

IS&T International Symposium on Electronic Imaging Science and Technology 2018

Visual Information Processing and
Communication IX

Burlingame, California, USA
28 January – 1 February 2018

Editors:

**Fengqing Maggie Zhu
Amy Reibman
Edward Delp**

ISBN: 978-1-5108-6933-2

Printed from e-media with permission by:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571



Some format issues inherent in the e-media version may also appear in this print version.

Copyright© (2018) by Society for Imaging Science & Technology
All rights reserved.

Printed by Curran Associates, Inc. (2018)

For permission requests, please contact Society for Imaging Science & Technology
at the address below.

Society for Imaging Science & Technology
7003 Kilworth Lane
Springfield, Virginia 22151
USA

Phone: 703-642-9090

Fax: 703-642-9094

info@imaging.org

Additional copies of this publication are available from:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: 845-758-0400
Fax: 845-758-2633
Email: curran@proceedings.com
Web: www.proceedings.com

Visual Information Processing and Communication IX

Monday, January 29, 2018

Keynote: Image and Video Compression

Session Chair: Zoe Liu, Google, Inc. (United States)

10:40 – 11:20 am

Sandpebble A

VIPC-123

Technical overview of AV1: An open source video codec from the Alliance for Open Media, Yaowu Xu, Google, Inc. (United States)

Dr. Yaowu Xu is currently the tech lead manager of the video coding research team at Google. The team has been responsible for developing and defining VP9, the core video technology of the WebM project. Prior to joining Google, Dr. Xu was the vice president of codec development at On2 Technologies. He was the co-creator of On2's VPx series codecs including VP32, VP4, VP5, VP6, VP7 and VP8. These codecs were broadly adopted by the industry and have fueled the phenomenal growth of web video. Dr. Xu's education includes a BS in physics, an MS and a PhD in nuclear engineering from Tsinghua University at Beijing, China. He also holds an MS and a PhD in electrical and computer engineering from the University of Rochester. Dr. Xu has published many technical papers in the area of image processing on leading journals and international conferences. He also holds many patents and has numerous patent applications pending in the area of digital video compression. Dr. Xu's research and development experiences include digital video compression and processing, real time video encoding and decoding, mobile video, image processing, pattern recognition and machine learning. His current research focuses on advanced algorithms for digital video compression. [no paper]

Image and Video Compression

Session Chair: Zoe Liu, Google, Inc. (United States)

11:20 am – 12:40 pm

Sandpebble A

11:20

VIPC-153

CrossEncoders: A complex neural network compression framework, Chirag Agarwal, Mehdi Sharifzadeh, and Dan Schonfeld, University of Illinois at Chicago (United States)

11:40

VIPC-154

Multi-level machine learning-based early termination in VP9 partition search, Yang Xian¹, Yunqing Wang², Yingli Tian¹, Yaowu Xu², and James Bankoski²; ¹The City University of New York and ²Google Inc. (United States)

12:00

VIPC-155

Texture segmentation based video compression using convolutional neural networks, Chichen Fu¹, Di Chen¹, Edward Delp¹, Zoe Liu², and Fengqing Zhu¹; ¹Purdue University and ²Google, Inc. (United States)

12:20

VIPC-156

Multi-reference video coding using stillness detection, Di Chen¹, Zoe Liu², Yaowu Xu², Fengqing Zhu¹, and Edward Delp¹; ¹Purdue University and ²Google, Inc. (United States)

12:40 – 2:00 pm Lunch

Plenary Session

2:00 – 3:00 pm

Grand Peninsula Ballroom D

Overview of Modern Machine Learning and Deep Neural Networks - Impact on Imaging and the Field of Computer Vision, Greg Corrado, Google, Inc. (United States)

Dr. Greg Corrado, co-founder of Google Brain, principal scientist, and director of augmented intelligence research at Google, provides an overview of modern machine learning and deep neural networks, with particular attention to its impact on imaging and the field of computer vision.

Dr. Corrado is a senior research scientist interested in biological neuroscience, artificial intelligence, and scalable machine learning. He has published in fields ranging across behavioral economics, neuromorphic device physics, systems neuroscience, and deep learning. At Google he has worked for some time on brain inspired computing, and most recently has served as one of the founding members and the co-technical lead of Google's large scale deep neural networks project. Prior to joining Google, Dr. Corrado was a staff research scientist at IBM. He received his MS in computer science and PhD in neuroscience from Stanford University.

3:00 – 3:30 pm Coffee Break

Deep Learning for Image Analysis

Session Chair: Edward Delp, Purdue University (United States)

3:30 – 4:50 pm

Sandpebble A

3:30

VIPC-173

Event recognition in personal photo collections: An active learning approach, Kashif Ahmad, Mohamed Mekhaili, and Nicola Conci, Università degli Studi di Trento (Italy)

3:50

VIPC-174

Generative adversarial networks for open set historical Chinese character recognition, Xiaoyi Yu, Jun Sun, and Satoshi Naoi, Fujitsu R&D Co. Limited (China)

4:10

VIPC-175

Transfer learning for data triage applications, Felix Mayer, Marcel Schaefer, and Martin Steinebach, Fraunhofer SIT (Germany)

4:30

VIPC-176

Approach for machine-printed Arabic character recognition: The-state-of-the-art deep-learning method, Daegun Ko, Changhyung Lee, Donghyeong Han, Hyeongsu Ohk, Ki-Min Kang, and Seong-Wook Han, HP Inc. (Republic of Korea)

5:00 – 6:00 pm All-Conference Welcome Reception

Tuesday, January 30, 2018

7:15 – 8:45 am Women in Electronic Imaging Breakfast

Computer Vision Applications in Sports

Session Chair: Fengqing Zhu, Purdue University (United States)

9:10 – 10:10 am

Sandpebble A

9:10 VIPC-205

Toward automatic and objective evaluation of synchronization in synchronized diving video, *Xin Li and Yixin Du, West Virginia University (United States)*

9:30 VIPC-206

Convolutional neural networks for the analysis of broadcasted tennis games, *Grigorios Tsagakatakis¹, Mustafa Jaber², and Panagiotis Tsakalides^{1,3}*; ¹Foundation for Research and Technology (FORTH) (Greece), ²NantVision Inc. (United States), and ³University of Crete (Crete)

9:50 VIPC-207 [no paper]

A self powered device embedded in a sports ball for a better immersive experience, *Rony Ferzli and Adel Elshebini, Intel Corporation (United States)*

10:00 am – 7:30 pm Industry Exhibition

10:10 – 10:40 pm Coffee Break

Keynote: Image and Video Analytics

Session Chair: Grigorios Tsagakatakis, Foundation for Research and Technology (FORTH) (Greece)

10:40 – 11:20 pm

Sandpebble A

VIPC-215

Perceptual optimization in video coding - a systematic approach, *Ioannis Katsavounidis, Netflix (United States)*

Dr. Ioannis Katsavounidis received the Diploma (BS/MS) from the Aristotle University of Thessaloniki, Greece, (1991) and his MS and PhD from the University of Southern California, Los Angeles, (1992 and 1998 respectively), all in electrical engineering. From 1996 to 2000, he worked in Italy as an engineer for the high-energy physics department of the California Institute of Technology. From 2000 to 2007, he worked at InterVideo, Inc., in Fremont, CA, as director of software for advanced technologies, in charge of MPEG2, MPEG4 and H.264 video codec development. Between 2007 and 2008, he served as CTO of Cidana, a mobile multimedia software company in Shanghai, China, covering all aspects of DTV standards and codecs. From 2008 to 2015 he was an associate professor with the department of electrical and computer engineering at the University of Thessaly in Volos, Greece, teaching undergraduate and graduate courses in signals, controls, image processing, video compression, and information theory. He is currently a senior research scientist at Netflix, working on video quality and video codec optimization problems. His research interests include image and video quality, compression and processing, information theory, and software-hardware optimization of multimedia applications. [no paper]

Image and Video Analytics

Session Chair: Grigorios Tsagakatakis, Foundation for Research and Technology (FORTH) (Greece)

11:20 am – 12:40 pm

Sandpebble A

11:20 VIPC-251

Interactive hand pose estimation: Boosting accuracy in localizing extended finger joints, *Cairong Zhang, Guijin Wang, Hengkai Guo, Xinghao Chen, Fei Qiao, and Huazhong Yang, Tsinghua University (China)*

11:40 VIPC-252

Efficient preprocessing and feature extraction for robust face recognition, two-level clustering, *Huda M.S. Algharib and Osman Serdar Gedik, Ankara Yildirim Beyazıt University (Turkey)*

12:00 VIPC-253

Text/figure separation in document images using Docstrum descriptor and two-level clustering, *Valery Anisimovskiy, Ilya Kurilin, Andrey Shcherbinin, and Petr Pohl, Samsung R&D Institute Russia (Russian Federation)*

12:20 VIPC-254

Prediction system for activity recognition with compressed video, *Chengzhang Zhong and Amy Reibman, Purdue University (United States)*

12:40 – 2:00 pm Lunch

Plenary Session

2:00 – 3:00 pm

Grand Peninsula Ballroom D

Fast, Automated 3D Modeling of Buildings and Other GPS Denied Environments, *Avideh Zakhor, University of California, Berkeley (United States)*

Professor Avideh Zakhor discusses fast, automated 3D modeling of buildings and other GPS denied environments with examples from her work in 3D reality capture, and visual and metric documentation of building interiors. Dr. Zakhor is a serial entrepreneur with startups in outdoor mapping, indoor mapping, and micro-lithography, currently CEO and founder of Indoor Reality, a Silicon Valley startup with products in 3D reality capture, and visual and metric documentation of building interiors.

Dr. Zakhor has been faculty member at University of California, Berkeley since 1994 where she holds the Qualcomm Chair in the electrical engineering and computer science department. She co-founded OPC technology in 1996, which was acquired by Mentor Graphics in 1998, and UrbanScan Inc. in 2005, acquired by Google in 2007. UrbanScan created the first fully automated 3D outdoor mapping system for 3D exterior models of buildings in urban environments. She has received a number of best paper awards in 3D computer vision, image processing, signal processing, is an IEEE fellow, and received the presidential young investigator award in 1992. Dr. Zakhor received her BSc in electrical engineering, from the California Institute of Technology (1983), and her MS (1985) and PhD (1987) in electrical engineering and computer science from MIT.

3:00 – 3:30 pm Coffee Break

Imaging Systems

Session Chair: Amy Reibman, Purdue University (United States)

3:30 – 4:30 pm

Sandpebble A

3:30 VIPC-263

A robust and accurate calibration method for out-of-focus camera,
*Xiaowei Hu¹, Guijin Wang¹, Jinnan Wang¹, Pengfei Sun¹, Jingtao Fan¹,
Feng Chen¹, and Yiyuan Xie²; ¹Tsinghua University and ²Southwest
University of Science and Technology (China)*

3:50 VIPC-264

**Improving the efficiency of on-site operators in utility management:
Combining hololens and AR for real-time check of electricity meters,**
*Lorenzo Orlandi¹, Daniele Sevegnani¹, and Nicola Conci²; ¹ARCODA
and ²Università degli Studi di Trento (Italy)*

4:10 VIPC-265

**Generation of stereoscopic image sequences from monocular videos
using epipolar geometry,** *Vasundhara Goyal and Dan Schonfeld,
University of Illinois at Chicago (United States)*

Symposium Demonstration Session

5:30 – 7:30 pm

Grand Peninsula Ballroom E

Wednesday, January 31, 2018

10:00 am – 4:00 pm Industry Exhibition

Plenary Session

2:00 – 3:00 pm

Grand Peninsula Ballroom D

Ubiquitous, Consumer AR Systems to Supplant Smartphones, *Ronald
T. Azuma, Intel, Corp. (United States)*

Dr. Ronald T. Azuma, researcher and augmented reality pioneer, shares his vision for achieving ubiquitous, consumer AR systems. Recent large investments in augmented reality reflect the commercial interest in its inherent potential to replace current smartphone technology, but much remains to be done. In his talk, Dr. Azuma gives a vision for achieving this goal, which requires not just solving numerous technical challenges but also determining new, compelling AR experiences that will establish AR as a new platform and novel form of media.

Dr. Azuma leads a team in Intel Labs that designs and prototypes novel experiences and key enabling technologies to enable new forms of media. These technology areas include computational imaging and photography, computational displays, and head-worn displays. Dr. Azuma is recognized as a pioneer and innovator in augmented reality, and has held prominent leadership roles in that research area, including leading and implementing research projects and demonstrations in areas such as AR, visualization, and mobile applications. Dr. Azuma received his BSc (1988) in electrical engineering from University of California, Berkeley, and MS (1990) and PhD (1995) in computer science from University of North Carolina, Chapel Hill. Prior to joining Intel, he was a research leader at Nokia Research Center Hollywood, and a senior researcher at Hughes Research Laboratories.

Symposium Interactive Papers (Poster) Session

5:30 – 7:30 pm

The Grove

The following works will be presented at the EI 2018 Symposium Interactive Papers Session.

VIPC-414

An interrupted projection using seam carving for 360-degree images,
Ikuko Tsubaki and Kazuo Sasaki, Tokyo University of Technology (Japan)

**Meet the Future: A Showcase of Student and Young Professionals
Research**

5:30 – 7:30 pm

The Grove