

*Commenced Publication in 1973*

Founding and Former Series Editors:

Gerhard Goos, Juris Hartmanis, and Jan van Leeuwen

## Editorial Board

David Hutchison

*Lancaster University, Lancaster, UK*

Takeo Kanade

*Carnegie Mellon University, Pittsburgh, PA, USA*

Josef Kittler

*University of Surrey, Guildford, UK*

Jon M. Kleinberg

*Cornell University, Ithaca, NY, USA*

Friedemann Mattern

*ETH Zurich, Zurich, Switzerland*

John C. Mitchell

*Stanford University, Stanford, CA, USA*

Moni Naor

*Weizmann Institute of Science, Rehovot, Israel*

C. Pandu Rangan

*Indian Institute of Technology, Madras, India*

Bernhard Steffen

*TU Dortmund University, Dortmund, Germany*

Demetri Terzopoulos

*University of California, Los Angeles, CA, USA*

Doug Tygar

*University of California, Berkeley, CA, USA*

Gerhard Weikum

*Max Planck Institute for Informatics, Saarbrücken, Germany*

More information about this series at <http://www.springer.com/series/7409>

Laurent Amsaleg · Gylfi Þór Guðmundsson  
Cathal Gurrin · Björn Þór Jónsson  
Shin'ichi Satoh (Eds.)

# MultiMedia Modeling

23rd International Conference, MMM 2017  
Reykjavik, Iceland, January 4–6, 2017  
Proceedings, Part II

*Editors*

Laurent Amsaleg  
CNRS–IRISA  
Rennes  
France

Gylfi Þór Guðmundsson  
Reykjavík University  
Reykjavík  
Iceland

Cathal Gurrin  
Dublin City University  
Dublin  
Ireland

Björn Þór Jónsson  
Reykjavík University  
Reykjavík  
Ireland

Shin'ichi Satoh  
National Institute of Informatics  
Tokyo  
Japan

ISSN 0302-9743                      ISSN 1611-3349 (electronic)  
Lecture Notes in Computer Science  
ISBN 978-3-319-51813-8              ISBN 978-3-319-51814-5 (eBook)  
DOI 10.1007/978-3-319-51814-5

Library of Congress Control Number: 2016962021

LNCS Sublibrary: SL3 – Information Systems and Applications, incl. Internet/Web, and HCI

© Springer International Publishing AG 2017

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Printed on acid-free paper

This Springer imprint is published by Springer Nature  
The registered company is Springer International Publishing AG  
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

# Preface

These proceedings contain the papers presented at MMM 2017, the 23rd International Conference on MultiMedia Modeling, held at Reykjavik University during January 4–6, 2017. MMM is a leading international conference for researchers and industry practitioners for sharing new ideas, original research results, and practical development experiences from all MMM related areas, broadly falling into three categories: multimedia content analysis; multimedia signal processing and communications; and multimedia applications and services.

MMM conferences always include special sessions that focus on addressing new challenges for the multimedia community. The following four special sessions were held at MMM 2017:

- SS1: Social Media Retrieval and Recommendation
- SS2: Modeling Multimedia Behaviors
- SS3: Multimedia Computing for Intelligent Life
- SS4: Multimedia and Multimodal Interaction for Health and Basic Care Applications

MMM 2017 received a total 198 submissions across four categories; 149 full-paper submissions, 34 special session paper submissions, eight demonstration submissions, and seven submissions to the Video Browser Showdown (VBS 2017). Of all submissions, 68% were from Asia, 27% from Europe, 3% from North America, and 1% each from Oceania and Africa.

Of the 149 full papers submitted, 35 were selected for oral presentation and 33 for poster presentation, which equates to a 46% acceptance rate overall. Of the 34 special session papers submitted, 24 were selected for oral presentation and two for poster presentation, which equates to a 76% acceptance rate overall. In addition, five demonstrations were accepted from eight submissions, and all seven submissions to VBS 2017. The overall acceptance percentage across the conference was thus 54%, but 46% for full papers and 23% of full papers for oral presentation.

The submission and review process was coordinated using the ConfTool conference management software. All full-paper submissions were reviewed by at least three members of the Program Committee. All special session papers were reviewed by at least three reviewers from the Program Committee and special committees established for each special session. All demonstration papers were reviewed by at least three reviewers, and VBS papers by two reviewers. We owe a debt of gratitude to all these reviewers for providing their valuable time to MMM 2017.

We would like to thank our invited keynote speakers, Marcel Worring from the University of Amsterdam, The Netherlands, and Noriko Kando from the National Institute of Informatics, Japan, for their stimulating contributions.

We also wish to thank our organizational team: Demonstration Chairs Esra Acar and Frank Hopfgartner; Video Browser Showdown Chairs Klaus Schoeffmann, Werner Bailer, Cathal Gurrin and Jakub Lokoč; Sponsorship Chairs Yantao Zhang and Tao Mei;

Proceedings Chair Gylfi Þór Guðmundsson; and Local Organization Chair Marta Kristín Lárusdóttir.

We would like to thank Reykjavik University for hosting MMM 2017. Finally, special thanks go to our supporting team at Reykjavik University (Arnar Egilsson, Ýr Gunnlaugsdóttir, Þórunn Hilda Jónasdóttir, and Sigrún Heba Ómarsdóttir) and CP Reykjavík (Kristjana Magnúsdóttir, Elísabet Magnúsdóttir and Ingibjörg Hjálmsfríðardóttir), as well as to student volunteers, for all their contributions and valuable support.

The accepted research contributions represent the state of the art in multimedia modeling research and cover a very diverse range of topics. A selection of the best papers will be invited to submit extended versions to a special issue of *Multimedia Tools and Applications*. We wish to thank all authors who spent their valuable time and effort to submit their work to MMM 2017. And, finally, we thank all those who made the (sometimes long) trip to Reykjavík to attend MMM 2017 and VBS 2017.

January 2017

Björn Þór Jónsson  
Cathal Gurrin  
Laurent Amsaleg  
Shin'ichi Satoh  
Gylfi Þór Guðmundsson

# Organization

## Organizing Committee

### General Chairs

Björn Þór Jónsson  
Cathal Gurrin

Reykjavik University, Iceland  
Dublin City University, Ireland

### Program Chairs

Laurent Amsaleg  
Shin'ichi Satoh

CNRS-IRISA, France  
NII, Japan

### Demonstration Chairs

Frank Hopfgartner  
Esra Acar

University of Glasgow, UK  
Technische Universität Berlin, Germany

### VBS 2017 Chairs

Klaus Schöffmann  
Werner Bailer  
Cathal Gurrin  
Jakub Lokoč

Klagenfurt University, Austria  
Joanneum Research, Austria  
Dublin City University, Ireland  
Charles University in Prague, Czech Republic

### Sponsorship Chairs

Yantao Zhang  
Tao Mei

Snapchat  
Microsoft Research Asia

### Proceedings Chair

Gylfi Þ. Guðmundsson

Reykjavik University, Iceland

### Local Chair

Marta K. Lárusdóttir

Reykjavik University, Iceland

### Local Support

Reykjavik University Event Services, CP Reykjavik

## Steering Committee

Phoebe Chen (Chair)  
Tat-Seng Chua

La Trobe University, Australia  
National University of Singapore, Singapore

Kiyoharu Aizawa	University of Tokyo, Japan
Cathal Gurrin	Dublin City University, Ireland
Benoit Huet	EURECOM, France
R. Manmatha	University of Massachusetts, USA
Noel E. O'Connor	Dublin City University, Ireland
Klaus Schöffmann	Klagenfurt University, Austria
Yang Shiqiang	Tsinghua University, China
Cees G.M. Snoek	University of Amsterdam, The Netherlands
Meng Wang	Hefei University of Technology, China

## Special Session Organizers

### SS1: Social Media Retrieval and Recommendation

Liqliang Nie	National University of Singapore, Singapore
Yan Yan	University of Trento, Italy
Benoit Huet	EURECOM, France

### SS2: Modeling Multimedia Behaviors

Peng Wang	Tsinghua University, China
Frank Hopfgartner	University of Glasgow, UK
Liang Bai	National University of Defense Technology, China

### SS3: Multimedia Computing for Intelligent Life

Zhineng Chen	Chinese Academy of Sciences, China
Wei Zhang	Chinese Academy of Sciences, China
Ting Yao	Microsoft Research Asia, China
Kai-Lung Hua	National Taiwan University of Science and Technology, Taiwan, R.O.C.
Wen-Huang Cheng	Academia Sinica, Taiwan, R.O.C.

### SS4: Multimedia and Multimodal Interaction for Health and Basic Care Applications

Stefanos Vrochidis	ITI-CERTH, Greece
Leo Wanner	Pompeu Fabra University, Spain
Elisabeth André	University of Augsburg, Germany
Klaus Schöffmann	Klagenfurt University, Austria

## Program Committee

Esra Acar	Technische Universität Berlin, Germany
Amin Ahmadi	DCU/Insight Centre for Data Analytics, Ireland
Le An	UNC Chapel Hill, USA
Ognjen Arandjelović	University of St. Andrews, UK
Anant Bajjal	SAMSUNG Electronics, South Korea



Werner Bailer	Joanneum Research, Austria
Ilaria Bartolini	University of Bologna, Italy
Jenny Benois-Pineau	University of Bordeaux/LABRI, France
Milan Bjelica	University of Belgrade, Serbia
Laszlo Böszörményi	Klagenfurt University, Austria
Benjamin Bustos	University of Chile, Chile
K. Selçuk Candan	Arizona State University, USA
Shiyu Chang	UIUC, USA
Savvas Chatzichristofis	Democritus University of Thrace, Greece
Edgar Chávez	CICESE, Mexico
Wen-Huang Cheng	Academia Sinica, Taiwan, R.O.C.
Gene Cheung	National Institute of Informatics, Japan
Wei-Ta Chu	National Chung Cheng University, Taiwan, R.O.C.
Vincent Claveau	IRISA-CNRS, France
Kathy M. Clawson	University of Sunderland, UK
Claudiu Cobarzan	Babes-Bolyai University, Romania
Michel Crucianu	Cnam, France
Peng Cui	Tsinghua University, China
Rossana Damiano	Università di Torino, Italy
Petros Daras	Centre for Research and Technology Hellas, Greece
Wesley De Neve	Ghent University, Belgium
François Destelle	DCU Insight, Ireland
Cem Direkolu	Middle East Technical University, Northern Cyprus Campus, Turkey
Lingyu Duan	Peking University, China
Jianping Fan	UNC Charlotte, USA
Mylène Fariès	University of Brasília, Brazil
Gerald Friedland	ICSI/UC Berkeley, USA
Weijie Fu	Hefei University of Technology, China
Lianli Gao	University of Electronic Science and Technology, China
Yue Gao	Tsinghua University, China
Guillaume Gravier	CNRS, IRISA, and Inria Rennes, France
Ziyu Guan	Northwest University of China, China
Gylfi Þór Guðmundsson	Reykjavik University, Iceland
Silvio Guimarães	PUC Minas, Brazil
Allan Hanbury	TU Wien, Austria
Shijie Hao	Hefei University of Technology, China
Alex Hauptmann	Carnegie Mellon University, USA
Andreas Henrich	University of Bamberg, Germany
Nicolas Hervé	Institut National de l'Audiovisuel, France
Richang Hong	Hefei University of Technology, China
Frank Hopfgartner	University of Glasgow, UK
Michael Houle	National Institute of Informatics, Japan
Jun-Wei Hsieh	National Taiwan Ocean University, Taiwan, R.O.C.
Zhenzhen Hu	Nanyang Technological University, Singapore

Kai-Lung Hua	National Taiwan University of Science and Technology, Taiwan, R.O.C.
Jen-Wei Huang	National Cheng Kung University, Taiwan, R.O.C.
Benoit Huet	EURECOM, France
Wolfgang Hürst	Utrecht University, The Netherlands
Ichiro Ide	Nagoya University, Japan
Adam Jatowt	Kyoto University, Japan
Rongrong Ji	Xiamen University, China
Peiguang Jing	Tianjin University, China
Håvard Johansen	University of Tromsø, Norway
Mohan Kankanhalli	National University of Singapore, Singapore
Jiro Katto	Waseda University, Japan
Yoshihiko Kawai	NHK, Japan
Yiannis Kompatsiaris	CERTH-ITI, Greece
Harald Kosch	University of Passau, Germany
Markus Koskela	University of Helsinki, Finland
Duy-Dinh Le	National Institute of Informatics, Japan
Michael Lew	Leiden University, The Netherlands
Haojie Li	Dalian University of Technology, China
Teng Li	Anhui University, China
Yingbo Li	Tripnester, France
Rainer Lienhart	University of Augsburg, Germany
Suzanne Little	Dublin City University, Ireland
Bo Liu	Rutgers, USA
Xueliang Liu	HFUT, China
Zhenguang Liu	National University of Singapore, Singapore
Guojun Lu	Federation University Australia, Australia
Changzhi Luo	Hefei University of Technology, China
Stéphane Marchand-Maillet	University of Geneva, Switzerland
Jean Martinet	University of Lille, France
José M. Martínez	Universidad Autónoma de Madrid, Spain
Kevin McGuinness	Dublin City University, Ireland
Robert Mertens	HSW University of Applied Sciences, Germany
Vasileios Mezaris	CERTH, Greece
Rui Min	Google, USA
Dalibor Mitrović	mediamid digital services GmbH, Austria
Henning Muller	HES-SO, Switzerland
Phivos Mylonas	Ionian University, Greece
Chong-Wah Ngo	City University of Hong Kong, SAR China
Liqiang Nie	Shandong University, China
Naoko Nitta	Osaka University, Japan
Noel O'Connor	Dublin City University, Ireland
Neil O'Hare	Yahoo, USA
Vincent Oria	New Jersey Institute of Technology, USA
Tse-Yu Pan	National Cheng Kung University, Taiwan, R.O.C.
Fernando Pereira	Instituto Superior Técnico, Portugal

Yannick Prié	University of Nantes, France
Jianjun Qian	Nanjing University of Science and Technology, China
Xueming Qian	Xi'an Jiaotong University, China
Georges Quénot	LIG-CNRS, France
Miloš Radovanović	University of Novi Sad, Serbia
Michael Riegler	Simula Research Lab, Norway
Stevan Rudinac	University of Amsterdam, The Netherlands
Mukesh Kumar Saini	Indian Institute of Technology Ropar, India
Jitao Sang	Chinese Academy of Sciences, China
Klaus Schöffmann	Klagenfurt University, Austria
Pascale Sébillot	IRISA/INSA Rennes, France
Jie Shao	University of Electronic Science and Technology, China
Xiaobo Shen	Nanjing University of Science and Technology, China
Koichi Shinoda	Tokyo Institute of Technology, Japan
Mei-Ling Shyu	University of Miami, USA
Alan Smeaton	Dublin City University, Ireland
Lifeng Sun	Tsinghua University, China
Yongqing Sun	NTT Media Intelligence Laboratories, Japan
Sheng Tang	Institute of Computing Technology, Chinese Academy of Sciences, China
Shuhei Tarashima	NTT, Japan
Wei-Guang Teng	National Cheng Kung University, Taiwan, R.O.C.
Georg Thallinger	Joanneum Research, Austria
Qi Tian	University of Texas at San Antonio, USA
Christian Timmerer	Alpen-Adria-Universität Klagenfurt, Austria
Dian Tjondronegoro	Queensland University of Technology, Australia
Shingo Uchihashi	Fuji Xerox Co., Ltd., Japan
Feng Wang	East China Normal University, China
Jinqiao Wang	Chinese Academy of Sciences, China
Shikui Wei	Beijing Jiaotong University, China
Lai Kuan Wong	Multimedia University, Malaysia
Marcel Worring	University of Amsterdam, The Netherlands
Hong Wu	University of Electronic Science and Technology of China, China
Xiao Wu	Southwest Jiaotong University, China
Changsheng Xu	Chinese Academy of Sciences, China
Toshihiko Yamasaki	The University of Tokyo, Japan
Bo Yan	Fudan University, China
Keiji Yanai	University of Electro-Communications, Tokyo, Japan
Kuiyuan Yang	Microsoft Research, China
You Yang	HUST, China
Jun Yu	Hangzhou Dianzi University, China
Maia Zaharieva	Vienna University of Technology, Austria
Matthias Zeppelzauer	University of Applied Sciences St. Poelten, Austria
Zheng-Jun Zha	University of Science and Technology of China, China

Cha Zhang	Microsoft Research, USA
Hanwang Zhang	National University of Singapore, Singapore
Tianzhu Zhang	CASIA, Bangladesh
Cairong Zhao	Tongji University, China
Ye Zhao	Hefei University of Technology, China
Lijuan Zhou	Dublin City University, Ireland
Shiai Zhu	University of Ottawa, Canada
Xiaofeng Zhu	Guangxi Normal University, China
Arthur Zimek	University of Southern Denmark, Denmark
Roger Zimmermann	National University of Singapore, Singapore

### **Demonstration, Special Session and VBS Reviewers**

Shanshan Ai	Beijing Jiaotong University, China
Alberto Messina	RAI CRIT, Italy
François Brémont	Inria, France
Houssem Chatbri	Dublin City University, Ireland
Jingyuan Chen	National University of Singapore, Singapore
Yi Chen	Helsinki Institute for Information Technology, Finland
Yiqiang Chen	Chinese Academy of Sciences, China
Zhineng Chen	Chinese Academy of Sciences, China
Zhiyong Cheng	National University of Singapore, Singapore
Mariana Damova	Mozaika, Romania
Stamatia Dasiopoulou	Pompeu Fabra University, Spain
Monika Dominguez	Pompeu Fabra University, Spain
Ling Du	Tianjin Polytechnic University, China
Jana Eggink	BBC R&D, UK
Bailan Feng	Chinese Academy of Sciences, China
Fuli Feng	Chinese Academy of Sciences, China
Min-Chun Hu	National Cheng Kung University, Taiwan, R.O.C.
Lei Huang	Ocean University of China, China
Marco A. Hudelist	Klagenfurt University, Austria
Bogdan Ionescu	Politehnica University of Bucharest, Romania
Eleni Kamateri	CERTH, Greece
Hyowon Lee	Singapore University of Technology and Design, Singapore
Andreas Leibetseder	Alpen-Adria-Universität Klagenfurt, Austria
Na Li	Dublin City University, Ireland
Xirong Li	Renmin University of China, China
Wu Liu	Beijing University of Posts and Telecommunications, China
Jakub Lokoč	Charles University in Prague, Czech Republic
Mathias Lux	Klagenfurt University, Austria
Georgios Meditskos	CERTH, Greece
Wolfgang Minker	Ulm University, Germany
Bernd Münzer	Klagenfurt University, Austria

Adrian Muscat	University of Malta, Malta
Yingwei Pan	University of Science and Technology of China, China
Zhengyuan Pang	Tsinghua University, China
Stefan Petscharnig	Alpen-Adria Universität Klagenfurt, Austria
Manfred Jürgen Primus	Alpen-Adria-Universität Klagenfurt, Austria
Zhaofan Qiu	University of Science and Technology of China, China
Amon Rapp	University of Toronto, Canada
Fuming Sun	Liaoning University of Technology, China
Xiang Wang	National University of Singapore, Singapore
Hongtao Xie	Chinese Academy of Sciences, China
Yuxiang Xie	National University of Defense Technology, China
Shiqiang Yang	Tsinghua University, China
Yang Yang	University of Electronic Science and Technology of China, China
Changqing Zhang	Tianjin University, China

### **External Reviewers**

Duc Tien Dang Nguyen	Dublin City University, Ireland
Yusuke Matsui	National Institute of Informatics, Japan
Sang Phan	National Institute of Informatics, Japan
Jiang Zhou	Dublin City University, Ireland

## Contents – Part II

### Full Papers Accepted for Poster Presentation

A Comparative Study for Known Item Visual Search Using Position Color Feature Signatures . . . . .	3
<i>Jakub Lokoč, David Kuboň, and Adam Blažek</i>	
A Novel Affective Visualization System for Videos Based on Acoustic and Visual Features. . . . .	15
<i>Jianwei Niu, Yiming Su, Shasha Mo, and Zeyu Zhu</i>	
A Novel Two-Step Integer-pixel Motion Estimation Algorithm for HEVC Encoding on a GPU. . . . .	28
<i>Keji Chen, Jun Sun, Zongming Guo, and Dachuan Zhao</i>	
A Scalable Video Conferencing System Using Cached Facial Expressions . . .	37
<i>Fang-Yu Shih, Ching-Ling Fan, Pin-Chun Wang, and Cheng-Hsin Hsu</i>	
A Unified Framework for Monocular Video-Based Facial Motion Tracking and Expression Recognition . . . . .	50
<i>Jun Yu</i>	
A Virtual Reality Framework for Multimodal Imagery for Vessels in Polar Regions . . . . .	63
<i>Scott Sorensen, Abhishek Kolagunda, Andrew R. Mahoney, Daniel P. Zitterbart, and Chandra Kambhamettu</i>	
Adaptive and Optimal Combination of Local Features for Image Retrieval . . .	76
<i>Neelanjan Bhowmik, Valérie Gouet-Brunet, Lijun Wei, and Gabriel Bloch</i>	
An Evaluation of Video Browsing on Tablets with the ThumbBrowser . . . . .	89
<i>Marco A. Hudelist and Klaus Schoeffmann</i>	
Binaural Sound Source Distance Reproduction Based on Distance Variation Function and Artificial Reverberation . . . . .	101
<i>Jiawang Xu, Xiaochen Wang, Maosheng Zhang, Cheng Yang, and Ge Gao</i>	
Color-Introduced Frame-to-Model Registration for 3D Reconstruction . . . . .	112
<i>Fei Li, Yunfan Du, and Rujie Liu</i>	
Compressing Visual Descriptors of Image Sequences. . . . .	124
<i>Werner Bailer, Stefanie Wechtitsch, and Marcus Thaler</i>	

Deep Convolutional Neural Network for Bidirectional Image-Sentence Mapping . . . . .	136
<i>Tianyuan Yu, Liang Bai, Jinlin Guo, Zheng Yang, and Yuxiang Xie</i>	
Discovering Geographic Regions in the City Using Social Multimedia and Open Data . . . . .	148
<i>Stevan Rudinac, Jan Zahálka, and Marcel Worring</i>	
Discovering User Interests from Social Images . . . . .	160
<i>Jiangchao Yao, Ya Zhang, Ivor Tsang, and Jun Sun</i>	
Effect of Junk Images on Inter-concept Distance Measurement: Positive or Negative?. . . . .	173
<i>Yusuke Nagasawa, Kazuaki Nakamura, Naoko Nitta, and Noboru Babaguchi</i>	
Exploiting Multimodality in Video Hyperlinking to Improve Target Diversity . . . . .	185
<i>Rémi Bois, Vedran Vukotić, Anca-Roxana Simon, Ronan Sicre, Christian Raymond, Pascale Sébillot, and Guillaume Gravier</i>	
Exploring Large Movie Collections: Comparing Visual Berrypicking and Traditional Browsing. . . . .	198
<i>Thomas Low, Christian Hentschel, Sebastian Stober, Harald Sack, and Andreas Nürnbergger</i>	
Facial Expression Recognition by Fusing Gabor and Local Binary Pattern Features . . . . .	209
<i>Yuechuan Sun and Jun Yu</i>	
Frame-Independent and Parallel Method for 3D Audio Real-Time Rendering on Mobile Devices. . . . .	221
<i>Yucheng Song, Xiaochen Wang, Cheng Yang, Ge Gao, Wei Chen, and Weiping Tu</i>	
Illumination-Preserving Embroidery Simulation for Non-photorealistic Rendering. . . . .	233
<i>Qiqi Shen, Dele Cui, Yun Sheng, and Guixu Zhang</i>	
Improving the Discriminative Power of Bag of Visual Words Model. . . . .	245
<i>Achref Ouni, Thierry Urruty, and Muriel Visani</i>	
M-SBIR: An Improved Sketch-Based Image Retrieval Method Using Visual Word Mapping . . . . .	257
<i>Jianwei Niu, Jun Ma, Jie Lu, Xuefeng Liu, and Zeyu Zhu</i>	
Movie Recommendation via BLSTM. . . . .	269
<i>Song Tang, Zhiyong Wu, and Kang Chen</i>	

Multimodal Video-to-Video Linking: Turning to the Crowd for Insight and Evaluation . . . . .	280
<i>Maria Eskevich, Martha Larson, Robin Aly, Serwah Sabetghadam, Gareth J.F. Jones, Roeland Ordelman, and Benoit Huet</i>	
Online User Modeling for Interactive Streaming Image Classification . . . . .	293
<i>Jiagao Hu, Zhengxing Sun, Bo Li, Kewei Yang, and Dongyang Li</i>	
Recognizing Emotions Based on Human Actions in Videos . . . . .	306
<i>Guolong Wang, Zheng Qin, and Kaiping Xu</i>	
Rocchio-Based Relevance Feedback in Video Event Retrieval . . . . .	318
<i>G.L.J. Pingen, M.H.T. de Boer, and R.B.N. Aly</i>	
Scale-Relation Feature for Moving Cast Shadow Detection. . . . .	331
<i>Chih-Wei Lin</i>	
Smart Loudspeaker Arrays for Self-Coordination and User Tracking . . . . .	343
<i>Jungju Jee and Jung-Woo Choi</i>	
Spatial Verification via Compact Words for Mobile Instance Search . . . . .	356
<i>Bo Wang, Jie Shao, Chengkun He, Gang Hu, and Xing Xu</i>	
Stochastic Decorrelation Constraint Regularized Auto-Encoder for Visual Recognition. . . . .	368
<i>Fengling Mao, Wei Xiong, Bo Du, and Lefei Zhang</i>	
The Perceptual Lossless Quantization of Spatial Parameter for 3D Audio Signals . . . . .	381
<i>Gang Li, Xiaochen Wang, Li Gao, Ruimin Hu, and Dengshi Li</i>	
Unsupervised Multiple Object Cosegmentation via Ensemble MIML Learning . . . . .	393
<i>Weichen Yang, Zhengxing Sun, Bo Li, Jiagao Hu, and Kewei Yang</i>	
Using Object Detection, NLP, and Knowledge Bases to Understand the Message of Images . . . . .	405
<i>Lydia Weiland, Ioana Hulpus, Simone Paolo Ponzetto, and Laura Dietz</i>	
Video Search via Ranking Network with Very Few Query Exemplars . . . . .	419
<i>De Cheng, Lu Jiang, Yihong Gong, Nanning Zheng, and Alexander G. Hauptmann</i>	
 <b>Demonstrations</b>	
A Demo for Image-Based Personality Test . . . . .	433
<i>Huaiwen Zhang, Jiaming Zhang, Jitao Sang, and Changsheng Xu</i>	



A Web-Based Service for Disturbing Image Detection . . . . .	438
<i>Markos Zampoglou, Symeon Papadopoulos, Yiannis Kompatsiaris, and Jochen Spangenberg</i>	
An Annotation System for Egocentric Image Media . . . . .	442
<i>Aaron Duane, Jiang Zhou, Suzanne Little, Cathal Gurrin, and Alan F. Smeaton</i>	
DeepStyleCam: A Real-Time Style Transfer App on iOS . . . . .	446
<i>Ryosuke Tanno, Shin Matsuo, Wataru Shimoda, and Keiji Yanai</i>	
V-Head: Face Detection and Alignment for Facial Augmented Reality Applications . . . . .	450
<i>Zhiwei Wang and Xin Yang</i>	
<b>Video Browser Showdown</b>	
Collaborative Feature Maps for Interactive Video Search . . . . .	457
<i>Klaus Schoeffmann, Manfred Jürgen Primus, Bernd Muenzer, Stefan Petscharnig, Christof Karisch, Qing Xu, and Wolfgang Huerst</i>	
Concept-Based Interactive Search System. . . . .	463
<i>Yi-Jie Lu, Phuong Anh Nguyen, Hao Zhang, and Chong-Wah Ngo</i>	
Enhanced Retrieval and Browsing in the IMOTION System. . . . .	469
<i>Luca Rossetto, Ivan Giangreco, Claudiu Tănase, Heiko Schuldt, Stéphane Dupont, and Omar Seddati</i>	
Semantic Extraction and Object Proposal for Video Search . . . . .	475
<i>Vinh-Tiep Nguyen, Thanh Duc Ngo, Duy-Dinh Le, Minh-Triet Tran, Duc Anh Duong, and Shin'ichi Satoh</i>	
Storyboard-Based Video Browsing Using Color and Concept Indices . . . . .	480
<i>Wolfgang Hürst, Algernon Ip Vai Ching, Klaus Schoeffmann, and Manfred J. Primus</i>	
VERGE in VBS 2017 . . . . .	486
<i>Anastasia MOUNTZIDOU, Theodoros Mironidis, Fotini Markatopoulou, Stelios Andreadis, Ilias Gialampoukidis, Damianos Galanopoulos, Anastasia Ioannidou, Stefanos Vrochidis, Vasileios Mezaris, Ioannis Kompatsiaris, and Ioannis Patras</i>	
Video Hunter at VBS 2017 . . . . .	493
<i>Adam Blažek, Jakub Lokoč, and David Kuboň</i>	
<b>Author Index</b> . . . . .	499

# Contents – Part I

## Full Papers Accepted for Oral Presentation

3D Sound Field Reproduction at Non Central Point for NHK 22.2 System. . . . .	3
<i>Song Wang, Ruimin Hu, Shihong Chen, Xiaochen Wang, Yuhong Yang, Weiping Tu, and Bo Peng</i>	
A Comparison of Approaches for Automated Text Extraction from Scholarly Figures . . . . .	15
<i>Falk Bösch and Ansgar Scherp</i>	
A Convolutional Neural Network Approach for Post-Processing in HEVC Intra Coding . . . . .	28
<i>Yuanying Dai, Dong Liu, and Feng Wu</i>	
A Framework of Privacy-Preserving Image Recognition for Image-Based Information Services . . . . .	40
<i>Kojiro Fujii, Kazuaki Nakamura, Naoko Nitta, and Noboru Babaguchi</i>	
A Real-Time 3D Visual Singing Synthesis: From Appearance to Internal Articulators . . . . .	53
<i>Jun Yu</i>	
A Structural Coupled-Layer Tracking Method Based on Correlation Filters . . . . .	65
<i>Sheng Chen, Bin Liu, and Chang Wen Chen</i>	
Augmented Telemedicine Platform for Real-Time Remote Medical Consultation . . . . .	77
<i>David Anton, Gregorij Kurillo, Allen Y. Yang, and Ruzena Bajcsy</i>	
Color Consistency for Photo Collections Without Gamut Problems . . . . .	90
<i>Qi-Chong Tian and Laurent D. Cohen</i>	
Comparison of Fine-Tuning and Extension Strategies for Deep Convolutional Neural Networks . . . . .	102
<i>Nikiforos Pittaras, Foteini Markatopoulou, Vasileios Mezaris, and Ioannis Patras</i>	
Describing Geographical Characteristics with Social Images . . . . .	115
<i>Huangjie Zheng, Jiangchao Yao, and Ya Zhang</i>	
Fine-Grained Image Recognition from Click-Through Logs Using Deep Siamese Network . . . . .	127
<i>Wu Feng and Dong Liu</i>	

Fully Convolutional Network with Superpixel Parsing for Fashion Web Image Segmentation . . . . . 139  
*Lixuan Yang, Helena Rodriguez, Michel Crucianu, and Marin Ferecatu*

Graph-Based Multimodal Music Mood Classification in Discriminative Latent Space. . . . . 152  
*Feng Su and Hao Xue*

Joint Face Detection and Initialization for Face Alignment . . . . . 164  
*Zhiwei Wang and Xin Yang*

Large-Scale Product Classification via Spatial Attention Based CNN Learning and Multi-class Regression . . . . . 176  
*Shanshan Ai, Caiyan Jia, and Zhineng Chen*

Learning Features Robust to Image Variations with Siamese Networks for Facial Expression Recognition . . . . . 189  
*Wissam J. Baddar, Dae Hoe Kim, and Yong Man Ro*

M3LH: Multi-modal Multi-label Hashing for Large Scale Data Search. . . . . 201  
*Guan-Qun Yang, Xin-Shun Xu, Shanqing Guo, and Xiao-Lin Wang*

Model-Based 3D Scene Reconstruction Using a Moving RGB-D Camera . . . . . 214  
*Shyi-Chyi Cheng, Jui-Yuan Su, Jing-Min Chen, and Jun-Wei Hsieh*

Modeling User Performance for Moving Target Selection with a Delayed Mouse . . . . . 226  
*Mark Claypool, Ragnhild Eg, and Kjetil Raaen*

Multi-attribute Based Fire Detection in Diverse Surveillance Videos . . . . . 238  
*Shuangqun Li, Wu Liu, Huadong Ma, and Huiyuan Fu*

Near-Duplicate Video Retrieval by Aggregating Intermediate CNN Layers. . . . . 251  
*Giorgos Kordopatis-Zilos, Symeon Papadopoulos, Ioannis Patras, and Yiannis Kompatsiaris*

No-Reference Image Quality Assessment Based on Internal Generative Mechanism. . . . . 264  
*Xinchun Qian, Wengang Zhou, and Houqiang Li*

On the Exploration of Convolutional Fusion Networks for Visual Recognition . . . . . 277  
*Yu Liu, Yanming Guo, and Michael S. Lew*

Phase Fourier Reconstruction for Anomaly Detection on Metal Surface Using Salient Irregularity . . . . . 290  
*Tzu-Yi Hung, Sriram Vaikundam, Vidhya Natarajan, and Liang-Tien Chia*

ReMagicMirror: Action Learning Using Human Reenactment with the Mirror Metaphor . . . . . 303  
*Fabian Lorenzo Dayrit, Ryosuke Kimura, Yuta Nakashima, Ambrosio Blanco, Hiroshi Kawasaki, Katsushi Ikeuchi, Tomokazu Sato, and Naokazu Yokoya*

Robust Image Classification via Low-Rank Double Dictionary Learning . . . . . 316  
*Yi Rong, Shengwu Xiong, and Yongsheng Gao*

Robust Scene Text Detection for Multi-script Languages Using Deep Learning . . . . . 329  
*Ruo-Ze Liu, Xin Sun, Hailiang Xu, Palaiahnakote Shivakumara, Feng Su, Tong Lu, and Ruoyu Yang*

Robust Visual Tracking Based on Multi-channel Compressive Features . . . . . 341  
*Jianqiang Xu and Yao Lu*

Single Image Super-Resolution with a Parameter Economic Residual-Like Convolutional Neural Network . . . . . 353  
*Ze Yang, Kai Zhang, Yudong Liang, and Jinjun Wang*

Spatio-Temporal VLAD Encoding for Human Action Recognition in Videos . . . . . 365  
*Ionut C. Duta, Bogdan Ionescu, Kiyoharu Aizawa, and Nicu Sebe*

Structure-Aware Image Resizing for Chinese Characters . . . . . 379  
*Chengdong Liu, Zhouhui Lian, Yingmin Tang, and Jianguo Xiao*

Supervised Class Graph Preserving Hashing for Image Retrieval and Classification . . . . . 391  
*Lu Feng, Xin-Shun Xu, Shanqing Guo, and Xiao-Lin Wang*

Visual Robotic Object Grasping Through Combining RGB-D Data and 3D Meshes. . . . . 404  
*Yiyang Zhou, Wenhai Wang, Wenjie Guan, Yirui Wu, Heng Lai, Tong Lu, and Min Cai*

What Convnets Make for Image Captioning? . . . . . 416  
*Yu Liu, Yanming Guo, and Michael S. Lew*

What are Good Design Gestures? –Towards User- and Machine-friendly Interface– . . . . . 429  
*Ryo Kawahata, Atsushi Shimada, and Rin-ichiro Taniguchi*

**SS1: Social Media Retrieval and Recommendation**

Collaborative Dictionary Learning and Soft Assignment for Sparse Coding of Image Features . . . . .	443
<i>Jie Liu, Sheng Tang, and Yu Li</i>	
LingoSent — A Platform for Linguistic Aware Sentiment Analysis for Social Media Messages . . . . .	452
<i>Yuting Su and Huijing Wang</i>	
Multi-Task Multi-modal Semantic Hashing for Web Image Retrieval with Limited Supervision . . . . .	465
<i>Liang Xie, Lei Zhu, and Zhiyong Cheng</i>	
Object-Based Aggregation of Deep Features for Image Retrieval. . . . .	478
<i>Yu Bao and Haojie Li</i>	
Uyghur Language Text Detection in Complex Background Images Using Enhanced MSERs . . . . .	490
<i>Shun Liu, Hongtao Xie, Chuan Zhou, and Zhendong Mao</i>	

**SS2: Modeling Multimedia Behaviors**

CELoF: WiFi Dwell Time Estimation in Free Environment . . . . .	503
<i>Chen Yan, Peng Wang, Haitian Pang, Lifeng Sun, and Shiqiang Yang</i>	
Demographic Attribute Inference from Social Multimedia Behaviors: A Cross-OSN Approach. . . . .	515
<i>Liancheng Xiang, Jitao Sang, and Changsheng Xu</i>	
Understanding Performance of Edge Prefetching . . . . .	527
<i>Zhengyuan Pang, Lifeng Sun, Zhi Wang, Yuan Xie, and Shiqiang Yang</i>	
User Identification by Observing Interactions with GUIs . . . . .	540
<i>Zaher Hinbarji, Rami Albatal, and Cathal Gurrin</i>	
Utilizing Locality-Sensitive Hash Learning for Cross-Media Retrieval . . . . .	550
<i>Jia Yuhua, Bai Liang, Wang Peng, Guo Jinlin, Xie Yuxiang, and Yu Tianyuan</i>	

**SS3: Multimedia Computing for Intelligent Life**

A Sensor-Based Official Basketball Referee Signals Recognition System Using Deep Belief Networks . . . . .	565
<i>Chung-Wei Yeh, Tse-Yu Pan, and Min-Chun Hu</i>	

Compact CNN Based Video Representation for Efficient Video Copy  
 Detection . . . . . 576  
*Ling Wang, Yu Bao, Haojie Li, Xin Fan, and Zhongxuan Luo*

Cross-Modal Recipe Retrieval: How to Cook this Dish? . . . . . 588  
*Jingjing Chen, Lei Pang, and Chong-Wah Ngo*

Deep Learning Based Intelligent Basketball Arena with Energy Image. . . . . 601  
*Wu Liu, Jiangyu Liu, Xiaoyan Gu, Kun Liu, Xiaowei Dai,  
 and Huadong Ma*

Efficient Multi-scale Plane Extraction Based RGBD Video Segmentation. . . . . 614  
*Hong Liu, Jun Wang, Xiangdong Wang, and Yueliang Qian*

Human Pose Tracking Using Online Latent Structured Support Vector  
 Machine. . . . . 626  
*Kai-Lung Hua, Irawati Nurmala Sari, and Mei-Chen Yeh*

Micro-Expression Recognition by Aggregating Local Spatio-Temporal  
 Patterns . . . . . 638  
*Shiyu Zhang, Bailan Feng, Zhineng Chen, and Xiangsheng Huang*

egoPortray: Visual Exploration of Mobile Communication Signature from  
 Egocentric Network Perspective . . . . . 649  
*Qing Wang, Jiansu Pu, Yuanfang Guo, Zheng Hu, and Hui Tian*

i-Stylist: Finding the Right Dress Through Your Social Networks . . . . . 662  
*Jordi Sanchez-Riera, Jun-Ming Lin, Kai-Lung Hua, Wen-Huang Cheng,  
 and Arvin Wen Tsui*

**SS4: Multimedia and Multimodal Interaction for Health and  
 Basic Care Applications**

Boredom Recognition Based on Users’ Spontaneous Behaviors  
 in Multiparty Human-Robot Interactions . . . . . 677  
*Yasuhiro Shibasaki, Kotaro Funakoshi, and Koichi Shinoda*

Classification of sMRI for AD Diagnosis with Convolutional Neuronal  
 Networks: A Pilot 2-D+ $\epsilon$  Study on ADNI . . . . . 690  
*Karim Aderghal, Manuel Boissenin, Jenny Benois-Pineau,  
 Gwenaëlle Catheline, and Karim Afdel*

Deep Learning for Shot Classification in Gynecologic Surgery Videos. . . . . 702  
*Stefan Petscharnig and Klaus Schöffmann*

Description Logics and Rules for Multimodal Situational Awareness  
in Healthcare . . . . . 714  
*Georgios Meditskos, Stefanos Vrochidis, and Ioannis Kompatsiaris*

Speech Synchronized Tongue Animation by Combining Physiology  
Modeling and X-ray Image Fitting . . . . . 726  
*Jun Yu*

Erratum to: ReMagicMirror: Action Learning  
Using Human Reenactment with the Mirror Metaphor . . . . . E1  
*Fabian Lorenzo Dayrit, Ryosuke Kimura, Yuta Nakashima,  
Ambrosio Blanco, Hiroshi Kawasaki, Katsushi Ikeuchi, Tomokazu Sato,  
and Naokazu Yokoya*

**Author Index** . . . . . 739