# 2019 Annual Report (June 2018 – May 2019) Visual Signal Processing and Communications Technical Committee IEEE Circuits and Systems Society

Chairman: Qi Tian Secretary: Wen-Hsiao Peng

# **Executive Summary**

Overall, the TC is in a very good shape. Major activities include:

- <u>Conference organization</u>: We have successfully supported ISCAS 2019 where the VSPC track attracted 43 submissions and accepted 23 papers. We have successfully organized VCIP'2018 in Taiwan, which attracted 260+ submissions. VCIP'2019 will take place in Sydney, Australia. We have also identified the team for VCIP'2020, which will be held in Hong Kong.
- Membership: We have retired 16 members in 2018 and elected 13 new members in 2019, reaching a total of 50 active members.
- 3) <u>TC website</u>: We have updated our TC website to include TC bylaw, member information as well as TC NewsLetter.
- 4) <u>Members' achievements</u>: Our members played several leading roles in IEEE conferences/journals and others, with 65+ leading roles in conferences (9 as conference chairs, 6 as TPC Chairs, 50 as Track/Area Chairs, RCM, Special Session Chairs, etc.) and 52 leading positions for journals (2 EiCs and 50 AEs/GEs/Editorial Board Members). Other notable recognitions include 2 newly elected IEEE Fellows in 2018, 1 Distinguished Lecturer, 13 distinguished IEEE leader positions, and 13 invited talks including keynote speeches.

# 5) **<u>Potential collaborations with other TCs</u>**:

- Talks: Our members have strong expertise on multimedia AI applications, visual analytics and visual systems. We can offer cutting-edge talks / tutorials in conferences / workshops.
- Special issue: Our TC members Mathias Wien and Wen-Hsiao Peng successfully organized an IEEE JETCAS special issue on Immersive Video Coding and Transmission, which was published in March 2019.
- Event organization: We welcome other TCs to have joint tracks or special sessions in our TC's main conferences: ISCAS and VCIP.

# 1. Technical Committee Meeting

The Visual Signal Processing and Communications Technical Committee in the IEEE Circuits and Systems Society organizes two annual TC meetings, one in ISCAS and the other in VCIP. The detailed information is given as follows.

## 1.1. Upcoming TC Meeting in ISCAS 2019

Date: May 29th, 2019 (Wednesday, during ISCAS 2019) Venue: Meeting Room 3 (Sapporo Convention Center) ISCAS 2019, Sapporo, Japan Time: 11:30 am -12:45 pm

## 1.2. Proposed Agenda

(To be finalized)

- Approval of the agenda
- BOG & TC Chair meeting reports & Review TC activities
- Conference Subcommittee
  - Update on IEEE CAS financial support
  - VCIP 2019 Status Report
  - VCIP 2020 Status Report
- Publicity Subcommittee
- Award Subcommittee
- Membership Subcommittee
  - Nomination of new TC members
- Other business
- Adjourn

# 2. Members Who Submitted Annual Reports

First Name	Last Name	Affiliation	Email
Qi	Tian	University of Texas at San Antonio	qi.tian@utsa.edu
Wen-Hsiao	Peng	National Chiao Tung University	wpeng@cs.nctu.edu.tw
Hsueh-Ming	Hang	National Chiao Tung University	hmhang@nctu.edu.tw
Ying	Li	IBM	yingli@us.ibm.com
Kai-Kuang	Ma	Nanyang Technological University	EKKMA@ntu.edu.sg
Nam	Ling	Santa Clara University	nling@scu.edu
Hsu-Feng	Hsiao	National Chiao Tung University	hillhsiao@cs.nctu.edu.tw
Chris Gwo Giun	Lee	National Cheng Kung University	clee@mail.ncku.edu.tw
Carl James	Debono	University of Malta	c.debono@ieee.org
Daniel	Lun	Hong Kong Polytechnic University	enpklun@polyu.edu.hk
Ebroul	Izquierdo	Queen Mary, University of London	Ebroul.izquierdo@qmul.ac.uk
Eduardo	Da Silva	Federal University of Rio de Janeiro	eduardo@smt.ufrj.br
Fei	Qiao	Tsinghua University	qiaofei@tsinghua.edu.cn
Ichiro	Ide	Nagoya University	ide@i.nagoya-u.ac.jp
Jian	Zhang	University of Technology Sydney	jian.zhang@uts.edu.au
Jiwen	Lu	Tsinghua University	lujiwen@tsinghua.edu.cn
Lei	Zhang	Microsoft	leizhang@microsoft.com
Lu	Yu	Zhejiang University	yul@zju.edu.cn
Mathias	Wien	RWTH Aachen University	wien@lfb.rwth-aachen.de
Тао	Mei	AI Research, JD.COM	tmei@live.com
Enrico	Magli	Politecnico di Torino	Enrico.magli@polito.it
Zhibo	Chen	University of Science and Technology of China	chenzhibo@ustc.edu.cn

Shiliang	Zhang	Peking University	slzhang.jdl@pku.edu.cn
Weiyao	Lin	Shanghai Jiao Tong University	wylin@sjtu.edu.cn
Guo-Jun	Qi	Huawei Cloud	guojunq@gmail.com
Jianfei	Cai	Nanyang Technological	asjfcai@ntu.edu.sg
		University	

# 3. Accomplished Technical Activities (June 2018 to May 2019)

# 3.1. Conference/Event organizations

Your Name	Conference	Conference/Event Title	Your Role	
0' T'	Sponsors			
Qi Tian	ACM	ACM Multimedia 2018	Area Chair	
Qi Tian	IEEE	Data 2019	Publicity Chair	
Qi Tian	IEEE	ICME 2018	Panel Chair and Area Chair	
Qi Tian	ACM	ICMR	TPC Chair	
Wen-Hsiao Peng	IEEE	ICIP 2019	Publication Chair	
Wen-Hsiao Peng	IEEE	ISCAS 2019	Track Chair	
Wen-Hsiao Peng	IEEE	ICME 2018	Area Chair	
Wen-Hsiao Peng	APSIPA	ASC 2018	Technical Program Co-chair	
Wen-Hsiao Peng	IEEE	VCIP 2018, ICIP 2019, ICME 2019, ICASSP 2019	Reviewer	
Hsueh-Ming Hang	IEEE	ICIP 2019	General Co-chair	
Kai-Kuang Ma	APSIPA	APSIPA 2019	APSIPA BoG Member	
Kai-Kuang Ma	IEEE	ISPACS 2019	Award Chair	
Kai-Kuang Ma	IEEE	ICIP 2019, ICASSP 2019, ISCAS 2019	Reviewer	
Nam Ling	IEEE	Umedia 2018	General Co-Chair	
Nam Ling	IEEE	SiPS 2018	Special Session Co-Organizer	
Hsu-Feng Hsiao	IEEE	DSP 2018	Special Session Organizer and Session Chair	
Hsu-Feng Hsiao	IEEE	ISCAS 2018	Review Committee Member and Session Chair	
Hsu-Feng Hsiao	IEEE	Globecom 2018	Technical Program Committee Member	
Hsu-Feng Hsiao	IEEE	ICUFN 2018	Technical Program Committee Member	
Hsu-Feng Hsiao	IEEE	IEEE 88th Vehicular Technology Conference, 2018	Technical Program Committee Member	
Hsu-Feng Hsiao	IEEE	WF-5G'18 (IEEE 1st 5G WORLD FORUM 2018)	Technical Program Committee Member	
Hsu-Feng Hsiao	IEEE	IEEE Globecom 2019	Technical Program Committee Member	
Hsu-Feng Hsiao	IEEE	ICUFN 2019	Technical Program Committee Member	
Hsu-Feng Hsiao	IEEE	ISCAS 2019	Review Committee Member and Session Chair	
Chris Gwo Giun Lee	IEEE	AICAS 2019/Special Session	Special Session Chair	
Chris Gwo Giun Lee	IEEE	AICAS 2019/Industry Forum	Organizer and Coordinator	
Chris Gwo Giun Lee	IEEE	ISCAS 2019/Industry Forum	Coordinator	
Chris Gwo Giun Lee	APSIPA	ASC 2019/Special Session	Special Session Co-chair	
Carl James Debono	IEEE	ICME 2019	Technical Program Committee Member	
Carl James Debono	IEEE	ICIP 2019	Technical Program Committee Member	
Carl James Debono	IEEE	ISCAS 2019	Review Committee Member	
Carl James Debono	IEEE	VCIP 2018	Technical Program Committee Member	
Ebroul Izquierdo	IEEE	VCIP 2019	General Co-chair	
Ebroul Izquierdo	IEEE	ICIP 2019	Area Chair	
Ebroul Izquierdo	IEEE	VCIP 2018	Technical Program Committee Member	
Fei Qiao	IEEE	SiPS 2018	Technical Program Committee Member	
Fei Qiao	IEEE	VCIP 2018	Technical Program Committee Member	
Fei Oiao	IEEE	ATS 2018	Special Session Chair	
Fei Oiao	IEEE	SiPS 2019	Special Session Chair	
Ichiro Ide	ACM	8th Int. Conf. on Multimedia Retrieval (ICMR2018)	Organization chair	
Ichiro Ide	IJCAI/IEICE	10th Workshop on Cooking and	General chair	

		Eating Activities (CEA2018)		
Ichiro Ide	ACM	26th Int. Conf. on Multimedia	Web & social media chair	
		(ACMMM2018)		
Ichiro Ide	IEEE	VCIP 2018	Demo chair	
Ichiro Ide	-	9th Int. Symposium on Information	Program chair	
		and Communication Technology		
T.1.'	IFFF	(SOICT 2018)		
Ichiro Ide	IEEE	ISM 2018	Workshop chair	
Ichiro Ide	IEEE	2nd Workshop on Data Engineering	Chair	
		Recipe (DECOR) 2019		
Ichiro Ide	-	16th IAPR Int Conf on Machine	Local arrangement chair	
1011110 140		Vision Applications (MVA2019)		
Jiwen Lu	IEEE	ICIP 2018	Area Chair	
Jiwen Lu	IEEE	ICME 2018	Area Chair	
Jiwen Lu	IEEE	ICPR 2018	Area Chair	
Lei Zhang	IEEE	ICME 2019	TPC Co-chair	
Lei Zhang	IEEE	ICIP 2019	Area Chair	
Lu Yu		PCS 2019	General Chair	
Mathias Wien	IEEE	ICIP 2019	Area Chair	
Mathias Wien	IEEE	ICME 2019	Area Chair	
Mathias Wien	IEEE	ICME 2019	Session Chair	
Mathias Wien	IEEE	ISCAS 2019	RCM	
Mathias Wien	IEEE	PCS 2019	TPC Co-chair	
Mathias Wien	IEEE	ICIP 2019	Reviewer	
Mathias Wien	IEEE	ICME 2019	Reviewer	
Mathias Wien	IEEE	VCIP 2018	Reviewer	
Mathias Wien	IEEE	PCS 2018	Reviewer	
Tao Mei	IEEE	ICME 2019	General Co-chair	
Tao Mei	ACM	Multimedia 2019	Grand Challenge Co-chair & Area Chair	
Tao Mei	ACM	SIGIR 2019	Senior PC	
Enrico Magli	IEEE	ICASSP	Track chair	
Enrico Magli	IEEE	ICME 2019	Special session co-chair	
Enrico Magli		PCS 2018	TPC member	
Enrico Magli	Eurasip	Eusipco 2018	TPC member	
Enrico Magli	IEEE	ICIP 2018	Area chair	
Enrico Magli	IEEE	VCIP 2018	RCM	
Zhibo Chen	IEEE	PCS 2019	TPC Chair	
Zhibo Chen	IEEE	ACM MM 2019	PC Member	
Zhibo Chen	IEEE	ISCAS 2019	RC Member/Session Chair	
Weiyao Lin	IEEE	ICME 2018	Area Chair	
Weiyao Lin	ACM	Multimedia 2018	Area Chair	
Weiyao Lin	IEEE	ICIP 2019	Area Chair	
Weiyao Lin	IEEE	ICME 2019	Student Program Chair	
Guo-Jun Qi	IEEE	ICIP	Area Chair	
Guo-Jun Qi	ACM	ACM Multimedia	Area Chair	
Guo-Jun Qi	Springer	International Conference on	Steering Committee Member	
Cure Ive O	LADD		Arres Chain	
Guo-Jun Qi	IAPK		Area Unair	
Guo-Jun Qi	ACM		Program Committee Member	
Guo-Jun Qi	ACM	AUM Multimedia	workshop Unair	
Guo-Jun Qi	IJCAI	Artificial Intelligence	Semor Program Committee Member	
Jianfei Cai	IEEE	IEEE MMSP 2019	Keynote and Panel Chair	
Jianfei Cai	CVF	ICCV 2019	Area Chair	
Jianfei Cai	ACM	ACM Multimedia 2018	Area Chair	

# 3.2. IEEE and Other Journal Editorships:

Your Name	Journal Sponsors	Journal Title	Your Role
Qi Tian	IEEE	Transactions on Circuits and Systems for Video Technology	Associate Editor
Qi Tian	IEEE	Transactions on Neural Networks and Learning Systems	Associate Editor
Qi Tian	IEEE	Transactions on Multimedia	Associate Editor
Qi Tian	ACM	ТОММ	Associate Editor
Qi Tian	Springer	Multimedia Systems	Associate Editor
Qi Tian	Springer	Journal of Machine Vision and Applications	Associate Editor
Wen-Hsiao Peng	IEEE	Transactions on Circuits and Systems for Video Technology	Associate Editor

Wen-Hsiao Peng	IEEE	Journal on Emerging and Selected Topics in Circuits and Systems	Guest Editor
Wen-Hsiao Peng	IEEE	Journal on Emerging and Selected Topics in Circuits and Systems	Senior Editorial Board
		••••••••••••••••••••••••••••••••••••••	Member
Wen-Hsiao Peng	IEEE	Transactions on Circuits and Systems II: Express Briefs	Guest Editor
Ying Li	Springer	International Journal of Multimedia Information Retrieval	Editorial on board
Ving Li	Elsevier	Journal of Visual Communication and Image Penresentation	Editorial on board
Voi Vuong Mo	IEEE	Transactions on Image Processing	Sonior Associate Editor
	IEEE		
Kai-Kuang Ma	IEEE	Transactions on Circuits & Systems for Video Technology	Associate Editor
Kai-Kuang Ma	IEEE	Signal Processing Letters	Senior Associate Editor
Nam Ling	IEEE	Access (Special Section on Recent Advances in Video Coding and Security)	Guest Editor
Chris Gwo Giun	IEEE	Transactions on Signal Processing	Associate Editor
Lee Chris Cure Cium	SDDINCED	Journal of Signal Dragogging Systems	Associate Editor
Lee	SPRINGER	Journal of Signal Processing Systems	Associate Eultoi
Carl James Debono	IEEE COMSOC	R-Letters	Review Board Member
Ebroul Izquierdo	Eurasip	Image and video processing	Associate Editor
Eduardo da Silva	IEEE	Transactions on Circuits and Systems I: Regular Papers	Guest Editor, Special Session on ISCAS 2018
Eduardo da Silva	Elsevier	Journal of the Franklin Institute	Associate Editor
Fei Qiao	SPRINGER	MEJ, Special Issue for Approximate Computing	Associate Editor
Ichiro Ide	Springer	Multimedia Systems	Associate Editor
Ichiro Ide	Slovensko društvo	Informatica, Spcial Issue on "The Eighth Int. Symposium on	Guest Editor
	Informatika	Information and Communication Technology-SoICT 2017"	
Ichiro Ide	IEICE	Transactions on Information and Systems, Special Section on "Multimedia for Cooking and Eating Activities"	Guest Editor-in-Chief
Jiwen Lu	Elsevier	Pattern Recognition Letters	Co-Editor-in-Chief
Jiwen Lu	IEEE	Transactions on Image Processing	Associate Editor
Jiwen Lu	IEEE	Transactions on Circuits and Systems for Video Technology	Associate Editor
Jiwen Lu	IEEE	Transactions on Biometrics, Behavior, and Identity Sciences	Associate Editor
Jiwen Lu	Elsevier	Pattern Recognition	Associate Editor
Jiwen Lu	Elsevier	Journal of Visual Communication and Image Representation	Associate Editor
Lei Zhang	IEEE	Transactions on Circuits & Systems for Video Technology	Associate Editor
Lei Zhang	IEEE	Transactions on Multimedia	Associate Editor
Lei Zhang	Springer	Multimedia Systems Journal	Associate Editor
Mathias Wien	IEEE	Transactions on Circuits & Systems for Video Technology	Associate Editor
Mathias Wien	IEEE	Journal on Emerging and Selected Topics in Circuits and Systems	Leading Guest Editor
Tao Mei	IEEE	Transactions on Image Processing	Associate Editor
Tao Mei	IEEE	Transactions on Circuits and Systems for Video Technology	Associate Editor
Tao Mei	ACM	Transactions on Multimedia Computing Communications and	Associate Editor
		Applications (TOMM)	
Tao Mei	ACM	Transactions on Intelligent Systems and Technology	Associate Editor
Weiyao Lin	Elsevier	Journal of Visual Communication and Image Representation	Associate Editor
Weiyao Lin	Elsevier	Signal Processing: Image Communication	Area Editor
Weiyao Lin	Springer	Circuits, Systems, and Signal Processing	Associate Editor
Weiyao Lin	IEEE	Transactions on Circuits and Systems for Video Technology	Associate Editor
Weiyao Lin	IEEE	Transactions on Intelligent Transportation Systems	Associate Editor
Guo-Jun Qi	IEEE	Transactions on Circuits and Systems for Video Technology	Associate Editor
Guo-Jun Qi	IEEE	Transactions on Image Processing	Associate Editor
Guo-Jun Qi	Elsevier	Pattern Recognition	Associate Editor
Guo-Jun Qi	ACM	Transactions on Knowledge Discovery from Data	Associate Editor
Jianfei Cai	IEEE	IEEE Transactions on Multimedia	Associate Editor

**3.3. Awards, Honors, and Recognition** (Fellow, Distinguished Lecturer, Outstanding Service, Best Paper Awards, etc.)

Your Name	Awards / Honors / Recognition	Period
Qi Tian	ACM CIKM 2019 Best Overall Paper Award	2019
Wen-Hsiao Peng	APSIPA ASC 2018 Best Paper Award Runner-up	Nov. 12-15, 2018
Nam Ling	Minjiang Scholar	2018 -
Chris Gwo Giun	IEEE Distinguished Lecturer (CAS)	2019-2020
Lee		
Chris Gwo Giun	Outstanding Technology Transfer Award Ministry of Science and	2018
Lee	Technology, Taiwan	
Lei Zhang	ACM ICMR 2018 Best Poster Award	2018
Tao Mei	IEEE Fellow	Class of 2019
Weiyao Lin	Outstanding Area Chair award of ICME 2018	2018

# 3.4. Keynote Speeches/Invited Talks:

Your Name	Invited by	<b>Event/Conference</b> Title	Talk Title	Date
Wen-Hsiao Peng	Sungkyunkwan Univ., South Korea.	APSIPA distinguished lecture	Reinforcement Learning for Video Encoder Control and Video Prediction	Mar. 14, 2019
Wen-Hsiao Peng	Tongji Univ., China.	APSIPA distinguished lecture	Reinforcement Learning for HEVC/H.265 Video Rate Control	Nov. 20, 2018
Nam Ling	Zhongyuan University of Technology	Distinguished lecture, Hong De Lecture Hall	Intra Prediction in Versatile Video Coding (VVC)	Aug. 20, 2018
Nam Ling	Xi'an U of Posts & Telecommunications	Invited talk	Intra Prediction in Versatile Video Coding (VVC)	Aug. 29, 2018
Nam Ling	Tianjin University	Invited talk	Intra Prediction in Versatile Video Coding (VVC)	Aug. 21, 2018
Nam Ling	IEEE CAS Society Singapore Chapter	Invited talk	Intra Prediction in Versatile Video Coding (VVC)	July 3, 2018
Nam Ling	Fuzhou University	Invited talk	Intra Prediction in Versatile Video Coding (VVC)	June 28, 2018
Chris Gwo Giun Lee	IEEE Benelux Section	Distinguished lecture	Algorithm/Architecture Co-design for Intelligent Hardware Systems in Parallel/Reconfigurable Edge	May 22, 2019
Chris Gwo Giun Lee	IEEE Portugal Section	Distinguished lecture	Algorithm/Architecture Co-design for Smart Signals and Systems in Cognitive Cloud/Edge	May 20, 2019
Chris Gwo Giun Lee	IEEE SiPS	Tutorial Speaker	Algorithm/Architecture Co-design for Smart Signals and Systems in Cognitive Cloud/Edge	Oct. 23, 2018
Ebroul Izquierdo	Organizers	International Conference on Information and Communications Technology	Visual Information Retrieval	July 24, 2019
Lei Zhang	IEEE BigData 2018	IEEE BigData 2018 Workshop on Big Data Transfer Learning	Object Detection at Scale as Cloud Services	Dec. 10, 2018
Enrico Magli	OBPDC 2018	2018 Onboard Payload Data Compression workshop	Deep learning: a great fit for onboard data processing?	Sep. 21, 2018

**3.5. Other distinguished IEEE services** (e.g., CAS BoG, Region presidents, VP, TC Chairs/Secretary,... )

Your Name	Organization	Position/Activities	Period
Qi Tian	IEEE CASS VSPC TC	Chair	2018-2020
Wen-Hsiao Peng	IEEE CASS VSPC TC	Secretary	2018-2020
Chris Gwo Giun	IEEE R10	Chair, IEEE Region 10 ExCom, Industry Relation	2019-present
Lee		Committee	
Chris Gwo Giun	IEEE LSTC	BoG Member	2018-present
Lee			
Chris Gwo Giun	IEEE Tainan Section	Vice Chair, Technical Activities	2018-present
Lee			
Chris Gwo Giun	IEEE Member & Geographic	Member	2018-2019
Lee	Activities, Member Benefit		
	Portfolio Committee		
Chris Gwo Giun	IEEE Region 10 ExCom	Chair, IEEE Region 10 ExCom, Individual Benefits	2017-2018
Lee		& Services Committee	
Eduardo da Silva	SBrT - Brazilian	Vice President Technical Activities	2014-2018
	Telecommunications Society		
Eduardo da Silva	IEEE	Regional Director – R7 and R9 – Signal Processing	2018-2019
		Society	
Fei Qiao IEEE CASS VSPC TC		Publicity Sub-committee Co-Chair	2017-2018
Fei Qiao IEEE SPS DISPS TC		Election Sub-committee Co-Chair	2018-2019
Jian Zhang IEEE CASS VSPC TC		Conference Sub-committee Co-Chair	2017-2019
Lu Yu	IEEE CASEO TC	Secretary	2017-2019
Guo-Jun Qi	IEEE VSPC TC	Publicity Chair	2019-present

### **5.** TC Significant Activities List

[Please list your 2 (or less) most significant activities in the past year (June 2018--May 2019), including paper, special session, special issue, workshop, conference, award, important position, etc]

### • Qi Tian

• [award] **Best Overall Paper Award**, the 27th ACM International Conference on Information and Knowledge Management (CIKM), October 2018.

J. Zhang, Y. Liu, S. Ma, and Q. Tian, "Relevance Estimation with Multiple Information Sources on Search Engine Result Pages," The 27th ACM International Conference on Information and Knowledge Management (CIKM), Best Overall Paper, Torino, Italy, October 22-26, 2018.

Relevance estimation is among the most important tasks in the ranking of search results because most search engines follow the Probability Ranking Principle. In this paper, we propose a novel framework named Joint Relevance Estimation model (JRE), which learns the visual patterns from screenshots of search results, explores the presentation structures from HTML source codes and also adopts the semantic information of textual contents.

#### • Lu Yu

[Important Position] Video Group Chair of ISO/IEC JTC1 SC29 WG11 (Moving Picture Experts Group, MPEG) since Jan. 2018

50-word summary: Organized standardization activities on enhanced 3-degree-of-freedom (3DoF+) and 6-degree-of-freedom (6DoF) video coding, Essential Video Coding (EVC) - a licensing friendly new generation of video coding, and compression of neural networks, etc.

#### • Mathias Wien

[Special Issue] Journal on Emerging and Selected Topics in Circuits and Systems, Special issue on Immersive Video Coding and Transmission, Guest Editors M. Wien, J. M. Boyce, T. Stockhammer and W. H. Peng.

[Paper] M. Wien, J. M. Boyce, T. Stockhammer and W. H. Peng, "Standardization Status of Immersive Video Coding," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 9, no. 1, pp. 5-17, March 2019.

doi: 10.1109/JETCAS.2019.2898948.

Based on increasing availability of capture and display devices dedicated to immersive media, coding, and transmission of these media has recently become a highest-priority subject of standardization. Different levels of immersiveness are defined with respect to an increasing degree of freedom in terms of movements of the observer within the immersive media scene. The level ranges from three degrees of freedom allowing the user to look around in all directions from a fixed point of view to six degrees of freedom, where the user can freely alter the viewpoint within the immersive media scene. The moving pictures experts group (MPEG) of ISO/IEC is developing a standards suite on "Coded Representation of Immersive Media," called MPEG-I, to provide technical solutions for building blocks of the media transmission chain, ranging from architecture, systems tools, coding of video and audio signals, to point clouds and timed text. In this paper, an overview on recent and ongoing standardization efforts in this area is presented. While some specifications, such as high efficiency video coding or version 1 of the omnidirectional media format, are already available, other activities are under development or in the exploration phase. This paper addresses the status of these efforts with a focus on video signals, indicates the development timelines, summarizes the main technical details, and provides pointers to further points of reference.

#### Daniel Lun

[Best paper award top-ten finalist in ISCAS 2018] Tingtian Li and Daniel P. K. Lun, "A Novel Reflection Removal Algorithm Using the Light Field Camera," in *Proc. IEEE International Symposium on Circuits and Systems* (ISCAS), May 27-30, 2018.

In daily photography, it is common that the captured images are superimposed with undesired reflection of another scene. Such reflection does not only reduce the visual quality of the target background scene, but also affects the subsequent processing on the image. Removing the reflection without any prior information of the background is a very challenging task. In this paper, we propose a novel method to remove the reflection based on light field (LF) imaging. The proposed algorithm does not have the assumptions of the existing methods; thus, it is more robust. Experimental results show that the proposed algorithm outperforms the existing approaches both qualitatively and quantitatively.

[Book chapter] Daniel P.K. Lun and B. Budianto, "Non-contact Three-Dimensional Measurement Using the Learning Approach," in Learning Approaches in Signal Processing (Wan-Chi Siu, Lap-Pui Chau, Liang Wang and Tieniu Tan, Ed.), Pan Stanford Publishing, 2018.

In this book chapter, robust non-contact three-dimensional measurement algorithms based on the learning method are presented. In particular, the fringe projection profilometry method was introduced and the ambiguity problem due to the discontinuity in the fringe image was explained. It was shown that by embedding codes patterns into the fringe patterns, the ambiguity problem can be solved. Unlike the conventional methods, the algorithms introduced in this book chapter does not require additional hardware setup or additional fringe pattern projection.

#### • Chris Gwo Giun Lee

[Industry Forum] IEEE AICAS 2019 Industry Forum (for YP & WIE). Title: Influences of EDGE Device's Instant Decision: From Bio-Tech, FinTech to Sustainable Energy & Beyond Time: March 19, 2019 15:00 ~ 16:00 Location: 10F, Ballroom B, Ambassador Hotel, Hsinchu, Taiwan

50-word summary: In view of aging societies, significance of digital economy and global warming, this Industry Forum is organized together with Young Professionals (YP) and Women in Engineering (WIE) within IEEE Circuits and Systems Society with industry leaders, investors, and venture capitalists invited to share their visions on how current state-of-the-art in EDGE/mobile devices, capable of near real-time decisions, may influence our daily life from the perspectives of healthcare, finance, and energy. Upon addressing the pain spots of these industry sectors, internship and potential industry/academia collaboration opportunities upon current Technology 4.0 entrepreneur landscape will also be speculated

[Chair] IEEE Region 10 ExCom, Industry Relation Committee. Vision: Leadership in bridging industry and academia Mission: Cross pollinate between industry and academia in fostering innovations, internship and entrepreneurship

50-word summary: In view of the global fast-changing landscape due to Industry 4.0 and with anticipation of aligning Region 10 with IEEE's strategic direction in advancing SMARTECH and HARDTECH for humanity with higher industry involvement, the Industry Relations Committee (IRC) provides a professional networking platform upon which interactive bi-directional industry and academia activities are fostered in servicing IEEE members within Region 10!

#### • Eduardo da Silva

[Paper] Ribeiro, F. M. L., de Oliveira, J. F. L., Ciancio, A. G., da Silva, E. A. B., Estrada, C. R. D., Tavares, L. G. C., Gois, J. N., Said, A., Martelotte, M. C., "Quality of Experience in a Stereoscopic Multiview Environment", *IEEE Transactions on Multimedia*, 20(1), pp. 1-14, January 2018.

In this paper we investigate how visualization factors, such as disparity, mobility, angular resolution and viewpoint interpolation, influence the Quality of Experience (QoE) in a stereoscopic multiview environment. The effects of angular resolution and viewpoint interpolation on the quality of experience produced by multiview systems have been assessed, providing relevant cues as to how the baselines of cameras and interpolation strategies in such systems affect user experience. Aspects such as visual comfort, model fluidity, sense of immersion, and the 3D experience as a whole have been assessed for several test cases. Obtained results suggest that user experience in a motion parallax environment is not as critically influenced by configuration parameters such as disparity as initially thought. In addition, while users are very sensitive to angular resolution in multiview 3D systems, this sensitivity tends not to be as critical when a user is performing a task that involves a great amount of interaction with the multiview content. These tests have also indicated that interpolating intermediate viewpoints can be effective in reducing the required view density without degrading the perceived QoE.

[Special Issue on ISCAS 2017] Thomaz, L., Jardim, E., da Silva, A. F., da Silva, E. A. B., Netto, S. L., Krim, H., "Anomaly Detection in Moving-Camera Video Sequences Using Principal Subspace Analysis", *IEEE Transactions on Circuits and Systems-I: Regular Papers*, Special Issue on ISCAS 2017, 65:(3), pp. 1003--1015, May 2018.

This paper presents a family of algorithms based on sparse decompositions that detect anomalies in video sequences obtained from slow moving cameras. These algorithms start by computing the union of subspaces that best represents all the frames from a reference (anomaly-free) video as a low-rank projection plus a sparse residue. Then they perform a low-rank representation of a target (possibly anomalous) video by taking advantage of both the union of subspaces and the sparse residue computed from the reference video. Such algorithms provide good detection results while at the same time obviating the need for previous video synchronization.

#### Wen-Hsiao Peng

[Best Paper Award Runner-up] H. P. Wang, W. J. Ko, and W. H. Peng, "Learning Priors for Adversarial Autoencoders," in *Proc. Asia-Pacific Signal and Information Processing Association (APSIPA) Annual Summit* 

Conference (ASC), Nov. 2018.

Most deep latent factor models choose simple priors for simplicity, tractability or not knowing what prior to use. Recent studies show that the choice of the prior may have a profound effect on the expressiveness of the model, especially when its generative network has limited capacity. This work introduces the notion of code generators to transform manually selected simple priors into ones that can better characterize the data distribution for adversarial autoencoders (AAEs). Experimental results show that the proposed model can generate better image quality and learn better disentangled representations than AAEs in both supervised and unsupervised settings.

[Paper] W. L. Chang, H. P. Wang, W. H. Peng and W. C. Chiu, "All about structure: adapting structural information across domains for boosting semantic segmentation," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, June 2019.

This paper tackles the problem of unsupervised domain adaptation for semantic segmentation, with the aim of transferring knowledge learned upon synthetic datasets with ground-truth labels to real-world images without annotations. Assuming that the structural content of images is most informative to semantic segmentation, this work disentangles images into domain-invariant structure and domain-specific texture representations for image-translation and label transfer across domains. Extensive experiments confirm its superiority over several state-of-the-art approaches.

#### • Ichiro Ide

[Paper] M.A. Kastner, I. Ide, Y. Kawanishi, T. Hirayama, D. Deguchi, and H. Murase, "Estimating the visual variety of concepts by referring to Web popularity", Multimedia Tools and Applications, Published online in Aug. 2018.

50-word summary: In this paper, a method to measure the visual variety of concepts is proposed to quantify the semantic gap between vision and language. For this, an image corpus is recomposed using ImageNet and Web data. Web-based metrics for measuring the popularity of sub-concepts are used as a weighting to ensure that the image composition in a dataset is as natural as possible. Using clustering methods, a score describing the visual variety of each concept is determined. A crowd-sourced survey is conducted to create ground-truth values applicable for this research. The evaluations show that the recomposed image corpus largely improves the measured variety compared to previous datasets.

[Conference] 8th ACM Int. Conf. on Multimedia Retrieval (ICMR2018)

50-word summary: Organized the 8th ACM Int. Conf. on Multimedia Retrieval (ICMR2018), which is one of the top-level conferences in the multimedia field in Yokohama, Japan. I contributed to the organization of this conference as the Organization chair, which involved solutions to various logistics issues. The conference attracted more than 220 participants, approximately 50% increase than the past conferences of the series.

#### Carl James Debono

[Paper] D. Seychell, C.J. Debono, "Ranking Regions of Visual Saliency in RGB-D Content," in *Proc. International Conference on 3D Immersion (IC3D)*, Dec. 2018.

50-word summary: Effective immersion takes place when the user can relate to the 3D environment presented and interact with key objects. Efficiently predicting which objects in a scene are in the user's attention, without using additional hardware, such as eye tracking solutions, provides an opportunity for creating more immersive scenes in real time and at lower costs. This is nonetheless algorithmically challenging. In this paper, we are proposing a technique that efficiently and effectively identifies the most salient objects in a scene. We show how it accurately matches user selection within 0.04s and is over 95% faster than other saliency algorithms while also providing a ranking of the most salient segments in a scene.

[Paper] L. Attard, C.J. Debono, G. Valentino, M. Di Castro, M.L. Baiguera Tambutti, "An RGB-D Video-based Wire Detection Tool to Aid Robotic Arms during Machine Alignment Measurement," in *Proc. IEEE International Conference on Imaging Systems and Techniques* (IST 2018), pp. 248 - 253, Oct. 2018.

50-word summary: Industrial equipment may require precise alignment of components in order to function properly. In this work, we consider the machine alignment of a beamline in a research facility where an autonomous measurement system is used to execute these machine alignment procedures. The system consists of two robotic arms equipped with laser sensors placed at their ends which with the help of fiducial markers, are used by a camera system to calculate the position of the machine with respect to a stretched wire installed on top of the machines. In this work we propose a wire detection tool that assists the measurement process by detecting the line automatically through RGB-D images. Although the system was implemented for the specific application of beam line component alignment in the LHC tunnel, the same principle can be used in various other applications requiring detection of fine lines at a close distance.

Jian Zhang

[Paper] Huang, X, Zhang, J, Fan, L, Wu, Q & Yuan, C 2017, "A Systematic Approach for Cross-Source Point Cloud Registration by Preserving Macro and Micro Structures," *IEEE Transactions on Image Processing*, vol. 26, no. 7, pp. 3261-3276.

We propose a systematic approach for registering cross-source point clouds that come from different kinds of sensors. This task is especially challenging due to the presence of significant missing data, large variations in point density, scale difference, large proportion of noise, and outliers. The robustness of the method is attributed to the extraction of macro and microstructures. Compared with eight state-of-the-art registration algorithms, the proposed method invariably outperforms on Pisa Cathedral and other challenging cases

[Paper] Yao, Y, Zhang, J, Shen, F, Hua, X, Xu, J & Tang, Z 2017, "Exploiting Web Images for Dataset Construction: A Domain Robust Approach", *IEEE Transactions on Multimedia*, vol. 19, no. 8, pp. 1771-1784.

To reduce the cost of manual labeling, there has been increased research interest in automatically constructing image datasets by exploiting web images. Datasets constructed by existing methods tend to have a weak domain adaptation ability, which is known as the "dataset bias problem." To address this issue, we present a novel image dataset construction framework that can be generalized well to unseen target domains. we build an image dataset with 20 categories. Extensive experiments on image classification, cross-dataset generalization, diversity comparison, and object detection demonstrate the domain robustness of our dataset.

#### • Lei Zhang

[Paper] X. Wang, Q. Huang, A. Celikyilmaz, J. Gao, D. Shen, Y.-F. Wang, W.-Y. Wang, L. Zhang, "Reinforced Cross-Modal Matching and Self-Supervised Imitation Learning for Vision-Language Navigation," in *Proc. IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019.

Vision-language navigation (VLN) is the task of navigating an embodied agent to carry out natural language instructions inside real 3D environments. In this paper, we study how to address three critical challenges for this task: the cross-modal grounding, the ill-posed feedback, and the generalization problems. We propose a novel Reinforced Cross-Modal Matching (RCM) approach that enforces cross-modal grounding both locally and globally via reinforcement learning (RL) and a Self-Supervised Imitation Learning (SIL) method to explore unseen environments by imitating its own past, good decisions. Evaluation on a VLN benchmark dataset shows that our model significantly outperforms existing methods by 10% on SPL and achieves the new state-of-the-art performance.

[Paper] Q. Huang, P. Zhang, D. Wu, L. Zhang, "Turbo Learning for Captionbot and Drawingbot," in *Proc. Neural Information Processing Systems (NeurIPS)*, 2018.

We study in this paper the problems of both image captioning and text-to-image generation, and present a novel turbo learning approach to jointly training an image-to-text generator (aka CaptionBot) and a text-to-image generator (aka DrawingBot). The key idea behind the joint training is that CaptionBot and DrawingBot as dual problems can form a closed loop to provide informative feedback to each other. This enables semi-supervised learning since the closed loop can provide peudo-labels for unlabeled samples. Experimental results on the COCO dataset demonstrate that the proposed turbo learning can significantly improve the performance of both CaptionBot and DrawingBot by a large margin.

#### • Zhibo Chen

[Paper] Zhibo Chen\*, Tianyu He, Xin Jin, Feng Wu, "Learning for Video Compression", *IEEE Transactions on Circuits and Systems for Video Technology*, 2019.

We propose the concept of PixelMotionCNN (PMCNN) which includes motion extension and hybrid prediction networks. PMCNN can model spatiotemporal coherence to effectively perform predictive coding inside the learning network. On the basis of PMCNN, we further explore a learning-based framework for video compression with additional components of iterative analysis/synthesis, binarization, etc. Experimental results demonstrate the effectiveness of the proposed scheme. Although entropy coding and complex configurations are not employed in this paper, we still demonstrate superior performance compared with MPEG-2 and achieve comparable results with H.264 codec. The proposed learning-based scheme provides a possible new direction to further improve compression efficiency and functionalities of future video coding.

[Paper] Zhibo Chen\*, Wei Zhou, and Weiping Li, "Blind stereoscopic video quality assessment: From depth perception to overall experience," *IEEE Transactions on Image Processing*, 2018.

We propose a new depth perception quality metric (DPQM) and verify that it outperforms existing metrics on our published 3D video extension of High Efficiency Video Coding (3D-HEVC) video database. Furthermore, we validate its effectiveness by applying the crucial part of the DPQM to a novel blind stereoscopic video quality evaluator (BSVQE) for overall 3D video quality assessment. In the DPQM, we introduce the feature of auto-regressive prediction-based disparity entropy (ARDE) measurement and the feature of energy weighted video content measurement, which are inspired by the free-energy principle and the binocular vision mechanism. In the

BSVQE, the binocular summation and difference operations are integrated together with the fusion natural scene statistic measurement and the ARDE measurement to reveal the key influence from texture and disparity. Experimental results on three stereoscopic video databases demonstrate that our method outperforms state-of-theart SVQA algorithms for both symmetrically and asymmetrically distorted stereoscopic video pairs of various distortion types.

#### Shiliang Zhang

[Paper] J. Ye, S. Zhang, T. Huang, and Y. Rui, "CDbin: Compact Discriminative Binary Descriptor Learned with Efficient Neural Network," *IEEE Transactions on Circuits and Systems for Video Technology*, 2019.

As an important computer vision task, image matching requires efficient and discriminative local descriptors. Most of existing descriptors like SIFT and ORB are hand-crafted. It is necessary to study more optimized descriptors through endto-end learning. This paper proposes compact binary descriptors learned with a lightweight Convolutional Neural Network (CNN). Specifically, we propose a CNN with no larger than five layers for descriptor learning. The resulting descriptors, i.e., Compact Discriminative binary descriptors (CDbin) are optimized with four complementary loss functions. Extensive experiments on two image patch datasets and three image retrieval datasets show that CDbin exhibits competitive performance compared with existing descriptors. For example, 64-bit CDbin substantially outperforms 256-bit ORB and 1024-bit SIFT on Hpatches dataset.

[Paper] H. Yao, S. Zhang, R. Hong, Y. Zhang, C. Xu, and Q. Tian, "Deep Representation Learning with Part Loss for Person Re-Identification," *IEEE Transactions on Image Processing*, 2019.

Learning discriminative representations for unseen person images is critical for person Re-Identification (ReID). Most of current approaches learn deep representations in classification tasks, which essentially minimize the empirical classification risk on the training set. Inspired by the structural risk minimization principle in SVM, we revise the traditional deep representation learning procedure to minimize both the empirical classification risk and the representation learning risk. The representation learning risk is evaluated by the proposed part loss, which automatically generates several parts for an image, and computes the person classification loss on each part separately. Simultaneously considering multiple part loss enforces the deep network to learn discriminative representations. Experimental results on three datasets, i.e., Market1501, CUHK03, VIPeR, show that our representation outperforms the existing deep representations.

[Paper] L. Wei, S. Zhang, H. Yao, W. Gao, and Q. Tian, "GLAD: Global-Local-Alignment Descriptor for Scalable Person Re-Identification," *IEEE Transactions on Multimedia*, April 2019.

The huge variance of human pose and the misalignment of detected human images significantly increase the difficulty of pedestrian image matching in person Re-Identification (Re-ID). Moreover, the massive visual data being produced by surveillance video cameras requires highly efficient person Re-ID systems. Targeting to solve the first problem, this work proposes a robust and discriminative pedestrian image descriptor, namely, the Global-Local-Alignment Descriptor (GLAD). For the second problem, this work treats person Re-ID as image retrieval and proposes an efficient indexing and retrieval framework. Extensive experimental results on widely used public benchmark datasets show GLAD achieves competitive accuracy compared to the state-of-the-art methods. On a large-scale person, with the Re-ID dataset containing more than 520 K images, our retrieval framework significantly accelerates the online Re-ID procedure while also improving Re-ID accuracy.

#### Weiyao Lin

.

[Paper] W. Lin , Y. Mi, J. Wu, K. Lu, H. Xiong, "Action recognition with coarse-to-fine deep feature integration and asynchronous fusion," in *Proc. AAAI Conf. Artificial Intelligence (AAAI)*, 2018.

50-word summary: Propose a novel deep-based framework for action recognition, which improves the recognition accuracy by: 1) deriving more precise features for representing actions, and 2) reducing the asynchrony between different information streams. We first introduce a coarse-to-fine network which extracts shared deep features at different action class granularities and progressively integrates them to obtain a more accurate feature representation for input actions. We further introduce an asynchronous fusion network. It fuses information from different streams by asynchronously integrating stream-wise features at different time points, hence better leveraging the complementary information in different streams. Experimental results on action recognition benchmarks demonstrate that our approach achieves the state-of-the-art performance.

[Paper] K. Chen, J. Li, W. Lin, J. See, J. Wang, L. Duan, Z. Chen, C. He, J. Zou, "Towards accurate one-stage object detection with AP-loss," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, 2019.

50-word summary: Alleviates the foreground-background class imbalance by proposing a novel framework to replace the classification task in one-stage detectors with a ranking task, and invent the Average-Precision loss (AP-loss) for the ranking problem. The AP-loss cannot be optimized directly due to its non-differentiability and non-convexity. We develop a novel optimization algorithm for this purpose, which seamlessly combines errordriven update scheme in perceptron learning and backpropagation algorithm in deep networks together. We verify good convergence property of the proposed algorithm theoretically and empirically. Experimental results demonstrate notable performance improvement of the AP-loss over different kinds of classification-loss on the state-of-the-art one-stage detectors at various benchmarks, without changing the network architectures.

#### • Guo-Jun Qi

[Paper] Liheng Zhang, Guo-Jun Qi\*, Liqiang Wang, Jiebo Luo, "AET vs. AED: Unsupervised Representation Learning by Auto-Encoding Transformations rather than Data," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, 2019.

The success of deep neural networks often relies on a large amount of labeled examples, which can be difficult to obtain in many real scenarios. To address this challenge, unsupervised methods are strongly preferred for training neural networks without using any labeled data. In this paper, we present a novel paradigm of unsupervised representation learning by Auto-Encoding Transformation (AET) in contrast to the conventional Auto-Encoding Data (AED) approach. Given a randomly sampled transformation, AET seeks to predict it merely from the encoded features as accurately as possible at the output end. The idea is the following: as long as the unsupervised features successfully encode the essential information about the visual structures of original and transformed images, the transformations, from parameterized, to non-parameterized and GAN-induced ones. Our experiments show that AET greatly improves over existing unsupervised approaches, setting new state-of-the-art performances being greatly closer to the upper bounds by their fully supervised counterparts on CIFAR-10, ImageNet and Places datasets.

[Paper] Liheng Zhang, Marzieh Edraki, Guo-Jun Qi\*, "CapProNet: Deep Feature Learning via Orthogonal Projections onto Capsule Subspaces," in *Proc. Neural Information Processing Systems (NeurIPS)*, Dec. 2018.

In this paper, we formalize the idea behind capsule nets of using a capsule vector rather than a neuron activation to predict the label of samples. To this end, we propose to learn a group of capsule subspaces onto which an input feature vector is projected. Then the lengths of resultant capsules are used to score the probability of belonging to different classes. We train such a Capsule Projection Network (CapProNet) by learning an orthogonal projection matrix for each capsule subspace, and show that each capsule subspace is updated until it contains input feature vectors corresponding to the associated class. We will also show that the capsule projection can be viewed as normalizing the multiple columns of the weight matrix simultaneously to form an orthogonal basis, which makes it more effective in incorporating novel components of input features to update capsule representations. In other words, the capsule projection can be viewed as a multi-dimensional weight normalization in capsule subspaces, where the conventional weight normalization is simply a special case of the capsule projection onto 1D lines. Only a small negligible computing overhead is incurred to train the network in low-dimensional capsule subspaces or through an alternative hyper-power iteration to estimate the normalization matrix. Experiment results on image datasets show the presented model can greatly improve the performance of the state-of-the-art ResNet backbones by 10-20% and that of the Densenet by 5-7% respectively at the same level of computing and memory expenses. The CapProNet establishes the competitive state-of-the-art performance for the family of capsule nets by significantly reducing test errors on the benchmark datasets.