, Brazil*

Aruna Bhat, *Delhi Technological University, India*

Azizi Abdullah, *University Kebangsaan Malaysia, Malaysia*

Bin Xue, *National University of Defense Technology, China*

Callum Canavan, *University of Cambridge, UK*

Chawalit Benjangkaprasert, *King Mongkut's Institute of Technology Ladkrabang, Thailand*

Daming Feng, *Advanced Micro Devices (AMD) in Austin, USA*

Huigang Wang, *Northwestern Polytechnical University, China*

TECHNICAL PROGRAM COMMITTEE

Ghufran Ahmed, *National University of Computer & Emerging Sciences, Pakistan*

Juan Suardíaz Muro, *Universidad Politécnica de Cartagena, Spain*

Lorna Uden, *Staffordshire University, UK*

Lu Tang, *Southeast University, China*

Luis Gómez, *Universidad de Las Palmas de Gran Canaria, Spain*

Nguyen Ngoc Phien, *Ton Duc Thang University, Vietnam*

Nusrat Sharmin, *Military Institute of Science and Technology, Bangladesh*

Pavlo Maruschak, *Ternopil Ivan Puluj National Technical University, Ukraine*

Sheli Sinha Chaudhuri, *Jadavpur University, India*

Shuai Zhao, *Jinan University, China*

Sompong Liangrocapart, *Mahanakorn University of Technology, Thailand*

Suphongsa Khetkeeree, *Mahanakorn University of Technology, Thailand*

Thaweesak Yingthawornsuk, *King Mongkut's University of Technology Thonburi, Thailand*

Vinh Truong Hoang, *Ho Chi Minh City Open University, Vietnam*

Wanwan Li, *University of South Florida, USA*

Wei Xue, *Anhui University of Technology, China*

Xiantao Jiang, *Shanghai Maritime University, China*

Xiaochun Cheng, *Middlesex University, UK*

Ying-Hao Yu, *National Chung Cheng University, Taiwan, China*

Zeno Geradts, *University of Amsterdam, Netherlands*

Agenda Overview

Friday, January 13, 2023

Meeting ID: 868 9479 2274 **Zoom Link:** https://us02web.zoom.us/j/86894792274

14:00-14:40	14:40-15:30	15:30-16:00
P018	P006	15:30-16:00: **Keynote speaker, listener, session chair, as well as other participants** who are unable to attend at allocated time.
P011	P016	
P014	P017	Note: a) All the presenters are required to join the test on **Friday, Jan. 13**, to ensure the next day meeting runs smoothly. b) We will test control panel including screen sharing, audio, video and “Raise Hand” feature, etc. Please get your presentation slides and computer equipment prepared beforehand. c) Desktop Wallpaper (jpg.) d) Zoom Background (jpg.)
P005	P1001	
P020	P302	
P008	P304	
P010	P306	
P023	P001	
P024	P307	
P003	P310	
/	P022	
/	P013	

Download

Agenda Overview

Saturday, January 14, 2023

Meeting ID: 868 9479 2274 Zoom Link: https://us02web.zoom.us/j/86894792274

SESSION TIME	ACTIVITY	SPEAKER
09:00-	Chairman: TBA	
09:00-09:10	**Opening Speech** by Local Organizing Chair	**Prof. Chi Man Pun** University of Macau, Macau, China
09:10-09:55	**Keynote I** Tear Proteomic Analysis	**Prof. Hong Lin** University of Houston-Downtown, Houston, Texas, USA
09:55-10:40	**Keynote II** Sparse Investment Portfolio Design and Finance Index Tracking	**Prof. LEUNG Andrew C S** City University of Hong Kong, Hong Kong, China
10:40-11:00		*Group Photo & Break Time*
11:00-11:45	**Keynote III** Meta-learning with Many Tasks	**Prof. James Tin-Yau Kwok** The Hong Kong University of Science and Technology, Hong Kong, China
11:45-12:30	**Keynote IV** Privacy Protection in Video Live Streaming	**Prof. Chi Man Pun** University of Macau, Macau, China
12:30-13:30		*Break Time*
13:30-16:00	**Session 1:** Machine Vision and Intelligent Image Processing	P018, P011, P014, P005, P020, P008, P010, P023, P024, P003
16:00-16:30		*Session 1 Group Photo & Break Time*
16:30-19:30	**Session 2:** Image and Signal Processing	P006, P016, P017, P1001, P302, P304, P306, P001, P307, P310, P022, P013
19:30-		*Session 2 Group Photo*

Hong Lin

Professor
University of Houston-Downtown
Houston, Texas, USA

January 14 (Sat.) 09:10-09:55

Hong Lin received his PhD in Computer Science from the University of Science and Technology of China. Before he joined the University of Houston-Downtown (UHD), he was a postdoctoral research associate at Purdue University, and an assistant research officer at the National Research Council, Canada. Dr. Lin is currently a Professor in Computer Science with UHD. His research interests include cognitive intelligence, human-centered computing, parallel/distributed computing, and big data analytics. He is the supervisor of the Grid Computing Lab and a co-founder of the Data Center at UHD. Dr. Lin currently serves as the program director for the Master of Science in Artificial Intelligence program at UHD. Dr. Lin is a senior member of the Association for Computing Machinery (ACM).

Tear Proteomic Analysis

Abstract: Contact lens-related ocular surface complications occur more often in teenagers and young adults. The purpose of this study was to determine changes in tear proteome of young patients wearing glasses (GL), orthokeratology lenses (OK), and soft contact lenses (SCL). Twenty-two young subjects (10-26 years of age) who were established GL, OK, and SCL wearers were recruited. Proteomic data were collected using a data independent acquisition-parallel accumulation serial fragmentation workflow. This research was highlighted by the following findings:

- Using the diaPASEF approach, we have identified 19 protein groups that showed significant differences in abundance in the SCL group compared to the OK group.
- Eighty-two proteins differed in abundance in the tears of children versus young adults in the GL group.
- Many of the significant proteins are related to inflammation and immunity as well as ocular surface homeostasis.

In this research, machine learning models, such as KNN, SVM, and LDA, were able to predict the study groups by analyzing MS data, and evaluate the importance of proteins in prediction. It was verified that machine learning models performed the best in discriminating children and young adult groups.

The research results indicated that tear proteomes were altered by orthokeratology and soft contact wear and age, which warrants further larger-scale study on the ocular surface responses of teenagers and young adults separately to contact lens wear. Overall, this study provides a deep coverage of tear proteome and suggests the need to investigate ocular responses to contact lens wear separately for teenagers and young adults. We also confirmed that tear proteomic analysis is useful in identifying important proteins in diabetes diagnosis.

LEUNG Andrew C S

Professor
City University of Hong Kong
Hong Kong, China

January 14 (Sat.) 09:55-10:40

Chi-Sing Leung received the Ph.D. degree in computer science from the Chinese University of Hong Kong in 1995. He is currently a Professor in the Department of Electronic Engineering, City University of Hong Kong. His research interests include neural computing and computer graphics. He has published over 120 journal papers in the areas of Digital Signal Processing, Neural Networks, and Computer Graphics. In 2005, he received the 2005 IEEE Transactions on Multimedia Prize Paper Award for his paper titled, “The Plenoptic Illumination Function”. He was a member of Organizing Committee of ICONIP2006. He was the Program Chair of ICONIP2009 and ICONIP2012. He is the Associate Editors for IEEE Transactions on Neural Networks and Learning Systems, and Neural Processing Letters. He is/was the guest editors of several journals, including Neural Computing and Applications, Neurocomputing, and Neural Processing Letters. He is a governing board member of the Asian Pacific Neural Network Society (APNNS) and Vice President of APNNS.

Sparse Investment Portfolio Design and Finance Index Tracking

Abstract: Nonnegative sparse recovery under sum-to-one constraint is a special form of linear regression problems, where the solution is required to simultaneously satisfy sparsity, nonnegativity and sum-to-one constraints. It has been the core problem of two applications, namely, sparse index tracking and sparse portfolio design. In this talk, we will discuss several algorithms to solve these problem. In some algorithms, we need to perform a nonconvex projection, the constraint set consists of the $\ell 0$-norm, nonnegativity and sum-to-one constraints. We discuss an efficient algorithm for performing the nonconvex projection and prove that the proposed algorithm produces the optimal solution for the nonconvex projection. Numerical experiments are given to demonstrate the effectiveness of the proposal algorithms.

James Tin-Yau Kwok

Professor, IEEE Fellow
The Hong Kong University of Science and Technology
Hong Kong, China

January 14 (Sat.) 11:00-11:45

Prof. Kwok is a Professor in the Department of Computer Science and Engineering, Hong Kong University of Science and Technology. Pr of. Kwok served/is serving as an Associate Editor for the IEEE Transactions on Neur al Networks and Learning Systems, Neurocomputing, Artificial Intelligence Journal, International Journal of Data Science and Analytics, and on the Editorial Board of Machine Learning. He is also serving / served as Senior Area Chairs of major machine learning / AI conferences including NeurIPS, ICML, ICLR, IJCAI, and as Area Chairs of conferences including AAAI and ECML. Prof Kwok will be the IJCAI-2025 Program Chair.

Meta-learning with Many Tasks

Abstract: Meta-learning accelerates the learning on new tasks by extracting meta-knowledge from historical tasks. Obviously, the larger the number of tasks to learn from, the more meta-knowledge can be learned. However, popular meta-learning algorithms like MAML only learn a globally-shared meta-model. This can be problematic when the task environments are complex, and a single meta-model is insufficient to capture the diversity of meta-knowledge. Moreover, sampling tasks from the large set of tasks in each training iteration also increases variance in the stochastic gradient, resulting in slow convergence. In this talk, we propose to address these problems by structuring the task model parameters into multiple subspaces, so that each subspace represents one type of meta-knowledge. Moreover, variance reduction is incorporated into meta-learning so as to achieve fast convergence. Experiments on various meta-learning tasks demonstrate its effectiveness over state-of-the-art algorithms.

Chi Man Pun

Professor
University of Macau
Macau, China

January 14 (Sat.) 11:45-12:30

Chi-Man Pun received his Ph.D. degree in Computer Science and Engineering from the Chinese University of Hong Kong in 2002. He had served as the Head of the Department of Computer and Information Science from 2014 to 2019, where he is currently a Professor and in charge of the Image Processing and Pattern Recognition Laboratory. Dr. Pun has investigated many externally funded research projects as PI, and has authored/co-authored more than 200 refereed papers in many top-tier journals (including T-PAMI, T-IFS, T-IP, T-DSC, and T-KDE) and conferences (including ACM MM, ECCV, AAAI, ICDE, and VR). He has served as the General Chair/Co-Chair, Program/Local Chair for many international conferences, and SPC/PC member for many top CS conferences such as AAAI, CVPR, ICCV, ECCV, etc. His research interests include Image Processing and Pattern Recognition; Multimedia Information Security, Forensic and Privacy; Adversarial Machine Learning and AI Security, etc.

Privacy Protection in Video Live Streaming

Abstract: With the prevailing of live video streaming, establishing an online pixelation method for privacy-sensitive objects is an urgency. Caused by the inaccurate detection of privacy-sensitive objects, simply migrating the tracking-by-detection structure into the online form will incur problems in target initialization, drifting, and over-pixelation. To cope with the inevitable but impacting detection issue, we propose a novel Privacy-sensitive Objects Pixelation (PsOP) framework for automatic personal privacy filtering during live video streaming. Leveraging pre-trained detection networks, our PsOP is extendable to any potential privacy-sensitive objects pixelation. Employing the embedding networks and the proposed Positioned Incremental Affinity Propagation (PIAP) clustering algorithm as the backbone, our PsOP unifies the pixelation of discriminating and indiscriminating pixelation objects through trajectories generation. In addition to the pixelation accuracy boosting, experiments on the streaming video data we built show that the proposed PsOP can significantly reduce the over-pixelation ratio in privacy-sensitive object pixelation.

Information for Presenter

01

No-Show Policy

A paper not presented or presented by a non-author without prior written approval by the Conference TPC will be removed from the final conference proceedings.

No refund will be approved to authors of those papers.

02

Duration of Presentation

15min each report, including 12min for presentation, and 2-3min for Q&A.

Presenter's certificate will be sent out by email, one week after the meeting.

03

Report File

i. PowerPoint file
ii. PDF file

Please join the meeting at least 10min before your session starts and get your presentation prepared beforehand.

04

"Best Presentation" Award

It will be selected from each session by the session chair.

Please check our official website a week after the meeting for information.

The presenter will receive a certificate of "Best Presentation".

Technical Session

January 14 (Sat.) 13:30-16:00 **Session 1: Machine Vision and Intelligent Image Processing**

MI: 868 9479 2274 Session Chair: TBA

TIME	PAPER ID	ABSTRACT
13:30-13:45	P018	Action Recognition with Non-Uniform Key Frame Selector ***Haohe Li**, Faculty of Electrical Engineering and Computer Science, Ningbo University, China* Abstract—Current approaches for spatiotemporal action recognition have achieved impressive progress, especially in temporal information processing. Meanwhile, the power of spatial information may be underestimated. Thus, a non-uniform key frame selector is proposed to pick the most representative frames according to the relationship between frames along the temporal dimension. Specifically, the reweight high-level frame features are used to generate an importance score sequence, while the key frames, in each temporal section, are selected based on the above scores. Such selected frames have richer semantic information, which has positive impact on the network training. The proposed model is evaluated on two action recognition, namely datasets HMDB51 and UCF101, and promising accuracy improvement is achieved.
13:45-14:00	P011	Semi-supervised Defect Segmentation with Uncertainty-aware Pseudo-labels from Multi-branch Network ***Dejene Mengistu Sime**, University of Electronic Science and Technology of China, China* Abstract—Semi-supervised learning methods have recently gained considerable attention for training deep learning networks with limited labeled samples and additional large label-free samples. Consistency regularization and pseudo-labeling methods are among the most widely used semi-supervised learning methods. However, unreliable pseudo labels will largely limit the model's performance when learning from unlabeled images. To alleviate this problem, we propose uncertainty-rectified pseudo labels generated from dynamically mixing predictions of multiple decoders with a shared encoder network. We estimated the uncertainty as the prediction discrepancy between the average prediction and the output of each decoder head. The estimated uncertainty then guides the consistency training as well as the pseudo-label-based supervision. The proposed method achieves significant performance improvements over the fully supervised baseline and other state-of-the-art semi-supervised segmentation methods on similar labeled data proportions. We also performed an extensive ablation study to demonstrate that the proposed method performs well under various setups.

TIME	PAPER ID	ABSTRACT
14:00-14:15	P014	Human Motion Prediction based on IMUs and MetaFormer ***Tian Xu**, Nanjing University of Science and Technology, China* Abstract—Human motion prediction forecasts future human poses from the histories, which is necessary for all tasks that need human-robot interactions. Currently, almost existing approaches make predictions based on visual observations, while vision-based motion capture (Mocap) systems have a significant limitation, e.g. occlusions. The vision-based Mocap systems will inevitably suffer from the occlusions. The first reason is the deep ambiguity of mapping the single-view observations to the 3D human pose; and then considering the complex environments in the wild, other objects will lead to the missing observations of the subject. Considering these factors, some researchers utilize non-visual systems as alternatives. We propose to utilize inertial measurement units (IMUs) to capture human poses and make predictions. To bump up the accuracy, we propose a novel model based on MetaFormer with spatial MLP and Temporal pooling (SMTPFormer) to learn the structural and temporal relationships. With extensive experiments on both TotalCapture and DIP-IMU, the proposed SMTPFormer has achieved superior accuracy compared with the existing baselines.
14:15-14:30	P005	A Fast CU Partitioning Algorithm Based on Texture Characteristics for VVC ***Wei Li**, Shanghai Maritime University, China* Abstract—Different from the traditional quaternary tree (QT) structure utilized in the previous generation video coding standard H.265/HEVC, a new partition structure named quadtree with nested multi-type tree (QTMT) is applied in the latest codec H.266/VVC. The introduction of QTMT brings in superior encoding performance at the cost of great time-consuming. Therefore, this paper proposes a fast coding unit (CU) partitioning algorithm based on CU texture complexity and texture direction. First, we terminate further splitting of a CU when its texture is judged as simple. Then, we use the gray level co-occurrence matrix (GLCM) to extract the texture direction of the block to decide whether to partition this CU by QT, thus terminating further MT partitions. Finally, a final partition type is selected from the four MT partitions in combination with the multi-level texture complexity and texture direction of the block. The simulation results show that the overall algorithm can significantly reduce the encoding time, while the loss of coding efficiency is reasonably low. In comparison with the reference model, the encoding time is reduced by up to 44.71%, while the BDBR is increased by only 0.84% on average.

TIME	PAPER ID	ABSTRACT
14:30-14:45	P020	Deep 3D-2D Convolutional Neural Networks Combined With Mobinenetv2 For Hyperspectral Image Classification ***Douglas Omwenga Nyabuga**, School of Computing and Informatics, Mount Kenya University, Rwanda* Abstract—Convolutional neural networks (CNNs), one of the most successful models for visual identification, have shown excellent performance outcomes in different visual recognition challenges, attracting much interest in recent years. However, deploying CNN models to hyperspectral imaging (HSI) data continues to be a challenge due to the strongly correlated bands and insufficient training sets. Furthermore, HSI categorization is hugely dependent on spectral-spatial information. Hence, a 2DCNNis a possible technique to analyze these features. However, because of the volume and spectral dimensions, a 3D CNN can be an option but is more computationally expensive. Furthermore, the models underperform in areas with comparable spectrums due to their inability to extract feature maps of high quality. This work, therefore, proposes a 3D/2D CNN combined with the MobineNetV2 model that uses both spectral-spatial feature maps to achieve competitive performance. First, the HIS data cube is split into small overlapping 3-D batches using the principal component analysis (PCA) to get the desired dimensions. These batches are then processed to build 3-D feature maps over many contiguous bands using a 3Dconvolutional kernel function, which retains the spectral properties. The performance of our model is validated using three benchmark HSI data sets (i.e., Pavia University, Indian Pines, and Salinas Scene). The results are then compared with different state-of-the-art (SOTA) methods.
14:45-15:00	P008	Research and application of digital design of Nanjing Ming City Wall “Meridian Gate” ***Yang Cao**, Nanjing Normal University, China* Abstract—Through the case study of 3D interactive design of Nanjing Tongji Gate Wall, this paper explores the application of 3D interactive design in relic visual reconstruction. This paper collects literature on Nanjing City Wall and conducts field research and interviews about Tongji Gate. On the basis of relevant data and literature, 3D interactive technology is used to design the interactive animation system of Tongji Gate. 3D interactive animation contributes to the visual reconstruction of material cultural heritage, the improvement of traditional cultural relics protection and research, the adaptation to communication and promotion in the new media era, the protection of cultural heritage, and the continuous cultural inheritance.

TIME	PAPER ID	ABSTRACT
15:00-15:15	P010	Automatic Gait Gender Classification Using Convolutional Neural Networks ***Lavanya Srinivasan**, Royal Holloway University of London, UK* Abstract—In this study, automatic gait gender classification using convolutional neural networks includes three phases: i) human gait signature generation, ii) which convolves the gait energy images with filters for feature extraction and iii) classified using feed-forward convolutional neural networks. Analysed performance of Gabor and Log Gabor features using classification accuracy. The Log Gabor filter's accuracy was 92.11% for the Normal vs Normal dataset, 74.14% for the Normal vs Bag dataset, 46.55% for the Normal vs Coat dataset, 72.41% for the Normal vs Case dataset and whiles Gabor filter's accuracy was 75% for the Normal vs Normal dataset, 60.34% for the Normal vs Bag dataset 65.52% for the Normal vs Coat dataset and 55.17% for the Normal vs Case dataset.
15:15-15:30	P023	Attention Based BiGRU-2DCNN with Hunger Game Search Technique for Low-Resource Document-Level Sentiment Classification ***Victor Kwaku Agbesi**, University of Electronic Science and Technology of China (UESTC), China* Abstract—Extracting features with deep learning models recorded considerable performance in classifying various sentiments. However, modeling lengthy sentences from low-resource document-level datasets, and examining the semantic relationships between these sentences is challenging. Also, recent document representation modules deployed rarely distinguish between the significance of various sentiment features from general features. To solve these challenges, we exploit a neural network-based feature extraction technique. The technique exploits an attention-based two-layer bidirectional gated recurrent unit with a two-dimension convolutional neural network (AttnBiGRU-2DCNN). Specifically, the Bi-GRU layers extract compositional semantics of the low resource document. These layers generate the semantic feature vector of the sentence and present the documents in a matrix form with a time-step dimension and a feature dimension. Due to the high dimensional multiple features generated, we propose a Hunger Games Search Algorithm to select essential features and subdue unnecessary features to increase classification accuracy. Extensive experiments on two low-resourced datasets indicate that the proposed method captures low-resource sentimental relations and also outperformed known state-of-the-art approaches.

TIME	PAPER ID	ABSTRACT
15:30-15:45	P024	Performance Evaluation of Object Detection Models for Traffic Safety Applications on Edge ***Anilcan Bulut**, Marmara University, Turkey* Abstract—Real-time objection detection is becoming more important and critical in all application areas, including Smart Transport and Smart City. From safety/security to resource efficiency, real-time image processing approaches are used more than ever. On the other hand, low-latency requirements and available resources present challenges. Edge computing integrated with cloud computing minimizes communication delays but requires efficient use of resources due to its limited resources. For example, although deep learning-based object detection methods give very accurate and reliable results, they require high computational power. This overhead reveals a need to implement deep learning models with less complex architectures for edge deployment. In this paper, the performance of evolving deep learning models with their lightweight versions such as YOLOv5-Nano, YOLOX-Nano, YOLOX-Tiny, YOLOv6-Nano, YOLOv6-Tiny, and YOLOv7-Tiny are evaluated on a commercially available edge device. The results show that YOLOv5-Nano and YOLOv6-Nano with their TensorRT versions can provide real-time applicability in approximately 35 milliseconds of inference time. It is also observed that YOLOv6-Tiny gives the highest average precision while YOLOv5-Nano gives the lowest energy consumption when compared to other models.
15:45-16:00	P003	Matching Images from Different Viewpoints with Deep Learning Based on LoFTR and MAGSAC++ ***Liang Tian**, Business-intelligence of Oriental Nations Corporation Ltd, Beijing, China* Abstract—Matching 2D images from different viewpoints plays a crucial role in the fields of Structure-from-Motion and 3Dreconstruction. However, image matching for assorted and unstructured images with a wide variety of viewpoints leads todifficulty for traditional matching methods. In this paper, we propose a Transformer-based feature matching approach tocapture the same physical points of a scene from two images with different viewpoints. The local features of images areextracted by the LoFTR, which is a detector-free deep-learning matching model on the basis of Transformer. The subsequent matching process is realized by the MAGSAC++ estimator, where the matching results are summarized in the fundamental matrix as the model output. By removing image feature points with low confidence scores and applying the test time augmentation, our approach can reach a mean Average Accuracy 0.81340 in the Kaggle competition Image Matching Challenge 2022. This score ranks 45/642 in the competition leaderboard, and can get a silver medal in this competition. Our work could help accelerate the research of generalized methods for Structure-from-Motion and 3D reconstruction, and would potentially deepen the understanding of image feature matching and related fields.

16:00-16:30 Break Time

Technical Session

January 14 (Sat.) 16:30-19:30 **Session 2: Image and Signal Processing**

MI: 868 9479 2274 Session Chair: TBA

TIME	PAPER ID	ABSTRACT
16:30-16:45	P006	VVC Coding Unit Partitioning Decision based on Naive Bayes Theory ***Leiming Duan**, Shanghai Maritime University, China* Abstract—Versatile Video Coding (VVC) is the latest video coding standard, which uses a hybrid codingmodel. VVC achieves 50% bitrate saving compared with High Efficiency Video Coding (HEVC) standard.However, the encoding complexity of VVC is higher. In this work, a fast partition decision algorithm isproposed to reduce the encoding complexity of VVC, and the CU splitting or no splitting is modeled as abinary classification problem based on Naive Bayes theory. This method has good performance and balancesencoding efficiency and encoding complexity. Experimental results show that, compared with the VVCreference software model, the proposed algorithm can reduce encoding time by 48.00%, while the loss of theBD-rate is only 1.69%.
16:45-17:00	P016	Traffic Accident Prediction from Dashcam Videos Based on the UString Model ***Lu-Chun Liu**, National Taipei University of Technology, Taipei, Taiwan* Abstract—For driving applications, such as advanced driver-assistance systems or autonomous driving, it is helpful to provide driving systems with crash information on the roads ahead as early as possible so that safe motion planning can be generated accordingly. In this paper, we provide an empirical study on traffic accident anticipation from dashcam videos based on spatial and temporal relations. In addition, to deal with uncertainties for practical applications, we adopt the UString model, which takes advantage of Bayesian neural networks to consider the probability of accidents. In this study, two public datasets, the Car Crash Dataset (CCD) and the Dashcam Accident Dataset (DAD) are used in model training and testing. Experimental results show that the UString model is more effective in the DAD dataset for ego vehicles not involved in traffic accidents.

TIME	PAPER ID	ABSTRACT
17:00-17:15	P017	Security Analysis of Visual based Share Authentication and Algorithms for Invalid Shares Generation in Malicious Model ***Krishnaraj Bhat**, Sardar Vallabhbhai National Institute of Technology Surat, Gujarat, India* Abstract—Several (k, n) Secret Image Sharing (SIS) schemes supporting authentication of reconstructed secret image and that of shares are proposed in the past. In these existing schemes, two similar state-of-the-art SIS schemes performing visual based share authentication have following merits compared to other schemes: no restriction on values of k and n to be 2, linear time complexity of share authentication, lossless secret image reconstruction, no pixel expansion, and support for share authentication in both dealer participatory and non-participatory environments. In this paper, we show that respective share authentication in these two similar state-of-the-art SIS schemes is computationally insecure in malicious model. We first identify the vulnerabilities in their respective share authentication through security analysis. Then, we propose two linear time algorithms for generating invalid shares from original shares by exploiting the identified vulnerabilities. These generated invalid shares are capable of passing respective authentication in the two analyzed SIS schemes. In addition, usage of a generated invalid share in place of original share during secret image reconstruction results in distorted secret image. Finally, we provide experimental results that accord with inferences of security analysis and linear time complexity of the proposed algorithms for invalid shares generation.
17:15-17:30	P1001	Strategies of Multi-Step-ahead Forecasting for Chaotic Time Series using Autoencoder and LSTM Neural Networks: A Comparative Study ***Nguyen Ngoc Phien**, Ton Duc Thang University, Viet Nam* Abstract—There has been a lot of research on the use of deep neural networks in forecasting time series and chaotic time series data. However, there exist very few works on multi-step ahead forecasting in chaotic time series using deep neural networks. Several strategies that deal with multi-step-ahead forecasting problem have been proposed in literature: recursive (or iterated) strategy, direct strategy, a combination of both the recursive and direct strategies, called DirRec, the Multiple-Input Multiple-Output (MIMO) strategy, and the fifth strategy, called DirMO which combines Direct and MIMO strategies. This paper aims to propose a new deep learning model for chaotic time series forecasting: LSTM-based stacked autoencoder and answer the research question: which strategy for multi-step ahead forecasting using LSTM-based stacked autoencoder yields the best performance for chaotic time series. We evaluated and compared in terms of two performance criteria: Root-Mean-Square Error (RMSE) and Mean-Absolute-Percentage Error (MAPE). The experimental results on synthetic and real-world chaotic time series datasets reveal that MIMO strategy provides the best predictive accuracy for chaotic time series forecasting using LSTM-based stacked autoencoder.

TIME	PAPER ID	ABSTRACT
17:30-17:45	P302	Novel IpDFT Method for Frequency Estimation of the Real-valued Sinusoid Signal with Jumping-changed Parameters ***Tianchang Yan**, Southeast University, China* Abstract— In order to make the system more stable and run more accurately and quickly, frequency estimation is used in many fields, especially in three-phase power system. However, the time-varying signals often suffer from the parameters step change due to external influences which may cause errors in the parameter estimation. In this paper, a novel method using the proposed derived equations is proposed which can eliminate the influence of step-changed parameters on the estimation of frequency in real-valued sinusoid signal. Through MATLAB simulation, it could be seen that the performance of our proposed method is quite good, and the MSEs value can be as low as about -80dB.
17:45-18:00	P304	Joint Angle-Polarization-Range Estimation for EVS-FDA-MIMO Radar ***Tiantian Zhong**, National Laboratory of Radar Signal Processing, Xidian University, China* Abstract—This paper deals with the problem of angle-polarization-range estimation with an electromagnetic vector sensor multiple-input multiple-output radar with frequency diverse array (EVS-FDA-MIMO). The EVS-FDA-MIMO ra-dar can provide additional target polarization information, compared with FDA-MIMO radar; distinguish different targets from the same range bin and angle but different ambiguous range regions, compared with EVS-MIMO ra-dar. First, the FDA-MIMO radar introduces full component EVS spaced a half-wavelength that provides high accuracy and unambiguous estimates of angle-polarization. Then, range estimation is obtained with the estimation of signal parameters via rotational invariance techniques (ESPRIT) algorithm using the rotational invariance of the transmitting array. Furthermore, the parameters pair matching is considered and the complexity of the proposed method is analyzed. Finally, numerical results are presented to validate the effectiveness of the proposed method.

TIME	PAPER ID	ABSTRACT
18:00-18:15	P306	An Enhanced Approach for Environmental Sound Classification using Multi-Window Augmentation ***Krishna Presannakumar**, Mahatma Gandhi University, Kottayam, India* Abstract—The design and deployment of environmental sound classification (ESC) sys-tems face substantial challenges relating to data availability and audio degra-dation brought on by background noises. Deep learning-based ESC systems becomes overfitted while modelling due to a lack of training data. In this pa-per, we propose an ESC model based on an enhanced multi-window spectro-gram augmentation scheme using a gated convolutional neural network. The proposed method continually generates the spectrograms of the audio signal required for training, using multiple windows in terms of size (overlapping values) and type. To address the issues of overfitting and vanishing gradi-ents, the study employs a hybrid model of convolutional neural networks and long short-term memories. To prevent overfitting, the LSTM architec-ture includes a drop-out layer and the features extracted are provided as the input to the convolutional neural networks. The results of the experiment show that our proposed approach produces promising results for ESC prob-lems.
18:15-18:30	P001	A View Direction-Driven Approach for Automatic Room Mapping in Mixed Reality ***Dong Jun Kim**, University of South Florida, USA* Abstract—Virtual Reality and Augmented Reality technologies have greatly improved recently, and developers are trying to make the experience as realistic as possible and close the gap between the physical world and the virtual world. In this paper, we propose an efficient and intuitive method to create an immersive Mixed Reality environment by automatically mapping your room. Our method is view direction driven, which allows the users to simply “look at” any indoor space to create a 3-dimensional model of the area the user is located in. This approach is easier and more intuitive for the users to use and reduces the time and effort compared to other MR environment generating methods. We use the Meta Quest 2’s cameras and gyroscope sensor and the Unity engine for the ray casting and the pass through API. We will present the mathematical details of our method and show that the proposed method achieves better results than previous methods through the user study results.

TIME	PAPER ID	ABSTRACT
18:30-18:45	P307	Two dimensional imaging of sea-surface targets in the Terahertz band ***Xiaofan Li**, National University of Defense Science and technology, China* Abstract—Corner reflectors are often used to simulate the attack of ship targets jam-ming guidance radar. The wavelength of the terahertz band is shorter, and therefore the identification ability of airborne terahertz radar for corner re-flectors is enhanced. In order to study the icosahedral corner reflectors pre-tending to be sea targets, the navigation system can not recognize the real ship targets due to interference with the navigation system, in this paper, the imaging characteristics of a ship and icosahedral reflectors are obtained by turntable imaging, and the large angle convolutional back projection(CBP) fast algorithm is used to image the electrically large-scale sea targets com-posed of a ship and corner reflectors. From the imaging results, the ship and corner reflector targets can be identified under the condition of prior knowledge. Considering the complex surface of ship targets and the layout of interference sources such as multiple types of corner reflectors and alumi-num bars, it can better imitate ship targets and induce the detection of the terahertz radar.
18:45-19:00	P310	Runtime Accuracy Tunable Approximate Floating-Point Multipliers ***Younggyun Cho**, University of Texas, Austin, USA* Abstract—Due to the widespread popularity of data-driven applications such as big data analytics, machine learning and computer vision, modern systems need extremely high computational performance to meet users' requirements. However, the conventional design of arithmetic units is not able to satisfy users with high performance and high energy efficiency simultaneously. Approximate computing is the new paradigm to design arithmetic units for high performance and high energy efficiency. For guarantees on a certain range of error rate from the approximate computing, approximate arithmetic units need an appropriate error estimator. According to our previous research, an error estimator based on ML (Machine Learning) classifiers is a good option since it can foresee the detailed feature of upcoming input data with high accuracy. On the other hand, an error estimator based on ML classifiers requires a training phase, which consumes extra computational power and energy. Besides, for different applications, the error estimator is often required to have another training phase to increase its accuracy. To overcome this shortcoming, we propose runtime accuracy tunable approximate floating-point multipliers in this paper. Our proposed design does not require data profiling, training and re-training phases.

TIME	PAPER ID	ABSTRACT
19:00-19:15	P022	Penetration Point Detection for Autonomous Trench Excavation Based on Binocular Vision *Jian**gying Zhao**, Chang'an University, China* Abstract—To autonomously detect the penetration point in the working area of trench excavation, a feature detection method of penetration point based on binocular cameras was proposed. First, the homogeneous coordinate transformation is established, which can convert the 3D point cloud of the excavation area from the camera coordinate system to the excavator global base coordinate system. Then, the global gradient consistency function is designed to describe the geometric feature of the penetration point of a trench, and the position coordinates of the penetration point are detected. Finally, the test of the penetration point detection of the excavation area is conducted. Within the range of the excavation operation, the maximum position error of the penetration point detection is less than 80 mm, and the average detection error is 46.2 mm, which proves that this method can effectively detect the penetration point.
19:15-19:30	P013	Tracking of Artillery Shell using Optical Flow ***Abdullah Al Masum**, Military Institute of Science Technology, Bangladesh* Abstract—The effectiveness of the Artillery shell is measured by its precision of hit the target. During hitting the target, the intended firing path can be affected by the inherited source as well as external factors like wind, temperature changes, etc. Thus tracking such a path becomes crucial. In this paper, we address the problem of artillery shell tracking with optical flow and adopted the solution for it with sparse and dense optical flow. The outcome of this work can be used in test firing to determine the accuracy of the artillery target. Experimental findings show that the proposed methodology is more effective than the current military artillery unguided shell's impact point locating, even if the shell ended up being an unexplored or blind shell.

Please scan the QR code to get more information.

扫描微信二维码，获取更多会议信息

www.ipmv.org |www.icesp.org