

# **2018 IEEE International Symposium on Multimedia (ISM 2018)**

**Taichung, Taiwan  
10 – 12 December 2018**



**IEEE Catalog Number: CFP18197-POD  
ISBN: 978-1-5386-6858-0**

**Copyright © 2018 by the Institute of Electrical and Electronics Engineers, Inc.  
All Rights Reserved**

*Copyright and Reprint Permissions:* Abstracting is permitted with credit to the source. Libraries are permitted to photocopy beyond the limit of U.S. copyright law for private use of patrons those articles in this volume that carry a code at the bottom of the first page, provided the per-copy fee indicated in the code is paid through Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923.

For other copying, reprint or republication permission, write to IEEE Copyrights Manager, IEEE Service Center, 445 Hoes Lane, Piscataway, NJ 08854. All rights reserved.

***\*\*\* This is a print representation of what appears in the IEEE Digital Library. Some format issues inherent in the e-media version may also appear in this print version.***

IEEE Catalog Number:	CFP18197-POD
ISBN (Print-On-Demand):	978-1-5386-6858-0
ISBN (Online):	978-1-5386-6857-3

**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  
E-mail: [curran@proceedings.com](mailto:curran@proceedings.com)  
Web: [www.proceedings.com](http://www.proceedings.com)

CURRAN ASSOCIATES INC.  
**proceedings**  
.com

# 2018 IEEE International Symposium on Multimedia (ISM) ISM 2018

## Table of Contents

Message from the General Co-Chairs .xii.....	
Message from the Program Co-Chairs .xiii.....	
Organizing Committee .xv.....	
Program Committee .xvii.....	

### Session 1: Bio-Related Applications

Tradeoffs Using Binary and Multiclass Neural Network Classification for Medical Multidisease Detection .1.....	
<i>Tor Jan Derek Berstad (ForzaSys AS), Michael Riegler (Simula Metropolitan Center for Digital Engineering, Norway), Håvard Espeland (ForzaSys AS), Thomas de Lange (Department of Transplantation, Oslo University Hospital, Norway), Pia Helen Smedsrud (Simula Research Laboratory), Konstantin Pogorelov (Simula Research Laboratory), Håkon Kvale Stensland (Simula Research Laboratory), and Pål Halvorsen (Simula Metropolitan Center for Digital Engineering, Norway)</i>	
A Robust Iris Segmentation Using Fully Convolutional Network with Dilated Convolutions .9.....	
<i>Yuting Yang (State Key Laboratory of Information Security, Institute of Information Engineering, Chinese Academy of Sciences), Peisong Shen (State Key Laboratory of Information Security, Institute of Information Engineering, Chinese Academy of Sciences), and Chi Chen (State Key Laboratory of Information Security, Institute of Information Engineering, Chinese Academy of Sciences)</i>	
Harnessing AI for Kidney Glomeruli Classification .17.....	
<i>Meghna Ayyar (IIIT Delhi), Puneet Mathur (NSIT), Rajiv Ratn Shah (IIIT Delhi), and Shree G Sharma (Arkana Laboratories)</i>	
Malignancy Classification of Lung Nodule Based on Accumulated Multi Planar Views and Canonical Correlation Analysis .21.....	
<i>Shimaa A. Abdelrahman (Egypt-Japan University of Science and Technology), Moataz M. Abdelwahab (Egypt-Japan University of Science and Technology), and Mohammed S. Sayed (Egypt-Japan University of Science and Technology)</i>	

## Session 2: Image, Video, and Other Applications

Voltage Controlled Oscillator with Impedance Spectroscopy for Non-invasive Glucose Application .25.....	
<i>Wen-Cheng Lai (National Penghu University of Science and Technology), Sheng-Lyang Jang (National Penghu University of Science and Technology), and Jyun-Jhih Wang (National Penghu University of Science and Technology)</i>	
MID: A Novel Contrast Metric for the MSER Detector .29.....	
<i>Martin Oelsch (Technical University of Munich), Basak Gulecyuz (Technical University of Munich), and Eckehard Steinbach (Technical University of Munich)</i>	
Gaze-Inspired Learning for Estimating the Attractiveness of a Food Photo .36.....	
<i>Akinori Sato (Nagoya University), Takatsugu Hirayama (Nagoya University), Keisuke Doman (Chukyo University), Yasutomo Kawanishi (Nagoya University), Ichiro Ide (Nagoya University), Daisuke Deguchi (Nagoya University), and Hiroshi Murase (Nagoya University)</i>	
Edge-Assisted Rendering of 360° Videos Streamed to Head-Mounted Virtual Reality .44.....	
<i>Wen-Chih Lo (National Tsing Hua University), Chih-Yuan Huang (National Tsing Hua University), and Cheng-Hsin Hsu (National Tsing Hua University)</i>	
Multi-modal Image Retrieval for Search-Based Image Annotation with RF .52.....	
<i>Petra Budikova (Masaryk University), Michal Batko (Masaryk University), and Pavel Zezula (Masaryk University)</i>	
REXplore: A Sketch Based Interactive Explorer for Real Estates Using Building Floor Plan Images .61.....	
<i>Divya Sharma (Indian Institute of Technology Jodhpur), Nitin Gupta (IBM Research Lab India), Chiranjoy Chattopadhyay (Indian Institute of Technology Jodhpur), and Sameep Mehta (IBM Research Lab India)</i>	
Tile-Based Rate Assignment for 360-Degree Video Based on Spatio-Temporal Activity Metrics .65.....	
<i>Robert Skupin (Fraunhofer HHI), Yago Sanchez (Fraunhofer HHI), Lei Jiao (Fraunhofer HHI), Cornelius Hellgo (Fraunhofer HHI), and Thomas Schierl (Fraunhofer HHI)</i>	
Content-Based Effectiveness Prediction of Video Advertisements .69.....	
<i>Qi Lou (University of California, Irvine), Somdeb Sarkhel (Adobe Research), Saayan Mitra (Adobe Research), and Viswanathan Swaminathan (Adobe Research)</i>	

## Session 3: Best Paper Candidates

Improved Image Description Via Embedded Object Structure Graph and Semantic Feature Matching .73.....	
<i>Li Ren (University of Central Florida) and Kien Hua (University of Central Florida)</i>	

Efficient Live and on-Demand Tiled HEVC 360 VR Video Streaming .81.....	
<i>Carsten Griwodz (University of Oslo), Mattis Jeppsson (ForzaSys AS), Håvard Espeland (ForzaSys AS), Tomas Kupka (ForzaSys AS), Ragnar Langseth (ForzaSys AS), Andreas Petlund (ForzaSys AS), Peng Qiaoqiao (Huawei), Chuansong Xue (Huawei), Konstantin Pogorelov (Simula Research Laboratory), Micheal Riegler (Simula Metropolitan Center for Digital Engineering, Norway), Dag Johansen (UIT - The Arctic University of Tromsø), and Pål Halvorsen (Simula Metropolitan Center for Digital Engineering)</i>	
HTTP/2-Based Streaming Solutions for Tiled Omnidirectional Videos .89.....	
<i>Mariem Ben Yahia (Orange Labs), Yannick Le Louedec (Orange Labs), Gwendal Simon (IMT-Atlantique), and Loutfi Nuaymi (IMT-Atlantique)</i>	

## Session 4: Deep Learning

SC-Conv: Sparse-Complementary Convolution for Efficient Model Utilization on CNNs .97.....	
<i>Chun-Fu (Richard) Chen (IBM T.J. Watson Research Center), Jinwook Oh (IBM T.J. Watson Research Center), Quanfu Fan (IBM T.J. Watson Research Center), and Marco Pistoia (IBM T.J. Watson Research Center)</i>	
Deep Reinforcement Learning with Parameterized Action Space for Object Detection .101.....	
<i>Zheng Wu (Ryerson University), Naimul Mefraz Khan (Ryerson University), Lei Gao (Ryerson University), and Ling Guan (Ryerson University)</i>	
Non-intrusive Estimation of Packet Loss Rates in Speech Communication Systems Using Convolutional Neural Networks .105.....	
<i>Gabriel Mittag (TU Berlin) and Sebastian Möller (TU Berlin, Language Technology, Deutsches Forschungszentrum für Künstliche Intelligenz (DFKI))</i>	
Neural Networks Based Fractional Pixel Motion Estimation for HEVC .110.....	
<i>Ehab M. Ibrahim (Egypt-Japan University for Science and Technology; Zagazig University, Egypt), Emad Badry (Egypt-Japan University for Science and Technology), Ahmed M. Abdelsalam (Polytechnique Montreal), Ibrahim L. Abdalla (Zagazig University, Egypt), Mohammed Sayed (Egypt-Japan University for Science and Technology; Zagazig University, Egypt), and Hossam Shalaby (Egypt-Japan University for Science and Technology; Alexandria University, Egypt)</i>	

## Session 5: Video Encoding & Quality

Towards Applying Game Adaptation to Decrease the Impact of Delay on Quality of Experience .114.....	
<i>Saeed Shafiee Sabet (Simula Research Laboratory), Steven Schmidt (Quality and Usability Lab, TU Berlin), Saman Zadtootaghaj (Telekom Innovation Labs), Carsten Griwodz (Simula Research Laboratory), and Sebastian Moller (Quality and Usability Lab, TU Berlin)</i>	

Investigation of Video Quality Dimensions for Different Type of Video Content .122.....	<i>Falk Schiffner (Quality and Usability Lab, TU Berlin, Germany), Bondarenko Vladimir (Quality and Usability Lab, TU Berlin, Germany), and Möller Sebastian (Quality and Usability Lab, TU Berlin, and German Research Center for Artificial Intelligence (DFKI))</i>
Geometry-Based Motion Vector Scaling for Omnidirectional Video Coding .127.....	<i>Ramin Ghaznavi-Youvalari (Nokia Technologies) and Alireza Aminlou (Nokia Technologies)</i>
NR-GVQM: A No Reference Gaming Video Quality Metric .131.....	<i>Saman Zadtootaghaj (Deutsche Telekom, Telekom Innovation Laboratories), Nabajeet Barman (Kingston University), Steven Schmidt (Technische Universität Berlin), Maria G. Martini (Kingston University), and Sebastian Möller (Technische Universität Berlin)</i>
Fast Line-Based Intra Prediction for Video Coding .135.....	<i>Santiago De-Luxán-Hernández (Fraunhofer HHI), Heiko Schwarz (Fraunhofer HHI), Detlev Marpe (Fraunhofer HHI), and Thomas Wiegand (Fraunhofer HHI)</i>
Light-Weight Video Coding Based on Perceptual Video Quality for Live Streaming .139.....	<i>Yusuke Sakamoto (Waseda University), Shintaro Saika (Waseda University), Masaru Takeuchi (Waseda University), Tatsuya Nagashima (Waseda University), Zhengxue Cheng (Waseda University), Kenji Kanai (Waseda University), Jiro Katto (Waseda University), Kaijin Wei (Huawei Technologies), Ju Zengwei (Huawei Technologies), and Xu Wei (Huawei Technologies)</i>

## Session 6: Audio, Music, Speech

Automatic Electronic Organ Reduction System Based on Melody Clustering Considering Melodic and Instrumental Characteristics .143.....	<i>Katunobu Itou (Hosei University) and Daiki Tanaka (Hosei University)</i>
MyLippper: A Personalized System for Speech Reconstruction using Multi-view Visual Feeds .151.....	<i>Yaman Kumar (Adobe), Rohit Jain (NSIT), Mohd Salik (NSIT), Rajiv ratn Shah (IIIT Delhi), Roger Zimmerman (NUS), and Yifang Yin (NUS)</i>
Spectrum Enhancement of Singing Voice Using Deep Learning .159.....	<i>Ryuka Nanzaka (Kobe University), Tsuyoshi Kitamura (Kobe University), Tetsuya Takiguchi (Kobe University), Yuji Adachi (MEC COMPANY LTD), and Kiyoto Tai (MEC COMPANY LTD.)</i>
Sound Recovery Considering the Vibration Direction of an Object in a Video .163.....	<i>Yohei Fuse (Kobe University), Yusuke Yasumi (Kobe University), and Tetsuya Takiguchi (Kobe University)</i>
Effect of Number of Stimuli on Users Perception of Different Speech Degradations. A Crowdsourcing Case Study .167.....	<i>Rafael Zequeira Jiménez (Technische Universität Berlin), Gabriel Mittag (Technische Universität Berlin), and Sebastian Möller (Technische Universität Berlin)</i>

### 3-Day Poster Showcases

Audio Feature Extraction Based on Sub-Band Signal Correlations for Music Genre Classification .172 .....	
<i>Takuya Kobayashi (Chuo University), Akira Kubota (Chuo University), and Yusuke Suzuki (Kawasaki City Office)</i>	
Live Demonstration: Kvazzip 4K HEVC Video Call .174 .....	
<i>Joni Räsänen (Tampere University of Technology), Marko Viitanen (Tampere University of Technology), Jarno Vanne (Tampere University of Technology), and Timo D. Hämmäläinen (Tampere University of Technology)</i>	
Eye-Controlled Region of Interest HEVC Encoding .176 .....	
<i>Joose Sainio (Tampere University of Tampere), Arttu Ylä-Outinen (Tampere University of Tampere), Marko Viitanen (Tampere University of Tampere), Jarno Vanne (Tampere University of Technology), and Timo Hämmäläinen (Tampere University of Technology)</i>	
Deep Learning of Human Perception in Audio Event Classification .178 .....	
<i>Yi Yu (National Institute of Informatics, SOKENDAI, Tokyo), Samuel Beuret (Ecole Polytechnique Fédérale de Lausanne), Donghuo Zeng (National Institute of Informatics, SOKENDAI, Tokyo), and Keizo Oyama (National Institute of Informatics, SOKENDAI, Tokyo)</i>	
Using Linear and Non-linear Magnifiers in Eyetracking-Based Human Computer Interaction .180 .....	
<i>Florian Schniederjann (HSW University of Applied Sciences), Jana Krahe (HSW University of Applied Sciences), Tobias Guth (HSW University of Applied Sciences), Johanna Wendel (HSW University of Applied Sciences), and Robert Mertens (HSW University of Applied Sciences)</i>	
A Text-Based CAPTCHA Cracking System with Generative Adversarial Networks .182 .....	
<i>Fan Liu (Hohai University), Zewen Li (Hohai University), Xueyi Li (Hohai University), and Tanyue Lv (Hohai University)</i>	

### Workshop: The Second IEEE International Workshop on Machine Learning and Computing for Visual Semantic Analysis (MLCSA)

Two Staged Machine Learning Network for Spine Segmentation and Recognition .184 .....	
<i>Pin-Hsien Liu (Asia University), Zhen-You Lian (Asia University), Chih-Yang Lin (Yuan Ze University), Cheng-Hung Chuang (Asia University), Chung-Lin Huang (Asia University), and Yuan-Yu Tsai (Asia University)</i>	
Comprehensive Study of Multiple CNNs Fusion for Fine-Grained Dog Breed Categorization .188 .....	
<i>Minori Uno (Yamaguchi University), Xian-Hua Han (Yamaguchi University), and Yen-Wei Chen (Ritsumeikan University)</i>	
Extraction of Movie Trailer Biases Based on Editing Features for Trailer Generation .194 .....	
<i>Honoka Kakimoto (Kwansei Gakuin University), Yuanyuan Wang (Yamaguchi University), Yukiko Kawai (Kyoto Sangyo University), and Kazutoshi Sumiya (Kwansei Gakuin University)</i>	
Mobile Scanner for Protein Crystallization Plates .199 .....	
<i>Ashok Shrestha (The University of Alabama in Huntsville), Truong Tran (The University of Alabama in Huntsville), Ramazan Aygun (The University of Alabama in Huntsville), and Marc Pusey (iXpressGenes, Inc.)</i>	

## Workshop: Multimodal Representation, Retrieval, and Analysis of Multimedia Content in Social Media (MR2ARMC)

A Novel Relative Camera Motion Estimation Algorithm with Applications to Visual Odometry .205.....	<i>Yue Jiang (Korea University), Mun-Cheon Kang (Korea University), Ming Fan (Korea University), Sung-Ho Chae (Korea University), and Sung-Jea Ko (Korea University)</i>
IceBreaker: Solving Cold Start Problem for Video Recommendation Engines .207.....	<i>Yaman Kumar (Adobe), Agniv Sharma (MIDAS Lab, DTU-Delhi), Abhigyan Khaund (MIDAS Lab, IIT-Mandi), Akash Kumar (MIDAS Lab, DTU-Delhi), Ponnurangam Kumaraguru (IIIT-Delhi), Rajiv Ratn Shah (IIIT-Delhi), and Roger Zimmermann (NUS-Singapore)</i>
Towards Improved Human Action Recognition Using Convolutional Neural Networks and Multimodal Fusion of Depth and Inertial Sensor Data .213.....	<i>Zeeshan Ahmad (Ryerson University) and Naimul Khan (Ryerson University)</i>
Supporting Multimedia Retrieval in Annotated Content using Hyperknowledge .221.....	<i>Marcio Ferreira Moreno (IBM Research), Wallas Henrique Sousa dos Santos (IBM Research), Rodrigo Costa Mesquita Santos (IBM Research), Patricia Torres Pereira Carrion (IBM Research), and Renato Fontoura de Gusmão Cerqueira (IBM Research)</i>
Supporting Knowledge Creation through HAS: The Hyperknowledge Annotation System .229.....	<i>Marcio Ferreira Moreno (IBM Research), Wallas Henrique Sousa Dos Santos (IBM Research), Rodrigo Costa Mesquita Santos (IBM Research), and Renato Fontoura De Gusmao Cerqueira (IBM Research)</i>

## Workshop: Multimedia Technologies for E-Learning (MTEL)

Player Types in Mobile Learning Games – Playing Patterns and Motivation .237.....	<i>Florian Schimanke (HSW - University of Applied Sciences), Robert Mertens (HSW - University of Applied Sciences), and Bettina Sophie Huck (HSW - University of Applied Sciences)</i>
3D Convolutional Network Based Foreground Feature Fusion .243.....	<i>Hanjian Song (Xi'an Jiaotong University Xi'an, China), Lihua Tian (Xi'an Jiaotong University Xi'an, China), and Chen Li (Xi'an Jiaotong University Xi'an, China)</i>
IBAtS - Image Based Attendance System: A Low Cost Solution to Record Student Attendance in a Classroom .249.....	<i>Setia Budi (Maranatha Christian University), Oscar Karnalim (Maranatha Christian University), Erico D. Handoyo (Maranatha Christian University), Sulaeman Santoso (Maranatha Christian University), Hapnes Toba (Maranatha Christian University), Huyen Nguyen (University of New South Wales), and Vishv Malhotra (Indian Institute of Technology Guwahati)</i>



# Workshop: The First IEEE International Workshop on State-of-the-art Speech Technologies in Multimedia and Mobile Environments (STeMME)

Acoustic Scene Classification Using Reduced MobileNet Architecture .257.....	
<i>Jun-Xiang Xu (National Central University), Tzu-Ching Lin (National Central University), Tsai-Ching Yu (National Central University), Tzu-Chiang Tai (National Central University), and Pao-Chi Chang (National Central University)</i>	
U-Healthcare System for Pre-Diagnosis of Parkinson's Disease from Voice Signal .261.....	
<i>Sylvio Barbon Junior (State University of Londrina), Victor G. Turrisi Costa (State University of Londrina), Shi-Huang Chen (Shu-Te University), and Rodrigo Capobianco Guido (São Paulo State University)</i>	
Single-Channel Speech Separation Based on Gaussian Process Regression .265.....	
<i>Nguyen Le (National Central University), Sih-Huei Chen (National Central University), Tzu-Chiang Tai (Providence University), and Jia-Ching Wang (National Central University)</i>	
<b>Author Index</b> 269.....	