

2017 IEEE International Symposium on Multimedia (ISM 2017)

**Taichung, Taiwan
11-13 December 2017**



**IEEE Catalog Number: CFP17197-POD
ISBN: 978-1-5386-2938-3**

**Copyright © 2017 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:	CFP17197-POD
ISBN (Print-On-Demand):	978-1-5386-2938-3
ISBN (Online):	978-1-5386-2937-6

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
Web: www.proceedings.com

CURRAN ASSOCIATES INC.
proceedings
.com

2017 IEEE International Symposium on Multimedia

ISM 2017

Table of Contents

Message from the General Co-Chairs	xiv
Message from the Program Co-Chairs.....	xvi
Organizing Committee	xviii
Program Committee.....	xx
Reviewers.....	xxiv

Session 1: Diverse Topics

Boundary-Preserving Depth Upsampling Without Texture Copying Artifacts and Holes	1
<i>Jiji Cai, Liang Chang, Hongbin Wang, Cheolkon Jung, and Joongkyu Kim</i>	
Multi-generation-robust Coding with JPEG XS	6
<i>Thomas Richter, Joachim Keinert, Antonin Descampe, Gael Rouvroy, and Alexandre Willeme</i>	
Occlusion Robust Light Field Depth Estimation Using Segmentation Guided Bilateral Filtering	14
<i>Qiuxia Hou and Cheolkon Jung</i>	
Priced-Based Fair Bandwidth Allocation for Networked Multimedia	19
<i>Hamed Hamzeh, Mahdi Hemmati, and Shervin Shirmohammadi</i>	
Performance Evaluation of Walking Imagery Training Based on Virtual Environment in Brain-Computer Interfaces	25
<i>Xiaolu Liu, Shuang Liang, Wenlong Hang, Baiying Lei, Qiong Wang, Jing Qin, and Kup-Sze Choi</i>	
Image Based Localization Based on Feature Scale Consistency in BOF Vector	31
<i>Mahdi Salarian, Mehdi Sharifzadeh, and Rashid Ansari</i>	

Session 2: 360° Video & Image

A New Adaptation Approach for Viewport-adaptive 360-degree Video Streaming	38
<i>Duc V. Nguyen, Huyen T. T. Tran, Anh T. Pham, and Truong Cong Thang</i>	
Estimation of Optimal Encoding Ladders for Tiled 360° VR Video in Adaptive Streaming Systems	45
<i>Cagri Ozcinar, Ana De Abreu, Sebastian Knorr, and Aljosa Smolic</i>	
Perceptual Analysis of Perspective Projection for Viewport Rendering in 360° Images	53
<i>Falah Jabar, João Ascenso, and Maria Paula Queluz</i>	

Session 3: Learning

Convolutional DLSTM for Crowd Scene Understanding	61
<i>Naifan Zhuang, Jun Ye, and Kien A. Hua</i>	
Recurrent Visual Relationship Recognition with Triplet Unit	69
<i>Kento Masui, Akiyoshi Ochiai, Shintaro Yoshizawa, and Hideki Nakayama</i>	
A Pre-Saliency Map Based Blind Image Quality Assessment via Convolutional Neural Networks	77
<i>Zhengxue Cheng, Masaru Takeuchi, and Jiro Katto</i>	
Human Action Classification Using Temporal Slicing for Deep Convolutional Neural Networks	83
<i>Nathan Henderson and Ramazan Aygun</i>	
Rate-Accuracy Optimization of Deep Convolutional Neural Network Models	91
<i>Alessandro Filini, João Ascenso, and Riccardo Leonardi</i>	
Automatic Classification of Microstructures in Thermal Barrier Coating Images	99
<i>Wei-Bang Chen, Yongjin Lu, James Li, and Ben Zimmerman</i>	

Session 4: Visual Aspects

Sustained Attention Function Evaluation During Cooking Based on Egocentric Vision	107
<i>Sho Ooi, Mutsuo Sano, Hajime Tabuchi, Fumie Saito, and Satoshi Umeda</i>	
Detecting Good Surface for Improvisatory Visual Projection	114
<i>Hoang Le, Thong Doan, Carl S. Marshall, Selvakumar Panneer, and Feng Liu</i>	
Coherent Visual Description of Textual Instructions	122
<i>Shashank Mujumdar, Nitin Gupta, Abhinav Jain, and Sameep Mehta</i>	
QoE Studies on Interactive 3D Tele-Immersion	130
<i>Kevin Desai, Suraj Raghuraman, Rong Jin, and Balakrishnan Prabhakaran</i>	

Spatio-Temporal Compositing of Video Elements for Immersive eLearning Classrooms	138
<i>Uma Gopalakrishnan, P. Venkat Rangan, Ramkumar N, and Balaji Hariharan</i>	

Session 5: Best Papers

Deep Attribute Driven Image Similarity Learning Using Limited Data	146
<i>Nitin Gupta, Ankush Gupta, Vikas Joshi, L. Venkata Subramaniam, and Sameep Mehta</i>	
A Real-Time Annotation of Motion Data Streams	154
<i>Petr Elias, Jan Sedmidubsky, and Pavel Zezula</i>	
Heterogeneous Features Fusion with Collaborative Representation Learning for 3D Action Recognition	162
<i>Chengwu Liang, Enqing Chen, Lin Qi, and Ling Guan</i>	
Towards Efficient 3D Pose Retrieval and Reconstruction from 2D Landmarks	169
<i>Hashim Yasin</i>	
Kara1k: A Karaoke Dataset for Cover Song Identification and Singing Voice Analysis	177
<i>Yann Bayle, Ladislav Maršík, Martin Rusek, Matthias Robine, Pierre Hanna, Katerina Slaninová, Jan Martinovic, and Jaroslav Pokorný</i>	
Computational and Perceptual Determinants of Film Mood in Different Types of Scenes	185
<i>Jussi Tarvainen, Jorma Laaksonen, and Tapio Takala</i>	

Session 6: Retrieval & Mining

Summarization of News Videos Considering the Consistency of Auditory and Visual Contents	193
<i>Ichiro Ide, Ye Zhang, Ryunosuke Tanishige, Keisuke Doman, Yasutomo Kawanishi, Daisuke Deguchi, and Hiroshi Murase</i>	
An Iterative Feature-Pair Updating Framework for Rigid Template Matching with Outliers	200
<i>Yang Yang, Qian Kou, Shaoyi Du, Shuang Luo, Yuehu Liu, and Bangyu Wu</i>	
Mining Culture-Specific Music Listening Behavior from Social Media Data	208
<i>Martin Pichl, Eva Zangerle, Günther Specht, and Markus Schedl</i>	
Static vs. Dynamic Content Descriptors for Video Retrieval in Laparoscopy	216
<i>Bernd Muenzer, Manfred J. Primus, Sabrina Kletz, Stefan Petscharnig, and Klaus Schoeffmann</i>	
A Cloud-Based Multi-threaded Implementation of View Synthesis System	224
<i>Parvaneh Pouladzadeh, Razib Iqbal, Shervin Shirmohammadi, and Omid Fatemi</i>	

In-Memory Distributed Indexing for Large-Scale Media Data Retrieval	232
<i>Yinmiao Ma, Danlu Liu, Grant Scott, Jeffrey Uhlmann, and Chi-Ren Shyu</i>	

Session 7-A: Retrieval, Recommendation, and Summarization

Enhancing Effectiveness of Descriptors for Searching and Recognition in Motion Capture Data	240
<i>Jan Sedmidubsky, Petr Elias, and Pavel Zezula</i>	
Blog Article Summarization with Image-Text Alignment Techniques	244
<i>Wei-Ta Chu and Ming-Chih Kao</i>	
Minimum-Risk Structured Learning of Video Summarization	248
<i>Fairoz Hussain and Massimo Piccardi</i>	
Feature Selection for FM-Based Context-Aware Recommendation Systems	252
<i>Xueyu Mao, Saayan Mitra, and Viswanathan Swaminathan</i>	
Personalized Video Recommendations for Shared Accounts	256
<i>Shuo Yang, Somdeb Sarkhel, Saayan Mitra, and Viswanathan Swaminathan</i>	

Session 7-B: Tracking & Matching

Moment-Based Dense Correspondence Matching Robust to Image Variation	260
<i>Inyong Yun, Seokhoon Boo, Joongkyu Kim, and Cheolkon Jung</i>	
Very Small Moving Objects Detection in Videos by Means of Fuzzy Logic and Reliability Coefficients: Application to Migrating Birds Counting	264
<i>Henri Nicolas and Valentina Guerin Detourville</i>	
A Gaze Tracking Based, Multi Modal Human Computer Interaction Concept for Efficient Input	268
<i>Leon Strapper, Robert Mertens, Sebastian Pospiech, Florian Bussmann, Arthur Grah, and Marius Mamsch</i>	
Robust and Fast Object Tracking for Challenging 360-degree Videos	274
<i>Ahmad Delforouzi and Marcin Grzegorzec</i>	
Temporal Matching Kernel with Embedded Stability-Sensitive Filter	278
<i>Fan Yang, Sbastien Poullot, and Shin'ichi Satoh</i>	

Session 7-C: Mining & Learning

A Color Prediction System for Interactive Drawing Based Image Retrieval on Mobile Devices	284
<i>Zhan Xu and Guoping Qiu</i>	
VenueNet: Fine-Grained Venue Discovery by Deep Correlation Learning	288
<i>Yi Yu, Suhua Tang, Kiyoharu Aizawa, and Akiko Aizawa</i>	
Adaptive Sparse Learning for Neurodegenerative Disease Classification	292
<i>Haijun Lei, Yujia Zhao, Yuting Wen, and Baiying Lei</i>	

A New Multimedia Documents Clustering Approach Based on Feature Patterns Similarity	296
<i>Pushpalatha K and Ananthanarayana V. S.</i>	
User Segment Identification Based on Similarity in Content Consumption	300
<i>Somdeb Sarkhel, Wreetabrata Kar, and Viswanathan Swaminathan</i>	

Session 8-A: Contents & Features

Automatic Selection of Web Contents Towards Automatic Authoring of a Video Biography	304
<i>Ichiro Ide, Yasutomo Kawanishi, Kyoka Kunishiro, Frank Nack, Daisuke Deguchi, and Hiroshi Murase</i>	
Indicators of Country Similarity in Terms of Music Taste, Cultural, and Socio-economic Factors	308
<i>Markus Schedl, Florian Lemmerich, Bruce Ferwerda, Marcin Skowron, and Peter Knees</i>	
DAMPAT: Dynamic Adaptation of Multimedia Presentations in Application Mobility	312
<i>Francisco Javier Velázquez-García and Frank Eliassen</i>	
Drawing Abrasive Hologram Animations with Auto-Generated Scratch Patterns	318
<i>Frode Eika Sandnes and Evelyn Eika</i>	
Survey of Visual Feature Extraction Algorithms in a Mars-like Environment	322
<i>Martin Oelsch, Dominik Van Opdenbosch, and Eckehard Steinbach</i>	

Session 8-B: Video Streaming

Streaming Multimedia via Overlay Networks Using Wi-Fi Peer-to-Peer Connections	326
<i>Justas Poderys and Jose Soler</i>	
SDN-enabled Game-Aware Network Management for Residential Gateways	330
<i>Maryam Amiri, Hussein Al Osman, and Shervin Shirmohammadi</i>	
Balancing Transcoding Against Quality-of-Experience to Limit Energy Consumption in Video-on-Demand Systems	334
<i>Jungwoo Lee, Hwangje Han, and Minseok Song</i>	
A Bitrate-Conservative Fast-Adjusting Rate Controller for Video Conferencing	338
<i>Abbas Javadtalab, Mona Omidyeganeh, Shervin Shirmohammadi, and Mojtaba Hosseini</i>	

Session 8-C: Enhancement & Security

Readability Enhancement of Low Light Videos Based on Discrete Wavelet Transform	342
<i>Tingting Sun, Cheolkon Jung, Peng Ke, Hyoseob Song, and Jungmee Hwang</i>	
The Impact of Feng Shui on House Price: A Data Perspective	346
<i>Tsuo-Chen Wu and Mei-Chen Yeh</i>	
Adaptive Quantization-Based HDR Video Coding with HEVC Main 10 Profile	350
<i>Qingtao Fu, Cheolkon Jung, and Ge Yang</i>	
Analyzing and Enhancing an Image Authentication Scheme	354
<i>Seyed Amir Hossein Tabatabaei, Ahmad Delforouzi, and Marcin Grzegorzek</i>	
Secure Secret Sharing in the Cloud	358
<i>Ching-Chun Chang and Chang-Tsun Li</i>	

Demo I: Video Related

Kvazaar: HEVC/H.265 4K30p Intra Encoder	362
<i>Arttu Ylä-Outinen, Ari Lemmetti, Marko Viitanen, Jarno Vanne, and Timo D. Hämmäläinen</i>	
Endometriosis Annotation in Endoscopic Videos	364
<i>Andreas Leibetseder, Bernd Münzer, Klaus Schoeffmann, and Jörg Keckstein</i>	
EndoXplore: A Web-Based Video Explorer for Endoscopic Videos	366
<i>Bernd Muenzer, Klaus Schoeffmann, and Laszlo Böeszöermenyi</i>	
The One Man Show	368
<i>Sai Samarth R Phaye, Love Mehta, and Mukesh Saini</i>	

Demo II: Interaction, Tracking, Network Related

A Multi Modal Interaction Paradigm Combining Gaze Tracking and Keyboard	370
<i>Luisa Brinkschulte, Robert Mertens, Leon Strapper, Sebastian Pospiech, and Lars Knipping</i>	
A Web Application for Subsequence Matching in 3D Human Motion Data	372
<i>Jan Sedmidubsky and Pavel Zezula</i>	
A History-Based TCP Throughput Prediction Incorporating Communication Quality Features by Support Vector Regression for Mobile Network	374
<i>Bo Wei, Wataru Kawakami, Kenji Kanai, and Jiro Katto</i>	
Vision-Based Automatic Identification Tracking of Steel Products for Intelligent Steel Manufacturing	376
<i>Chao-Yung Hsu, Li-Wei Kang, Teng-Yi You, and Wei-Chen Jhong</i>	

Workshop: Emerging Multimedia Applications and Services for Smart Cities (EMASC 2017)

Detection and Visualization of Arabic Emotions on Social Emotion Map	378
<i>Mohammed F. Alhamid, Saad Alsahli, Majdi Rawashdeh, and Mubarak Alrashoud</i>	
A Crowd-Sourced Adaptive Safe Navigation for Smart Cities	382
<i>Neeraj Goel, Rajat Sharma, N. Nikhil, S. D. Mahanoor, and Mukesh Saini</i>	
The Solar System as a 3D Metaphor to Visualize User Interactions in a Social Network	388
<i>Wanjun Pei, Benjamin Guthier, and Abdulmotaieb El Saddik</i>	
Privacy Analysis of Smart City Healthcare Services	394
<i>Amjad A. Alghanim, Sk Md Mizanur Rahman, and M. Anwar Hossain</i>	
Impact of Distributed Caching on Video Streaming Quality in Information Centric Networks	399
<i>Lingchao Kong, Jingyi Zhu, Rui Dai, and Mohammad Nazmus Sadat</i>	

Workshop: Intelligent Multimedia Applications and Design for Quality Living (IMAD 2017)

MAESTRO: Constructing a Reference Framework for Self Monitoring Devices Dedicated to Seniors	403
<i>Moussa Ouedrago, Wassila Aggoune-Mtalaa, and Djamel Khadraoui</i>	
Shall IoT User Interfaces Start Recommending Multimedia Devices as Well?	407
<i>Mukesh Kumar Saini, Ali Danesh, and Abdulmotaieb El Saddik</i>	
Smart Homes and Quality of Life for the Elderly: A Systematic Review	413
<i>Debajyoti Pal, Tuul Triyason, and Suree Funikul</i>	
Real-Time System for Human Activity Analysis	420
<i>Randy Tan, Naimul Khan, and Ling Guan</i>	
Does Spending More Time on Facebook Makes Users Engage in Politics?	426
<i>Ho-Yin Yue, Wai-Man Pang, Chiu-Yin Tam, and Clive Wai-Ngok Fan</i>	
A Memory-Friendly Multi-modal Emotion Analysis for Smart Toy	432
<i>Geoffrey Poon, Ki-Mei Li, and Wai-Man Pang</i>	
Where2Buy: A Location-Based Shopping App with Products-wise Searching	438
<i>Kin Chi Chan, Tak Leung Cheung, Siu Hong Lai, Kin Chung Kwan, Hoyin Yue, and Wai-Man Pang</i>	
Towards Interactive and Realistic Rendering of 3D Fetal Ultrasound via Photon Mapping	444
<i>Jinta Zheng, Jing Qin, and Kup-Sze Choi</i>	
Enhancing the Performance of Brain-Computer Interface with Haptics	450
<i>Kup-Sze Choi and Shuang Liang</i>	

QoS and QoE Evaluations of 2K and 4K DASH Contents Distributions	453
<i>Tatsuya Nagashima, Kenji Kanai, and Jiro Katto</i>	
Hand Gesture Recognition Based on Wavelet Invariant Moments	459
<i>Xi Liu, Chen Li, and Lihua Tian</i>	
Cross-Modal Transfer Learning for HEp-2 Cell Classification Based on Deep Residual Network	465
<i>Haijun Lei, Tao Han, Weifeng Huang, Jong Yih Kuo, Zhen Yu, Xinzi He, and Baiying Lei</i>	

Workshop: Mining and Applications on Multimedia (MAM 2017)

Data Mining Techniques vs. Policy Development: Evaluating Advanced Applied Technological Policies and Emerging Communication Technology	469
<i>Sheng-Chih Chen, Yi-Cheng Chen, and Wei-Lin Chen</i>	
Study of Touch Identify for Mobile Device Security	475
<i>Chung-Hua Chu, Hsiao-Ting Shih, and Chih-Hua Tai</i>	
Large-Scale Analysis of Group-Specific Music Genre Taste from Collaborative Tags	479
<i>Markus Schedl and Bruce Ferwerda</i>	
Using Stacked Denoising Autoencoder for the Student Dropout Prediction	483
<i>Jong Yih Kuo, Chia Wei Pan, and Baiying Lei</i>	

Workshop: Machine Learning and Computing for Visual Semantic Analysis (MLCSA 2017)

A Video Shot Boundary Detection Approach Based on CNN Feature	489
<i>Rui Liang, Qingxin Zhu, Honglei Wei, and Shujiao Liao</i>	
Deep Image Retrieval Applied on Kotenseki Ancient Japanese Literature	495
<i>Chairath Sirirattanapol, Yusuke Matsui, Shin'ichi Satoh, Kuninori Matsuda, and Kazuaki Yamamoto</i>	
Hyper-spectral Image Super-resolution Using Non-negative Spectral Representation with Data-Guided Sparsity	500
<i>Xian-Hua Han, Jan Wang, Boxin Shi, YinQiang Zheng, and Yen-Wei Chen</i>	
Tensor Sparse Representation of Temporal Features for Content-Based Retrieval of Focal Liver Lesions Using Multi-phase Medical Images	507
<i>Jian Wang, Xian-Hua Han, Yingying Xu, Lanfen Lin, Hongjie Hu, Chongwu Jin, and Yen-Wei Chen</i>	
Influence of Video Quality on Multi-view Activity Recognition	511
<i>Jun-Ho Choi, Manri Cheon, and Jong-Seok Lee</i>	
Context-aware Image Generation by Using Generative Adversarial Networks	516
<i>Kenki Nakamura and Qiang Ma</i>	

Workshop: Multimedia Search and Applications (MSA 2017)

A New Governance Architecture for Government Information Resources Based on Big Data Ecological Environment in China	524
<i>Xue Sixin, Yan Yayuan, Yang Ji Jiang, and Wang Yue</i>	
Distributed Representation for Neighborhood-Based Collaborative Filtering	531
<i>Yi Yang, Guigang Zhang, Jian Wang, and Weixing Huang</i>	
A Real-Time Chinese Calligraphy Creation System	536
<i>Yanzhou Gong, Ziqiang Ni, Weixing Huang, Jian Wang, and Guigang Zhang</i>	
A Robust Hand Cursor Interaction Method Using Kinect	543
<i>Zhiwen Lei, Xiaoxiao Yang, Yanzhou Gong, Weixing Huang, Jian Wang, and Guigang Zhang</i>	
Kvazzup: Open Software for HEVC Video Calls	549
<i>Joni Räsänen, Marko Viitanen, Jarno Vanne, and Timo D. Hämäläinen</i>	
LBP-SVD Based Copy Move Forgery Detection Algorithm	553
<i>Yuan Wang, Lihua Tian, and Chen Li</i>	
Key Frame Extraction Based on Entropy Difference and Perceptual Hash	557
<i>Mi Zhang, Lihua Tian, and Chen Li</i>	
Fast Binary Descriptor Search for Keypoint Matching by Norm Ordering	561
<i>Masahiko Sugimura, Takayuki Baba, and Ryuta Tanaka</i>	

Workshop: Multimedia Technologies for E-Learning (MTEL 2017)

Automated Assessment of Students' Conceptual Understanding: Supporting Students and Teachers Using Data from an Interactive Textbook	567
<i>Toby Dragon and Carrie Lindeman</i>	
Keyword Based Indexing of a Multimedia File	573
<i>Aman Chaudhary, Akshatha K, Kiran Kodlekere, and SuryaPrasad J</i>	
A Unit Testing Framework for Context Variant Code in a Mobile Learning App	577
<i>Florian Schimanke, Robert Mertens, and Leonard Hill</i>	
A Fast Video Shot Boundary Detection Employing OTSU's Method and Dual Pauta Criterion	583
<i>Zhe Yang, Lihua Tian, and Chen Li</i>	
Enabling Near Real-Time Collaboration in a Distributed Multimedia Editing Environment	587
<i>Johannes Klein, Jean Botev, and Steffen Rothkugel</i>	
Classification of Reading Patterns Based on Gaze Information	595
<i>Wen-Hung Liao, Chin-Wen Chang, and Yi-Chieh Wu</i>	
Author Index	601