

2020 2nd International Conference on Image, Video and Signal Processing (IVSP 2020)

March 20-22, 2020

Supported by



Indexed by



Scopus

<http://www.ivsp.net/>

Welcome Message from Organizing Committee

It is our great pleasure to invite you to join our international conference - 2020 2nd International Conference on Image, Video and Signal Processing (IVSP 2020). This event will provide a unique opportunity for editors and authors to get together and share their latest research findings and results.

We're confident that over the two days you'll get the theoretical grounding, practical knowledge, and personal contacts that will help you build long-term, profitable and sustainable communication among researchers and practitioners working in a wide variety of scientific areas with a common interest in Video, Image and Signal Processing, Multimedia and Computer Graphics.

On behalf of all the conference committees, we would like to thank all the authors as well as the technical program committee members and reviewers. Their high competence, their enthusiasm, their time and expertise knowledge, enabled us to prepare the high-quality final program and helped to make the conference become a successful event.

We truly hope you'll enjoy the conference and get what you expect from the conference.

Organizing Committee
March 12, 2020

Conference Introductions

Welcome to 2020 IVSP Singapore conference. This conference is organized by ACM Chapter Singapore. The objective of the conference is to provide a platform for researchers, engineers, academicians as well as industrial professionals from all over the world to present their research results and development activities in Video, Image and Signal Processing, Multimedia and Computer Graphics.

Papers will be published in the following proceeding:

International Conference Proceedings Series by ACM (ISBN 978-1-4503-7695-2), which will be archived in the ACM Digital Library, and indexed by Ei Compendex, Scopus and submitted to be reviewed by Thomson Reuters Conference Proceedings Citation Index (ISI Web of Science).

Conference website and email: <http://www.ivsp.net/> and ivsp@acm-sg.org

Table of Contents

Keynote Speakers Introductions.....	1
How to join online conference in Zoom	4
The Meeting ID for Morning Session and Afternoon Session	5
How to rename.....	6
How to chat with others in Zoom.....	7
How to raise your hands and ask questions in Zoom.....	8
How to share your screen	9
How to stop sharing your screen.....	10
Schedule for Conference	11
Morning Session	12
Opening Remarks & Testing (9:30-9:40)	12
Keynote Speech I (9:40-10:25)	12
Keynote Speech II (10:25-11:10)	13
Take a Break: 11:10-11:20	13
Keynote Speech III (11:20-12:05)	14
Lunch 12:05-14:00.....	14
Oral Presentation Abstracts.....	15
Session 1.....	15
ML5017 (14:00-14:15 Singapore local time)	15
ML0005 (14:15-14:30 Singapore local time/15:15-15:30 Japan local time)	16
ML5033 (14:30-14:45 Singapore local time/12:15-12:30 Nepal local time)	16
ML0006 (14:45-15:00 Singapore local time/15:45-16:00 Japan local time)	17
ML5016 (15:00-15:15 Singapore local time)	17
ML5018 (15:15-15:30 Singapore&China local time)	18
ML5032 (15:30-15:45 Singapore&China local time)	18
Session 2.....	19
ML0002 (14:00-14:15 Singapore local time/15:00-15:15 Japan local time)	19
ML0020-A (14:15-14:30 Singapore local time/15:15-15:30 Japan local time)	20
ML5022 (14:30-14:45 Singapore&China local time)	20
ML0015 (14:45-15:00 Singapore local time/13:45-14:00 Thailand local time)	21
ML0016 (15:00-15:15 Singapore local time/10:00-10:15 Saudi Arabia local time)	21
ML5005 (15:15-15:30 Singapore&China local time)	22
ML0003 (15:30-15:45 Singapore& Philippines local time)	22
Take a Break: 15:45-16:00	22

Session 3.....	23
ML0001 (16:00-16:15 Singapore local time/13:30-13:45 Sri Lanka local time)	23
ML5030 (16:15-16:30 Singapore local time/11:15-11:30 Saudi Arabia Local time)	24
ML5021 (16:30-16:45 Singapore& Malaysia local time).....	24
ML0008 (16:45-17:00 Singapore&China local time)	25
ML5008 (17:00 -17: 15 Singapore local time/10:00-10:15 Sweden local time)	25
ML0010-A(17:15 -17: 30 Singapore local time/14:45-15:00 India local time)	26
ML5012 (17:30-17:45 Singapore local time/15:15-15:30 Nepal local time)	26
Session 4.....	27
ML5002 (16:00-16:15 Singapore&China local time)	27
ML5003 (16:15-16:30 Singapore&China local time)	28
ML5011 (16:30-16:45 Singapore local time)	28
ML5027 (16:45-17:00 Singapore&China local time)	29
ML5004 (17:00-17:15 Singapore&China local time)	29
ML5025 (17:15-17:30 Singapore&China local time)	30
ML5024 (17:30-17:45 Singapore&China local time)	30
Video Display.....	31

Keynote Speakers Introductions

Keynote Speaker I



Prof. Kenneth K. M. Lam

The Hong Kong Polytechnic University, Hong Kong

Prof. Kin-Man Lam received his Associateship in Electronic Engineering from the Hong Kong Polytechnic University in 1986. He won the S.L. Poa Education Foundation Scholarship for overseas studies and was awarded an M.Sc. degree in communication engineering from the Department of Electrical Engineering, Imperial College of Science, Technology and Medicine, England, in 1987. In August 1993, he undertook a Ph.D. degree program in the Department of Electrical Engineering at the University of Sydney, Australia. He completed his Ph.D. studies in August 1996.

From 1990 to 1993, Prof. Lam was a lecturer at the Department of Electronic Engineering of The Hong Kong Polytechnic University. He joined the Department of Electronic and Information Engineering, The Hong Kong Polytechnic University again as an Assistant Professor in October 1996. He became an Associate Professor in 1999, and is now a Professor. He was actively involved in professional activities. He has been a member of the organizing committee or program committee of many international conferences. In particular, he was a General Chair of the 2012 IEEE International Conference on Signal Processing, Communications, & Computing (ICSPCC 2012), the APSIPA ASC 2015, and the 2017 International Conference on Multimedia and Expo, all which were held in Hong Kong. Prof. Lam was the Chairman of the IEEE Hong Kong Chapter of Signal Processing between 2006 and 2008. He was an Associate Editor of IEEE Trans. on Image Processing from 2009 to 2014. He received an Honorable Mention of the Annual Pattern Recognition Society Award for an outstanding contribution to the Pattern Recognition Journal in 2004. In 2008, he also received the Best Paper Award at the International Conference on Neural Networks and Signal Processing.

Prof. Lam was the Director-Student Services and the Director-Membership Services of the IEEE Signal Processing Society between 2012 and 2014, and between 2015 and 2017, respectively. He was an Associate Editor of IEEE Trans. on Image Processing between 2009 and 2014, and the Digital Signal Processing Journal between 2013 and 2018, an Area Editor of the IEEE Signal Processing Magazine between 2015 and 2017, and an Editor of HKIE Transactions between 2015 and 2018. Currently, he is the VP-Publications of the Asia-Pacific Signal and Information Processing Association (APSIPA). Prof. Lam serves as an Associate Editor APSIPA Trans. on Signal and Information Processing, and EURASIP International Journal on Image and Video Processing. His current research interests include human face recognition, image and video processing, and computer vision.

Keynote Speaker II



Prof. Xudong Jiang

Nanyang Technological University, Singapore

Prof. Xudong Jiang received the B.Sc. and M.Sc. degree from the University of Electronic Science and Technology of China, in 1983 and 1986, respectively, and received the Ph.D. degree from Helmut Schmidt University Hamburg, Germany in 1997, all in electrical and electronic engineering. From 1986 to 1993, he worked as Lecturer at the University of Electronic Science and Technology of China where he received two Science and Technology Awards from the Ministry for Electronic Industry of China. He was a recipient of the German Konrad-Adenauer Foundation young scientist scholarship. From 1993 to 1997, he was with Helmut Schmidt University Hamburg, Germany as scientific assistant. From 1998 to 2004, He worked with the Institute for Infocomm Research, A*Star, Singapore, as Senior Research Fellow, Lead Scientist and appointed as the Head of Biometrics Laboratory where he developed a fingerprint verification algorithm that achieved the fastest and the second most accurate fingerprint verification in the International Fingerprint Verification Competition (FVC2000). He joined Nanyang Technological University, Singapore as a faculty member in 2004 and served as the Director of the Centre for Information Security from 2005 to 2011. Currently, Dr Jiang is a tenured Associate Professor in School of Electrical and Electronic Engineering, Nanyang Technological University. Dr Jiang has published over hundred research papers in international refereed journals and conferences, some of which are well cited on Web of Science. He is also an inventor of 7 patents (3 US patents), some of which were commercialized. Dr Jiang is a senior member of IEEE and has been serving as Editorial Board Member, Guest Editor and Reviewer of multiple international journals, and serving as Program Committee Chair, Keynote Speaker and Session Chair of multiple international conferences. His research interest includes pattern recognition, computer vision, machine learning, image analysis, signal/image processing, machine learning and biometrics.

Keynote Speaker III

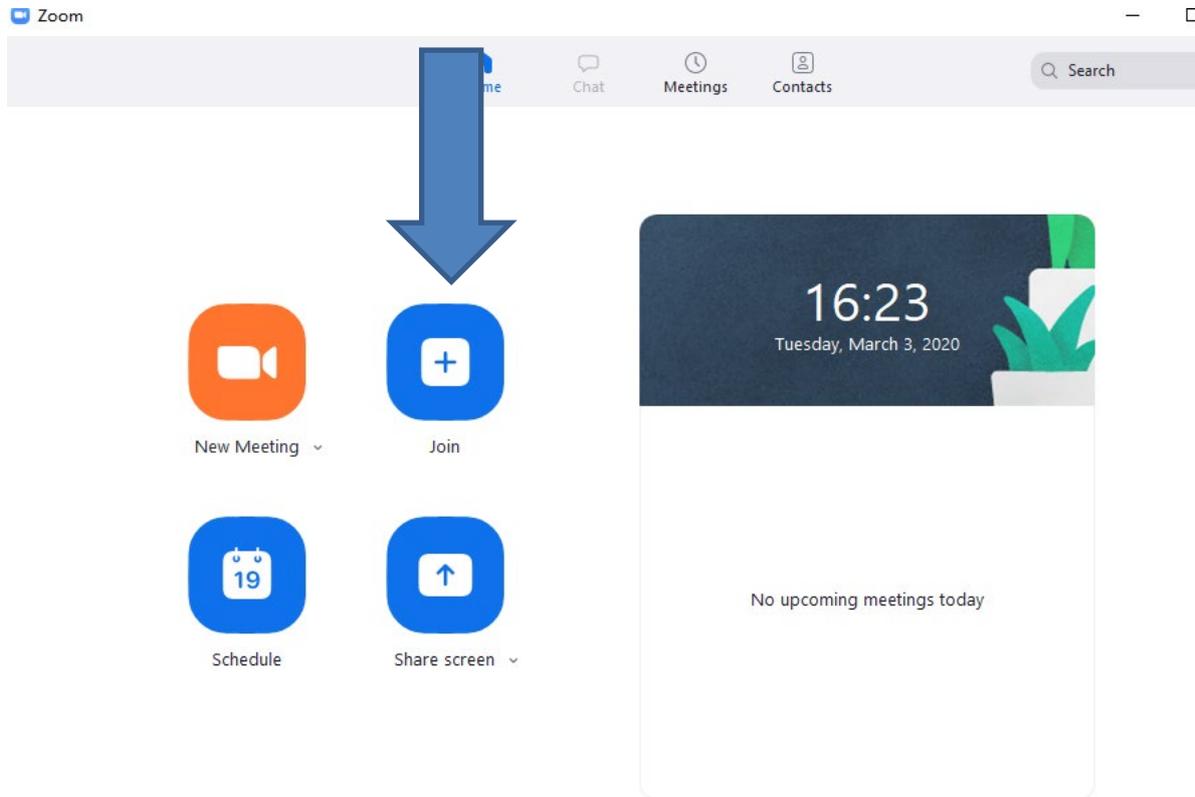


Prof. Kiyoshi Hoshino

Tsukuba University, Japan

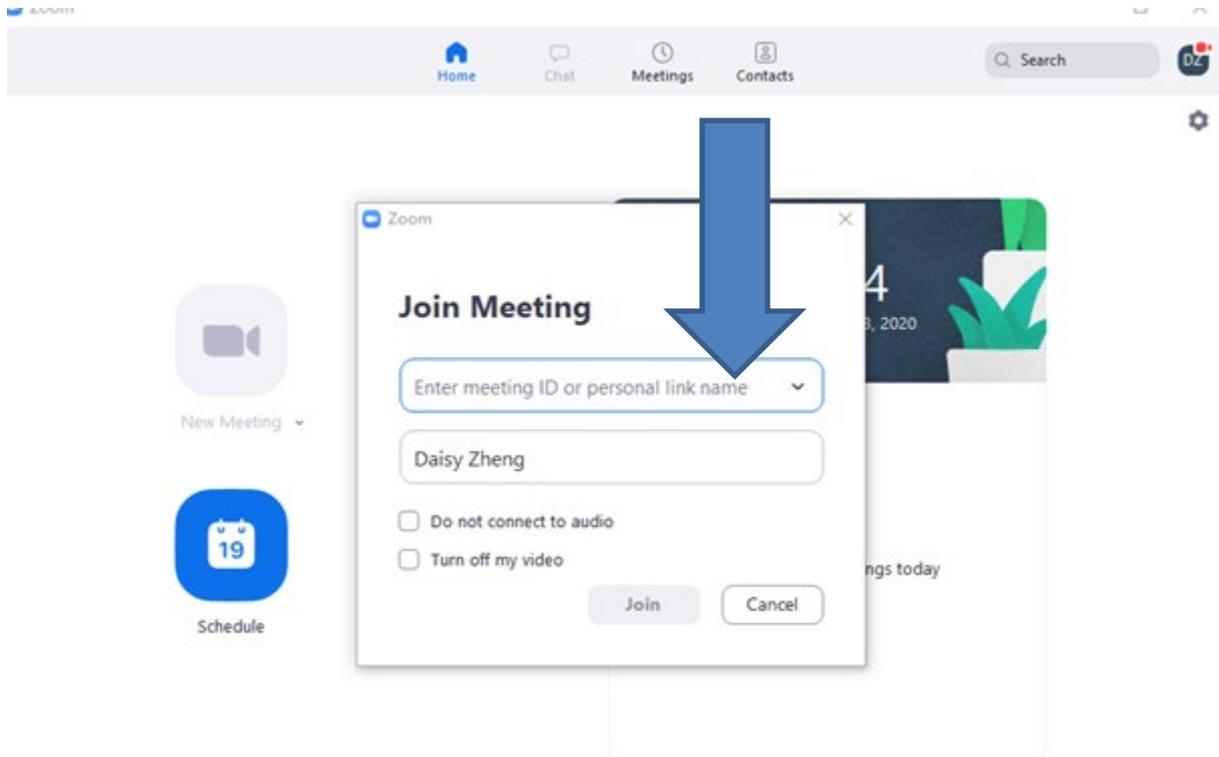
Prof. Kiyoshi Hoshino received two doctor's degrees; one in Medical Science in 1993, and the other in Engineering in 1996, from the University of Tokyo respectively. From 1993 to 1995, he was an assistant professor at Tokyo Medical and Dental University School of Medicine. From 1995 to 2002, he was an associate professor at University of the Ryukyus. From 2002, he was an associate professor at the Biological Cybernetics Lab of University of Tsukuba. He is now a professor. From 1998 to 2001, he was jointly appointed as a senior researcher of the PRESTO "Information and Human Activity" project of the Japan Science and Technology Agency (JST). From 2002 to 2005, he was a project leader of a SORST project of JST. His research interests include biomedical measurement and modelling, medical engineering, motion capture, computer vision, and humanoid robot design.

How to join online conference in Zoom



Please click “join”

The Meeting ID for Morning Session and Afternoon Session



Please fill in the meeting ID and join the online conference.

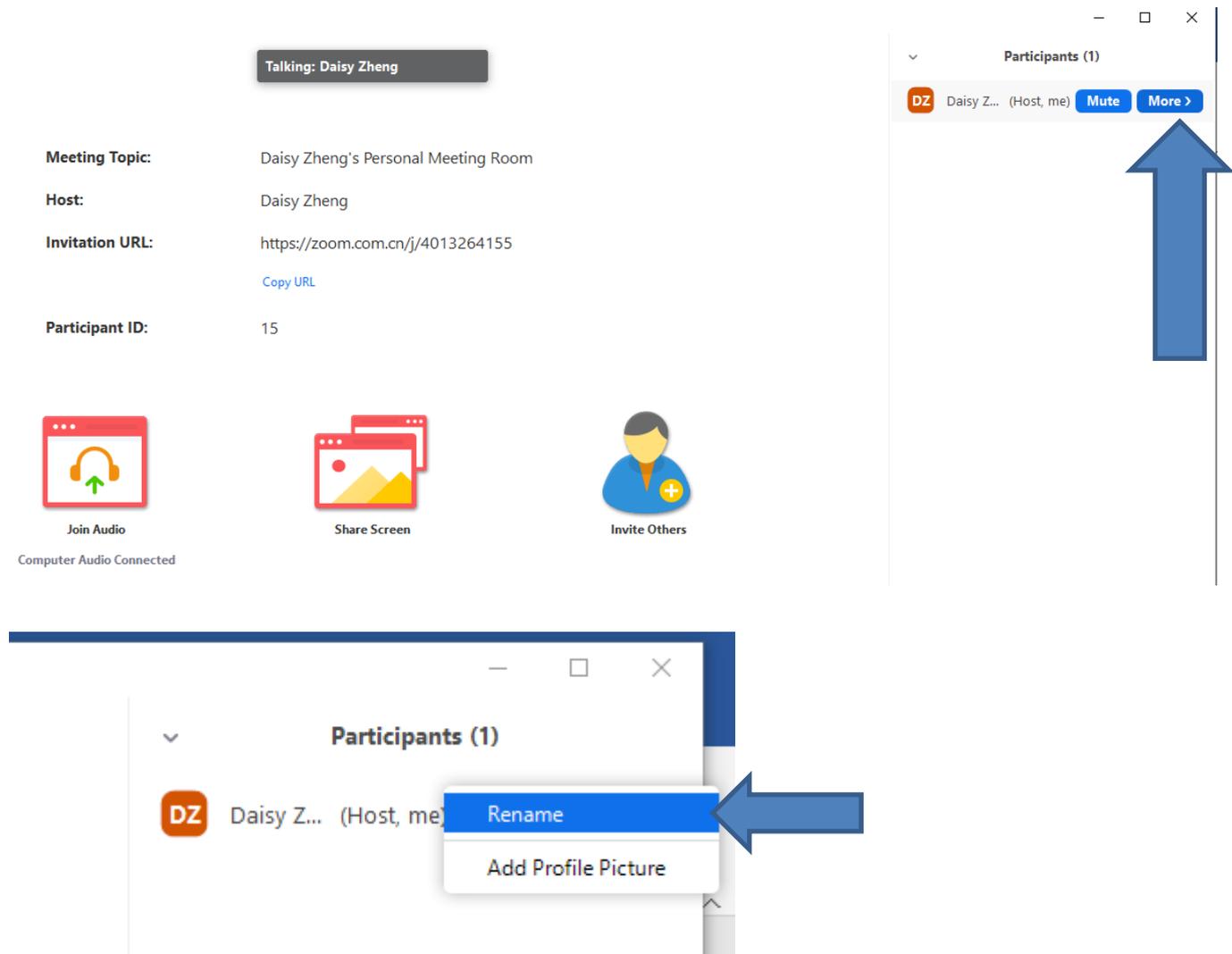
Morning Session: 401-326-4155

Afternoon Session

Session 1 and 3: 401-326-4155

Session 2 and 4: 606-418-2387

How to rename

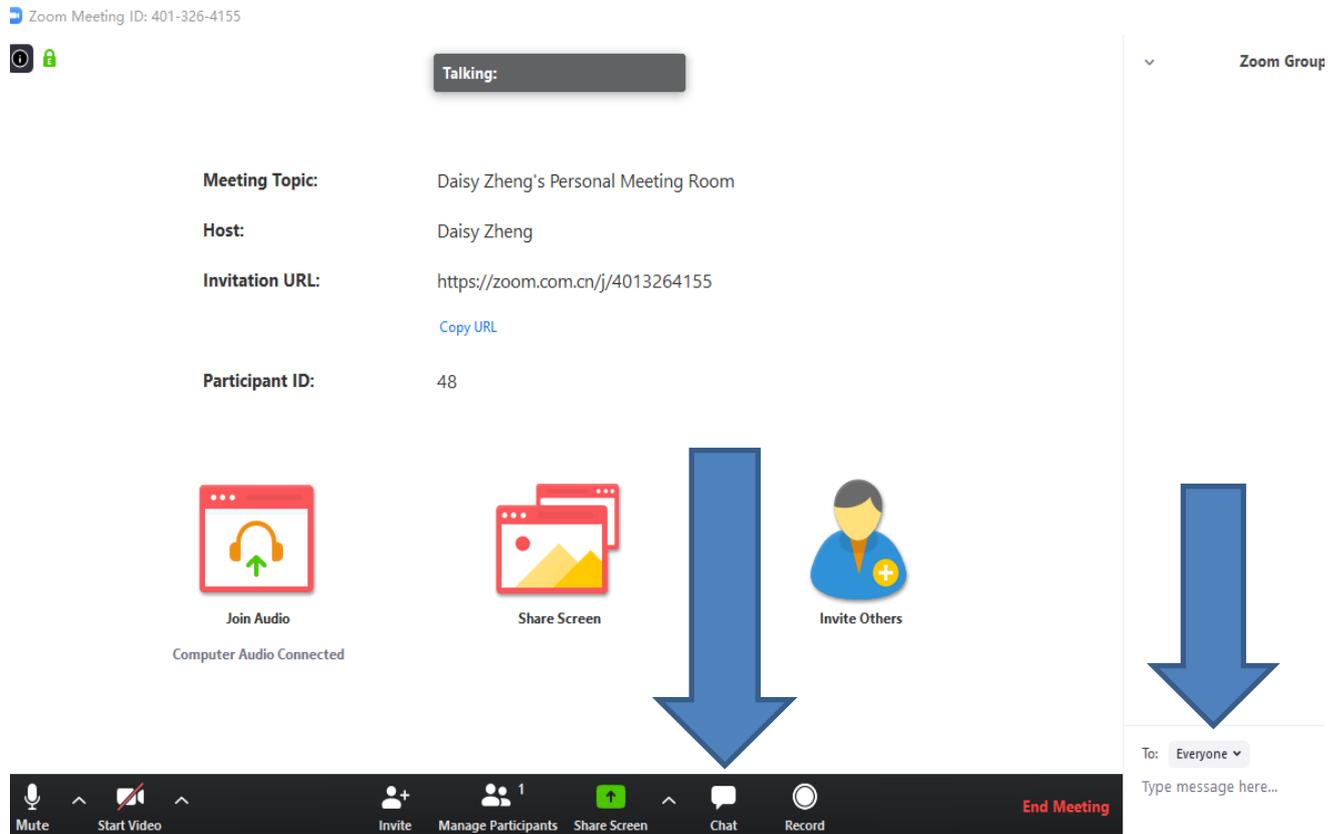


First, you can click “More”

Then, you can click “rename”.

After you join the conference, your conference id need to be renamed as your paper id (MLXXXX) + your name.

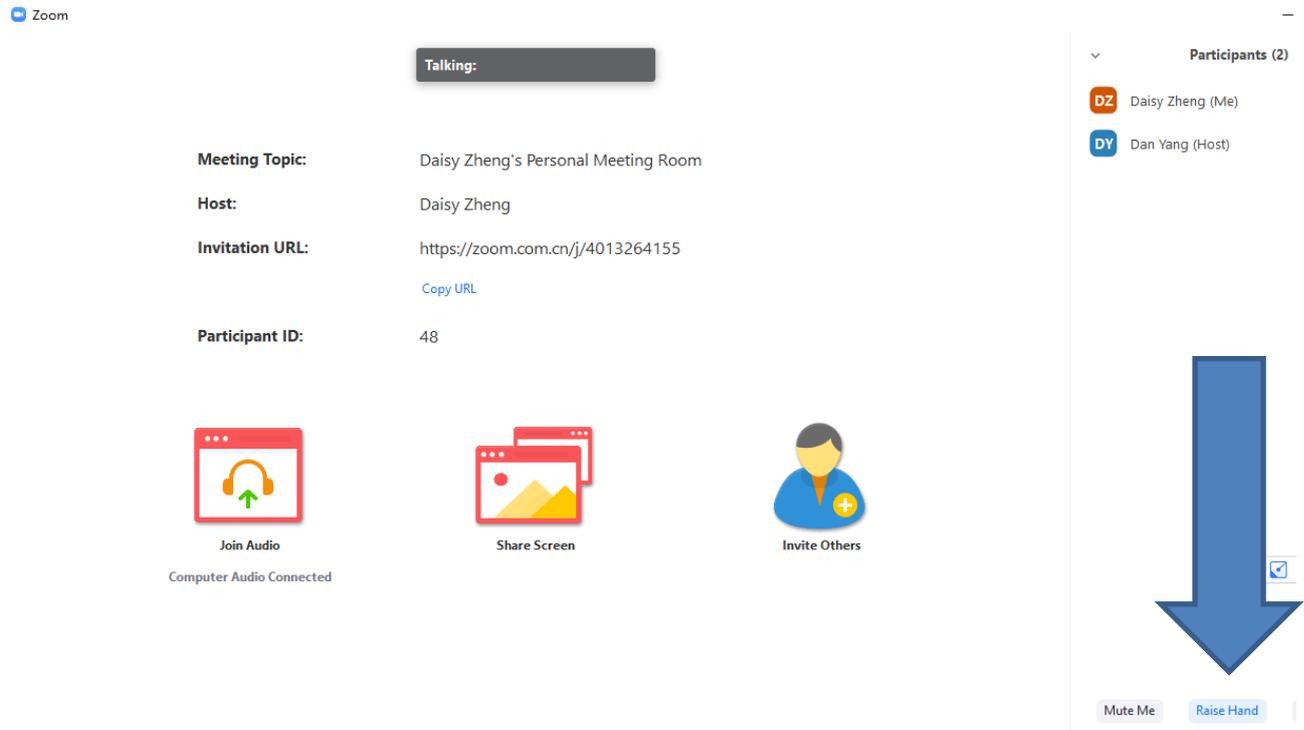
How to chat with others in Zoom



You can click “Chat” first.

Then, you can click “everyone” to choose who you want to talk with.

How to raise your hands and ask questions in Zoom



If you have any problems during the conference, you can click “raise your hands” or use “chat ”to communicate with the conference secretary and the conference secretary will help you.

When you have questions about keynote speeches and author’s oral presentation, you can also use “raise your hands”.

After the keynote speech or each author’s oral presentation, we have Q&A time.

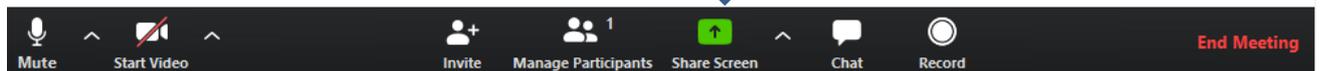
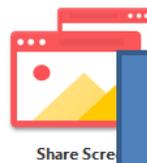
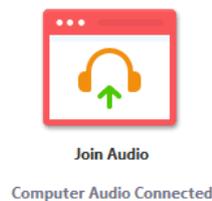
How to share your screen

Zoom Meeting ID: 401-326-4155



Talking:

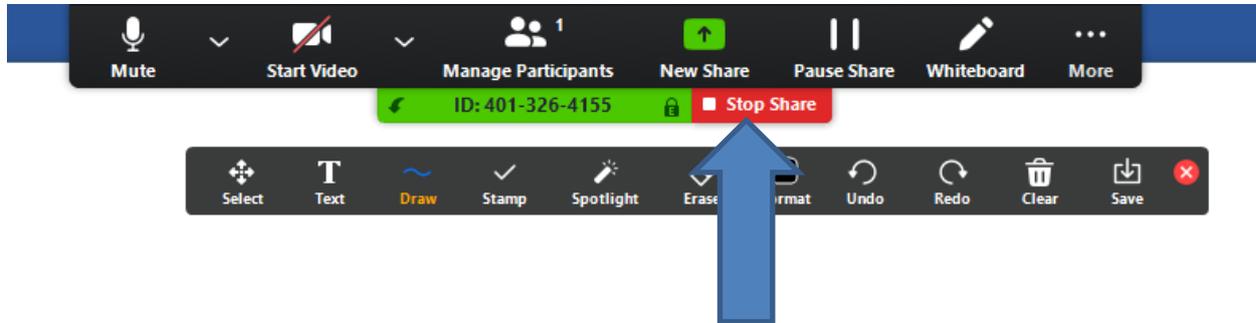
Meeting Topic: Daisy Zheng's Personal Meeting Room
Host: Daisy Zheng
Invitation URL: <https://zoom.com.cn/j/4013264155>
[Copy URL](#)
Participant ID: 40



When you do your presentation, you need to share your screen.

You can click “share screen”.

How to stop sharing your screen



When your oral presentation is over, you can click “stop share”.

Schedule for Conference

Tips: The time in the schedule is according to the local time in Singapore

March 20, 2020 (10:00 - 17:00)	
Online Test	
IVSP Meeting ID for test : 401-326-4155	
March 21, 2020 (9:30 - 12:05) Morning Session Meeting ID: 401-326-4155	
Opening Remark (9:30-9:40) & Testing	
Keynote Speech I (9:40-10:25)	
Title: Subspace Analysis from Conventional to Deep Learning Prof. Kenneth K. M. Lam, The Hong Kong Polytechnic University, Hong Kong	
Keynote Speech II (10:25-11:10)	
Title: Iterative Truncated Arithmetic Mean Filter and Its Properties Prof. Xudong Jiang, Nanyang Technological University, Singapore	
Take a Break (11:10-11:20)	
Keynote Speech III (11:20-12:05)	
Title: Eye Tracking and Measurement of Eye Rotation Prof. Kiyoshi Hoshino, University of Tsukuba, Japan	
Lunch (12:05-14:00)	
March 21, 2020	
Session 1 - Meeting ID: 401-326-4155 (14:00 – 15:45)	Session 2 - Meeting ID: 606-418-2387 (14:00 – 15:45)
Take a Break (15:45-16:00)	
March 21, 2020	
Session 3 - Meeting ID: 401-326-4155 (16:00 – 17:45)	Session 4 - Meeting ID: 606-418-2387 (16:00 – 17:45)
March 22, 2020	
Video Display	

Morning Session

Morning, March 21, 2020 (Saturday)

Time: 9:30~12:05

Online-Meeting ID: 401-326-4155

Opening Remarks & Testing (9:30-9:40)

Keynote Speech I (9:40-10:25)

Title: Subspace Analysis from Conventional to Deep Learning

Prof. Kenneth K. M. Lam

The Hong Kong Polytechnic University, Hong Kong

Abstract— For image recognition, an extensive number of methods have been proposed to overcome the high-dimensionality problem of feature vectors being used. Subspace analysis has been widely used to reduce the feature-vector dimensionality. These methods vary from unsupervised to supervised, and from statistics to graph-theory based. In this talk, the most popular and the state-of-the-art methods for dimensionality reduction, such as PCA, LDA, LPP, etc., are firstly reviewed, and then a new and more efficient subspace-analysis, or manifold-learning, method, namely Soft Locality Preserving Map (SLPM), is presented. All these methods can, in general, project feature vectors into a lower dimensional and more discriminant space, so more efficient and accurate performance can be achieved. With the recent advancements of deep neural networks, recent research has extended the concept of subspace analysis to deep learning. The trained deep networks can achieve a better recognition accuracy. To compare the performance of the conventional subspace-analysis methods and the deep-learning methods, all these methods are applied for facial expression recognition. They are evaluated and compared, using the same setup and databases.

Keynote Speech II (10:25-11:10)**Title: Iterative Truncated Arithmetic Mean Filter and Its Properties****Prof. Xudong Jiang****Nanyang Technological University, Singapore**

Abstract—The arithmetic mean and the order statistical median are two fundamental operations in image processing. They have their own merits and limitations in noise attenuation and image structure preservation. Comparing with the arithmetic operation, data sorting required by the median-based filters is a complex process and is intractable for multivariate data. This talk explores the relation between the two very often used fundamental statistics, namely, the arithmetic mean and the order statistical median. It unveils some simple statistics of a finite data set as the upper bounds of the deviation of the median from the mean. It is desirable to develop a filter having the merits of both the types of filters. The proposed Iterative Truncated arithmetic Mean filter, ITM filter, circumvents the data-sorting process but outputs a result approaching the median. Proper termination of the proposed ITM algorithm enables the filters to own merits of the both mean and median filters and, hence, to outperform both the filters in many image processing applications.

**Take a Break: 11:10-11:20**

Keynote Speech III (11:20-12:05)**Title: Eye Tracking and Measurement of Eye Rotation****Prof. Kiyoshi Hoshino****University of Tsukuba, Japan**

Abstract— In gaze estimation, the dark pupil eye tracking is the usual method. Please imagine when a camera for capturing the eyeball is placed almost directly to the side so that it does not obstruct the user's field of view, estimating the line of sight with the above-mentioned method. Because the eyeball is a sphere, the pupil center coordinate is distorted and this caused the issue of low eye tracking accuracy. The author will introduce a new eye tracking method that does not reduce the accuracy of estimates even when the pupil camera is almost at the side of the eyeball and that requires only a little calibration. In this method, the user is asked to look at a minimum of six points of light that include the corners of a pentagon that is slanted towards the inside corner of the eye and the center of gravity of that pentagon. This is a calibration method that maps the line of sight angles from the pupil center coordinate. The author will also introduce a method to measure rotational eye movement by using a small camera installed almost by the side of the eye in the same manner as the aforementioned eye tracking. This method uses the characteristic images of the blood vessels in the whites of the eyeball. A template image is selected based on three criteria: First, the blood vessel must be thick and have good dark/light contrast; second, there must not be any external light sources reflected; and third, there must not be blood vessels with a similar shape nearby. By continuously tracking the area where the degree of similarity in template matching is maximized, there is little interference from reflections of external light sources so that highly accurate measurement of eye rotation is possible.

**Lunch 12:05-14:00**

Oral Presentation Abstracts

Session 1

Tips: The schedule for each presentation is for reference only. In order not to miss your presentation, we strongly suggest that you attend the whole session.

Afternoon, March 21, 2020 (Saturday)

Time: 14:00-16:00

Online-Meeting ID: 401-326-4155

Session Chair: Dr. Maher Ala'raj, Zayed University, UAE

ML5017 (14:00-14:15 Singapore local time)

A YOLOv3-based Deep Learning Application Research for Condition Monitoring of Rail Thermite Welded Joints

Yang Liu, Xin Sun and John Hock Lye Pang

Nanyang Technological University, Singapore

Abstract—Thermite welded joints are commonly applied in running rails and any failures related to it will result in delay of time schedule or safety issues. The detection and monitoring of these joints are essential for pre-emptive maintenance and therefore to improve the structural integrity of the railway system. This study aims to produce an application prototype to detect and monitor the rail thermite welded joints based on YOLOv3 deep learning algorithm. By using YOLOv3 (You Only Look Once) as the detector, the classification and location boundary box can be determined for each image of thermite welded joint. Experiment on the training and validation of this algorithm has achieved a good result and this application prototype, to be integrated with the proposed camera monitoring system, is capable of detecting and monitoring crucial components of rail track system.

ML0005 (14:15-14:30 Singapore local time/15:15-15:30 Japan local time)

CNN-based Criteria for Classifying Artists by Illustration Style

Kazuma Kondo and Tatsuhito Hasegawa

University of Fukui, Japan

Abstract—In works of art, the drawing style appears in pictures as an artist feature. The style can be applied to identify fake works of art, and this task comes down to identifying the artists of the pictures. Artist classification is defined as a type of image classification task; therefore, it is suitable to use a convolutional neural network (CNN) which achieves high performance in image classification to identify an artist. However, the internal function of the CNN is not clear, and it is difficult to understand how a CNN identifies the artists of pictures. There is a possibility that a CNN identifies artists by not the style of the pictures but by other criteria. In this study, we investigated criteria of the CNN model which is performed end-to-end training. We clarified that a CNN classifies artists using the style of the illustration.

ML5033 (14:30-14:45 Singapore local time/12:15-12:30 Nepal local time)

Disaster Assessment from Satellite Imagery by Analysing Topographical Features Using Deep Learning

Saramsha Dotel, Avishekh Shrestha, Anish Bhusal, Ramesh Pathak, Aman Shakya and Sanjeeb Prasad Panday

Tribhuvan University, Nepal

Abstract—This paper explores the application of deep learning techniques in the task of assessing disaster impact from satellite imagery. Identifying the regions impacted by a disaster is critical for effective mobilization of relief efforts. Satellite images, with their vast coverage of ground surface, are a valuable resource that can be leveraged for this purpose. However, the task of analysing a satellite image to detect regions impacted by disaster is challenging. In recent years, the increasing availability of satellite images of a place presents an opportunity to utilize deep learning on these images to provide a preliminary insight on the impact of a disaster after its occurrence. We particularly focus on water related disasters like floods and hurricanes. To identify the impacted regions, we employ Convolutional Neural Networks to semantically segment topographical features like roads on pre and post-disaster satellite images and find the regions with maximal change. However, this approach is less applicable on rural landscapes due to the sparse distribution of topographical features like roads. To address such cases of imagery from rural areas, we propose a bitemporal image classification approach to compare pre and post disaster scenes and directly identify if the regions are impacted or not. On testing against a ground truth satellite image from DigitalGlobe with labeled regions depicting the impacts of Hurricane Harvey, the flooded road extraction approach achieved an accuracy of 0.845 with a F1-Score of 0.675. Similarly, the bitemporal image classification approach registered an accuracy of 0.94 when tested against a rural landscape impacted by South Asian Monsoon Flooding of 2017.

ML0006 (14:45-15:00 Singapore local time/15:45-16:00 Japan local time)

Fish Species Identification Using A CNN-Based Multimodal Learning Method

Hanano Masuda, Takahiro Jukei and Tatsuhito Hasegawa,
University of Fukui, Japan

Abstract—In this study, we propose an advanced fish species identification method that uses fish images and their meristic characters. Conventionally, fish species are only identified using feature values obtained from images. Because fish of the same species can have different colors or look very similar to other species, it is difficult to identify fish species based only on an image. We developed a multi-input model that features images and, additionally, trait data, aimed at becoming an advanced fish species identification system. We constructed a learning model using fish images and their meristic characters obtained by web scraping and compared its accuracy with a case that only used images. As a result, it was clarified that using two or more kinds of meristic characters provided higher accuracy than using only images.

ML5016 (15:00-15:15 Singapore local time)

A Hybrid Approach for Counting Templates in Images

Maher Ala'raj and Munir Majdalawieh
Zayed University, United Arab Emirates

Abstract—In the research, hybrid algorithm for counting repeated objects in the image is proposed. Proposed algorithm consists of two parts. Template matching sub-algorithm is based on normalized cross correlation function which is widely used in image processing application. Template matching can be used to recognize and/or locate specific objects in an image. Neural network sub-algorithm is needed to filter out false positives that may occur during cross correlation function evaluation.

In the last section of the paper experimental evaluation is carried out to estimate the performance of the proposed template matching algorithm for images of blood microscopy and chamomile field image. In the first case, the task is to count erythrocytes in the blood sample. In the second case, it is needed to count the flowers in the field.

For all 2 datasets we got precise results that coincides with actual number of objects in image. The reason of such performance is that convolutional neural network sub-algorithm improved initial results of template-matching sub-algorithm based on correlation function.

ML5018 (15:15-15:30 Singapore&China local time)

No-reference Image Quality Assessment Based on a Multi-feature Extraction Network

Hainan Zhang, Fang Meng and Yawen Han

Communication University of China, China

Abstract—Deep convolutional neural network (DCNN) has achieved high performance on computer vision. However, it's hard to directly apply to image quality assessment due to lack of enough subjective scores. In this paper, we tackle this problem by exploiting high-level semantic and low-level structural features of the pre-trained VGG16 model for no-reference image quality assessment. We first divide overlapping 224x224 patches to get a fixed input size, then use VGG16 to extract the patch features. We use a three fully-connected layers network to get each patch quality score and average the patch scores to predict the whole image quality score. Experimental results on the benchmark LIVE-II database show our methods are comparable with start-of-art algorithms and simpler than other DCNN feature-based methods.

ML5032 (15:30-15:45 Singapore&China local time)

Improved BLSTM RNN based Accent Speech Recognition Using Multi-task Learning and Accent Embeddings

Wenbi Rao¹, **Ji Zhang**¹ and Jianwei Wu²

1: Wuhan University of Technology, China

2: Wuhan FiberHome Digital Technology Co., Ltd, China

Abstract—A major challenge in Automatic speech recognition (ASR) systems for Mandarin is to be able to handle speakers with different kinds of accents. ASR systems that are trained using single-task learning underperform due to poor generalization ability when confronted with a new accent. In this paper, we explore how to use accent sentences information that accent embeddings and multi-task learning on the basis of the bidirectional long short term memory (BLSTM) to improve accent speech recognition. Firstly we consider augmenting the speech input with accent information in the form of embeddings extracted by a standalone network. Then we propose multi-task learning architecture that we jointly learn an accent classifier and a multi-accent acoustic model. Experiments with these methods demonstrate that we obtain a 4% average relative improvement in word error rate over a multi-accent baseline system

Oral Presentation Abstracts

Session 2

Tips: The schedule for each presentation is for reference only. In order not to miss your presentation, we strongly suggest that you attend the whole session.

Afternoon, March 21, 2020 (Saturday)

Time: 14:00-16:00

Online-Meeting ID: 606-418-2387

Session Chair: Prof. Kikuo Asai, The Open University of Japan, Japan

ML0002 (14:00-14:15 Singapore local time/15:00-15:15 Japan local time)

A Novel Excavation of Fortuitous Tourism Resources Candidates by Analyzing Emotional Information

Reiko Hishiyama and **Yuya Ieiri**

Waseda University, Japan

Abstract—Tourism is one of the most important industries in the world. An excavation of tourism resources is an effective approach to promote the tourism industry, and many excavation methods that have been devised for Fortuitous Tourism Resources Candidates. Therefore, it is necessary to excavate Fortuitous Tourism Resources loped for Potential Tourism Resources that are resources with low name recognition and high attractiveness. However, an excavation method of Fortuitous Tourism Resources that are resources with high attractiveness by chance, due to influences of environmental factors, has not yet been developed. Fortuitous Tourism Resources are extracted from Candidates in order to find new Fortuitous Tourism Resources. In this paper, an excavation method of Fortuitous Tourism Resources Candidates, by analyzing data focusing on differences in familiarity of tourists, is proposed. Furthermore, we conducted a demonstration experiment in Kyoto, Japan, to practice the proposed method. As a result, three Fortuitous Tourism Resources Candidates are excavated using the proposed method, and the effectiveness of the proposed method is shown by the experiment.

ML0020-A (14:15-14:30 Singapore local time/15:15-15:30 Japan local time)

Visualizing Topics Transition of Tweets in Time Series During Disaster

Kikuo Asai¹ and Yuji Sugimoto²

1: The Open University of Japan, Japan

2: Doshisha University, Japan

Abstract—A social networking service (SNS) such as Twitter has been increasingly used as a communication tool during natural disasters. Although time-series data of tweets include important information on understanding people's behavior and thinking, it is a challenging task to interpret a whole tendency of social media data because of change in time as well as volume and variety of the data. This work presents visualization of the topics transition of tweets in time series during a disaster.

First, the topics are retrieved by using a topic modeling technique with a latent dirichlet allocation (LDA) algorithm at certain time intervals. The different topics at each time interval are obtained depending on the disaster situations. Second, the transition of topics are generated by connecting the topics among the different intervals, based on degree of similarity. Third, the transition of topics are visualized with a mapping method based on heat-model and focus-and-context techniques. We believe that the visualization of the topics transition makes it easy to understand the tweets in time series.

ML5022 (14:30-14:45 Singapore&China local time)

A Lossless Compression Algorithm Based on Differential and Canonical Huffman Encoding for Spaceborne Magnetic Data

Li Li, Huijie Liu, Ye Zhu, Xuwen Liang and Lei Liu

Innovation Academy for Microsatellites of CAS, China

Abstract—Lossless compression for spaceborne magnetic data is significant, especially for small, low-cost satellites or deep space exploration spacecrafts. However, most of existing lossless compression algorithms are difficult to adapt to the embedded platform or the existing embedded algorithms have low compression ratio. In light of this, taking full advantage of the characteristics that the spaceborne magnetic data change slowly in a short time, a lossless compression algorithm based on differential encoding and canonical Huffman encoding is proposed. The original magnetic values are replaced with differential values with variable byte length to shorter the length of raw data. Canonical Huffman encoding can be performed twice after differential operation to further improve the compression ratio if the memory is sufficient. Experiments are carried out on spaceborne magnetic data of a satellite on an onboard computer. This algorithm can get compression ratio of 1.76 and 1.83 with canonical Huffman encoding disabled and enabled respectively, and the results demonstrate the superiority of the proposed algorithm in comparison with embedded LZW algorithms.

ML0015 (14:45-15:00 Singapore local time/13:45-14:00 Thailand local time)

Sentence Sentiment Classification using Convolutional Neural Network in Myanmar Texts

Saung Hnin Pwint Oo, Thanaruk Theeramunkong and Nguyen Duy Hung

Technology Thammasat University, Thailand

Abstract—There are still few works on application of deep learning for Myanmar language. This paper presents an approach to use a convolutional neural network (CNN) model to classify sentence sentiment in Myanmar texts. A CNN model is constructed on the top of a word embedding model (i.e., Word2Vec), which converts words into vectors. The model classifies the input sentences and labels each sentence with positive, negative, neutral, unrelated and unreadable sentiments. The model is learnt from 1,152 sentences taken from the customers' reviews of products, which is provided by a telecommunication company. Then, the model is tested on 495 unseen sentences with the result of 86.26% accuracy and 82.58% average f-measure in prediction. The model is compared with the traditional machine learning (ML) classifiers, especially support vector machine (SVM), naïve Bayes (NB), and logistic regression (LR). The model outperforms these classifiers since SVM results in 64.44% accuracy, NB obtains 60.20% accuracy and LR gets 55.15% accuracy.

ML0016 (15:00-15:15 Singapore local time/10:00-10:15 Saudi Arabia local time)

Machine Learning Algorithms for Diamond Price Prediction

Waad Alsuraihi, Kholoud Bawazeer, Ekram Al-hazmi and **Hanan Alghamdi**

King Abdulaziz University, Saudi Arabia

Abstract—Precious stones like diamond are in high demand in the investment market due to their monetary rewards. Thus, it is of utmost importance to the diamond dealers to predict the accurate price. However, the prediction process is difficult due to the wide variation in the diamond stones sizes and characteristics. In this paper, several machine learning algorithms were used to help in predicting diamond price, among them Linear regression, Random forest regression, polynomial regression, Gradient descent and Neural network. After training several models, testing their accuracy and analyzing their results, it turns out that the best of them is the random forest regression.

ML5005 (15:15-15:30 Singapore&China local time)

End-to-End Stereo Matching Network with Local Adaptive Awareness

Chenggang Guo, Dongyi Chen and Zhiqi Huang

University of Electronic Science and Technology of China, China

Abstract—End-to-end training of deep neural networks have been extensively utilized for the problem of stereo matching. Recent efforts mainly focus on the matching cost computation and the cost aggregation with additional guidance. In this paper, we propose a novel end-to-end stereo matching network that explores local adaptive awareness in the disparity upsampling stage and the loss function. An improved hierarchical upsampling module is presented to adaptively learns a disparity refinement function. A shared convolution structure is adopted in the upsampling module to adaptively discover local semantic correlations between image intensity pixels and disparity pixels. The same shared convolution structure is utilized in the proposed loss function to explicitly exploit local patterns among disparity pixels. Experiments on several recent large scale stereo matching benchmarks validate the effectiveness of our proposed network architecture. Comparisons with state-of-the-art networks also show that our network has a comparable accuracy and promising running speed.

ML0003 (15:30-15:45 Singapore& Philippines local time)

Real- Time Human Sitting Position Recognition using Wireless Sensors

Jheanel E. Estrada¹ and Larry A. Veal²

1: Technological Institute of the Philippines, Philippines

2: Mapua University, Philippines

Abstract—Human Factors denotes the application of psychological and physiological standards and theories that lead to the design of products, processes and systems. With that Human- Computer Interaction gain attention in the field of Healthcare, Education and Occupational Areas. This paper aims to develop a model that will recognize proper and improper sitting posture in front of a laptop using Digital Signal Processing and User- Modeling. This study used flex sensors placed in the upper areas of the body. In a go signal of the Licensed Physical Therapist, these flex sensors will capture the resistance of each point and sent to the server. Using Machine Learning Techniques, the study was able to develop ergonomic rules using some key features such as the distance of the table and chair, adjustable table height suited for each body frame and upper body points such as left trapezius muscle, left and right deltoids and stylomastid process. This gives way to the development of a more advanced computational systems that will prevent musculoskeletal disorders.

**Take a Break: 15:45-16:00**

Oral Presentation Abstracts

Session 3

Tips: The schedule for each presentation is for reference only. In order not to miss your presentation, we strongly suggest that you attend the whole session.

Afternoon, March 21, 2020 (Saturday)

Time: 16: 00-17:45

Online-Meeting ID: 401-326-4155

ML0001 (16:00-16:15 Singapore local time/13:30-13:45 Sri Lanka local time)

Analysing the Compatibility of Identifying Emotions by Facial Expressions and Text Analytics when Using Mobile Devices

Wasura Wattearachchi, Enosha Hettiarachchi and K. P. Hewagamage

University of Colombo School of Computing, Sri Lanka

Abstract—Affective computing is an imperative topic for Human-Computer Interaction, where user emotions and emotional communication can be utilized to improve the usability of a system. Several strategies are available to detect user emotions but it is questionable when identifying the most suitable and compatible strategy which can be used to detect emotions when using mobile devices. Multimodal emotion recognition paves the path to detect emotions by combining two or more strategies in order to identify the most meaningful emotion. Emotion identification through facial expressions and text analytics has given high accuracies but combining them and practically applying them in the context of a mobile environment should be done. Three prototypes were developed using evolutionary prototyping which can detect emotions from facial expressions and text data, using state of the art APIs and SDKs where the base of the prototypes was a keyboard known as “Emotional Keyboard” which is compatible with Android devices. Evaluations of Prototype 1 and 2 have been done based on participatory design and reviewed the compatibility of emotion identification through facial expressions and text data in the mobile context. Evaluation of Prototype 3 should be done in the future and a confusion matrix should be built to verify the accuracies by cross-checking with training and validation accuracies that have been obtained when developing the neural network.

ML5030 (16:15-16:30 Singapore local time/11:15-11:30 Saudi Arabia Local time)

Local Ensemble Approach for Meningioma Tumor Firmness Detection in MRI Images

Mohamed Maher Ben Ismail

King Saud University, Saudi Arabia

Abstract—Determining if a Meningioma, which is a primary tumor of the brain, is firm or not guides patient counseling and the operative strategy to be adopted. In this paper we propose a novel approach for automatic firmness detection for Meningioma tumor. The proposed framework relies on a possibilistic based ensemble method to classify MRI image instances as firm or non-firm cases. The proposed algorithm minimizes some objective function which combines context identification and multi-algorithm ensemble criteria. The optimization is intended to learn contexts as compact clusters in subspaces of the visual feature space via possibilistic clustering and feature discrimination. The yields optimal ensemble parameters for each context. The experiments were conducted and assessed using a real dataset. The obtained performance proved that the proposed approach outperforms relevant state of the art methods.

ML5021 (16:30-16:45 Singapore & Malaysia local time)

Fractional Active Contour Model for Edge Detector on Medical Image Segmentation

Norshaliza Kamaruddin, Nurazeen Maarop and Ganthan Narayana

University Technology Malaysia, Malaysia

Abstract—In computer terms, segmentation is a process to partition or to divide an image based on the number of objects within the image. The process of segmentation can be easy depending on the quality of the image such as the level of noise, the image contrast and etc. Segmentation on medical images has its own importance such as to extract an importance object like tumor or others, to assist the physician in making decision for surgery purposes. Currently many methods have been developed but to get accuracy in segmenting multi modalities of medical images are still remain unsuccessful. Among all methods, Active Contour Model shows good potential in medical image segmentation. But, accuracy in detecting edges along the object boundary is still remain unsuccessful. The used of fractional calculus that act as the first order integer to extract the missing pixels along the object boundary is seen to have the potential in solving the problem. This paper proposed a method called, Fractional Active Contour (FAC) model. The proposed method tends to highlight the role of fractional as an operator to detect and preserve the missing edges as well as giving the bending capability to the contour of the model. Experiments on several medical images from MRI, CT Scan and X-ray images demonstrates that the proposed FAC with the usage of the powerful fractional calculus as the edge detector model realizes an accurate boundary segmentation although under the constraint of missing edges within the environment of intensity inhomogeneity.

ML0008 (16:45-17:00 Singapore&China local time)

Study on the Identification of Irritability Emotion Based on the Percentage Change in Pupil Size

Minrui Zhao, Hongni Gao, Wei Wang, Jue Qu and Long Chen

Air Force Engineering University, China

Abstract—In order to explore the percentage change in pupil size that can represent the generation of irritability, the experiment of inducing irritability was designed. The pupil diameter of the participants was collected by eye movement instrument, and the data of percentage change in pupil size at different levels of irritability was analyzed by combining the subjective score table. The results showed that the percentage change in pupil size was different in different participants when the irritability level reached level 5, but generally ranged between 10% and 25%. Besides, with the constant deepening of the level of irritability, the percentage change in pupil size is increasing. Therefore, the percentage change in pupil size can be used as a measure of irritability, providing certain reference for the monitoring of irritability of operators in special positions.

ML5008 (17:00 -17: 15 Singapore local time/10:00-10:15 Sweden local time)

Condition Monitoring for Confined Industrial Process Based on Infrared Images by Using Deep Neural Network and Variants

Yuchong Zhang and Morten Fjeld

Chalmers University of Technology, Sweden

Abstract—Some industrial processes take place in confined settings only observable by sensors, e.g. infrared (IR) cameras. Drying processes take place while a material is transported by means of a conveyor through a “black box” equipped with internal IR cameras. While such sensors deliver data at high rates, this is beyond what human operators can analyze and calls for automation. Inspired by numerous implementations monitoring techniques that analyse IR images using deep learning, this paper shows how they can be applied to the confined microwave drying of porous foams, with benchmarking their effectiveness at condition monitoring to conduct fault detection. Convolutional neural networks, derived transfer learning, and deep residual neural network methods are already regarded as cutting-edge and are studied here, using a set of conventional approaches for comparative evaluation. Our comparison shows that state-of-the-art deep learning techniques significantly benefit condition monitoring, providing an increase in fault finding accuracy of up to 48% over conventional methods. Nevertheless, we also found that derived transfer learning and deep residual network techniques do not in our case yield increased performance over normal convolutional neural networks.

ML0010-A(17:15 -17: 30 Singapore local time/14:45-15:00 India local time)

Concept Drift Detection in Evolving Sensor Data Streams for Human Activity Recognition

Shikha Verma and Arnab Kumar Laha

Indian Institute of Management, Ahmedabad, India

Abstract—The pervasive presence of sensors in the Internet of Things produces information-rich, high-velocity data streams. These streams operate in dynamic environments and does not meet the requirements of a (time) stationary distribution which is often an important requirement for analysis of time series data. Such a phenomenon is known as concept drift and streaming data methods are expected to account for the existence of concept drift. Other challenges faced while mining streaming data include single pass at observations, need for concise representations due to limited disk space storage and limited/delayed access to ground truth labels for real life classification tasks.

Considering these challenges of streaming data, in this paper we introduce a novel way for concept drift detection using probabilistic classifier outputs involving metrics derived from the theory of optimal transport. An algorithm for active drift detection in data streams using sliding window approach under limited access to ground truth labels is proposed and tested with simulated as well as real-life human activity recognition data. While most of the papers in existing literature on human activity recognition in data streams deal entirely with unlabeled data our approach tries to leverage the existence of the labels in the training set in the drift detection process.

ML5012 (17:30-17:45 Singapore local time/15:15-15:30 Nepal local time)

Face Recognition Based on Shallow Convolutional Neural Network Classifier

Roshan Shrestha¹ and Sanjeeb Prasad Panday

1: Pokhara University, Nepal

2: Tribhuvan University, Nepal

Abstract—Recent works in semantic segmentation have focused on exploiting richer context through attention mechanisms or multi-scale fusion. Most of them ignore the low-level details, which greatly improve the recognition of pixels in the edge of the object. In this paper, we improve the quality of feature maps from two aspects: muti-level features fusion and global context encoding. A Context Attention Unit (CAU) is proposed to aggregate context information. The CAU utilizes self-attention model to enhance the semantic representation of high-level features. To make effective use of the different characteristics of multi-level feature maps, we further present a Gated Fuse Unit (GFU). A gate is calculated to control the information transfer of the feature maps of the adjacent layers. Base on these two units, a network with encoder-decoder structure is designed to gradually refine the feature map. Then the refined feature map generates the final segmentation prediction. We conduct extensive experiments on popular semantic segmentation benchmarks including Cityscapes and PASCAL VOC, which shows the effectiveness of our method.

Oral Presentation Abstracts

Session 4

Tips: The schedule for each presentation is for reference only. In order not to miss your presentation, we strongly suggest that you attend the whole session.

Afternoon, March 21, 2020 (Saturday)

Time: 16:00-17:30

Online-Meeting ID: 606-418-2387

ML5002 (16:00-16:15 Singapore&China local time)

Circular Template Matching Based on Improved Ring Projection Method

Songlin Bi, Yonggang Gu, Zhihong Zhang, Chao Zhai and Ming Gong

University of Science and Technology of China, China

Abstract— Circular template as matched target is universally applied in template matching and visual measurement as its isotropy, accurate positioning and other characteristics. Therefore, accurately establishing the correspondence between circular templates from different views is important. However, the circular template degenerates into an oval after perspective projection transformation and loses isotropy. Traditional ring projection method (RPM) is robust only to rotation and noise, establishing the correspondences is difficult. In this paper, an improved RPM, which combines circular template normalization and RPM, can be used to match circular templates of multi-view images. The algorithm is robust to scale, translation, shear, illumination, rotation. The robustness and high matching success rate are confirmed by a set of experiments.

ML5003 (16:15-16:30 Singapore&China local time)

An Improved Low-cost Binocular 3D Reconstruction Algorithm Based on Open AI Platform

Zhi Zhong, **Jinjing Yin**, Lei Yu and Mingguang Shan

Harbin Engineering University, China

Abstract—A simple, low-cost and good reconstruction-model method has always been a hot and difficult problem in the field of 3D reconstruction. A binocular 3D reconstruction algorithm based on image object detection technology is proposed. First, two low-cost web cameras are employed to acquire images. After the camera calibration is completed, FAST algorithm is applied to detect the feature points of the two images. Then, the image object is obtained by open AI platform. At the same time, image object is used to determine points: only the feature points located in the object area of the images are described by SIFT. In the next steps, the matching, calculation and triangulation of the world coordinate system are also completed by using these feature points. Finally, OpenGL is utilized to complete the 3D visualization of the object. The simulation result shows that the improved 3D reconstruction algorithm can effectively eliminate the ‘glitch’ caused by mismatching of feature points. The improved algorithm can greatly improve the accuracy of feature point matching without increasing the hardware cost. Feature-point matching accuracy and intuitive 3D reconstruction-model are demonstrated to show the feasibility and validity of this approach.

ML5011 (16:30-16:45 Singapore local time)

Automated Hand Hygiene Compliance Monitoring

Santle Camilus and Young M Lee

Johnson Controls, USA

Abstract—Success of any commercial computer vision solution necessitates the use of right sensors and analytics over the data. RGB image sensors are proved to be efficient for deep learning methods to solve object detection, segmentation and classification problems. Very minimal works are reported in the literature that use other imaging sensors such as depth or infra-red. This work uses a deep learning method over depth sensor data to solve one of the needed solutions of hospitals, battery plants, and restaurants, the Hand Hygiene Compliance Monitoring. Our technical contributions are of three-folds: (1) we evaluate the working of transfer learning using RGB data pre-training and depth data re-training (2) we design a generic strategy for the application as existing solutions in the market are performed under controlled environment (3) finally we compare our deep learning solution performance with a conventional machine learning solution.

ML5027 (16:45-17:00 Singapore&China local time)

CARNet: Context Attention Refine Network for Semantic Segmentation

Guanghai Wang, Song Wu and Guoqiang Xiao

Southwest University, China

Abstract—Recent works in semantic segmentation have focused on exploiting richer context through attention mechanisms or multi-scale fusion. Most of them ignore the low-level details, which greatly improve the recognition of pixels in the edge of the object. In this paper, we improve the quality of feature maps from two aspects: multi-level features fusion and global context encoding. A Context Attention Unit (CAU) is proposed to aggregate context information. The CAU utilizes self-attention model to enhance the semantic representation of high-level features. To make effective use of the different characteristics of multi-level feature maps, we further present a Gated Fuse Unit (GFU). A gate is calculated to control the information transfer of the feature maps of the adjacent layers. Based on these two units, a network with encoder-decoder structure is designed to gradually refine the feature map. Then the refined feature map generates the final segmentation prediction. We conduct extensive experiments on popular semantic segmentation benchmarks including Cityscapes and PASCAL VOC, which shows the effectiveness of our method.

ML5004 (17:00-17:15 Singapore&China local time)

Edge Voxel Erosion for Noise Removal in 3D Point Clouds Collected by Kinect®

Yan Wan¹, Yao Li¹, Jianpeng Jiang¹ and Bugao Xu²

1: Donghua University, China

2: University of North Texas, USA

Abstract—When the human body point cloud is captured by Kinect®, the edge voxels with lower accuracy need to be removed to ensure the precision of the 3D point cloud patches. However, there is no existing erosion algorithm for 3D voxel. This paper proposes a 3D contour voxels erosion algorithm based on 2D image erosion algorithm, which maps the erosion mode of 2D contour pixels to the 3D contour voxels, and eliminating the contour voxels with low accuracy to optimize the 3D human body point cloud.

ML5025 (17:15-17:30 Singapore&China local time)

Object Detection with Enhanced Feature Pyramid Network

Yingda Guan, Zuochang Ye and Yan Wang

Tsinghua University, China

Abstract—Scale variation is one of the most challenging problems in general object detection. Although current approaches have achieved significant progress by exploiting the multi-level information, they pay little attention to how to fuse feature maps and construct the feature pyramid more effectively. In this paper, we propose two novel modules to enhance the characteristics of object detection. First, a Pair-wise Attention Module (PAM) is proposed to introduce the two-way attention mechanism, which can emphasize informative features and filter less useful ones adaptively when fusing feature. Second, a Pyramid Reconfigure Module (PRM) is proposed to promote cross-level spatial information communication by the split-align-reconstruct operation. Then the feature among different levels can be complemented and enhanced with each other. The effectiveness of our proposed modules is evaluated on the COCO benchmark, and experimental results show that our approach achieves state-of-the-art results.

ML5024 (17:30-17:45 Singapore&China local time)

An End-to-end System for Pests and Diseases Identification

Ning Zhang, Zuochang Ye and Yan Wang

Tsinghua University, China

Abstract—The traditional pests and diseases identification methods do not work well for massive high-resolution remote sensing image data. Thus, we are expected to find an efficient way to automatically learn the presentations from the massive image data, and find the relationships among the data. This paper proposes an end-to-end system for pests and diseases identification in massive high-resolution remote sensing data based on deep learning. To achieve good performance on pests and diseases identification, this hierarchical model jointly learns the parameters of a neural network and the cluster assignments of the features. Our network named ClusterNet iteratively groups the features with a standard clustering algorithm k-means, and uses the subsequent assignments as supervision to update the weights of the network. Qualitatively, we only need to provide the remote sensing image of target area, and the system will automatically identify pests and diseases. This is more accurate and convenient compared to the traditional method of manual detection. Quantitatively the resulting model outperforms the traditional convolutional neural networks on our pests and diseases remote sensing dataset.

Video Display

For those authors who were unable to attend the conference the day before due to jet lag or other factors, we offer video display. Welcome to watch IVSP 2020 conference Video!

March 22, 2020 (9:30 - 12:05) Morning Session Meeting ID: 401-326-4155	
Keynote Speech I (9:40-10:25) Title: Subspace Analysis from Conventional to Deep Learning Prof. Kenneth K. M. Lam, The Hong Kong Polytechnic University, Hong Kong	
Keynote Speech II (10:25-11:10) Title: Iterative Truncated Arithmetic Mean Filter and Its Properties Prof. Xudong Jiang, Nanyang Technological University, Singapore	
Take a Break (11:10-11:20)	
Keynote Speech III (11:20-12:05) Title: Eye Tracking and Measurement of Eye Rotation Prof. Kiyoshi Hoshino, University of Tsukuba, Japan	
Lunch (12:05-14:00)	
March 22, 2020	
Session 1 - Meeting ID: 401-326-4155 (14:00 – 15:45)	Session 2 - Meeting ID: 606-418-2387 (14:00 – 15:45)
Take a Break (15:45-16:00)	
March 22, 2020	
Session 3 - Meeting ID: 401-326-4155 (16:00 – 17:45)	Session 4 - Meeting ID: 606-418-2387 (16:00 – 17:45)

