

Volume 4, Number 1
March 2012

Published by the Association for Computing Machinery
Special Interest Group on Multimedia

SIGMM RECORDS

Table of Contents

1	Volume 4, Number 1
1	Editorial
1	ACM Multimedia 2011
16	Award opportunities
16	SIGMM Award for Outstanding PhD Thesis
18	SIGMM Award for Outstanding Technical Contributions
19	ACM TOMCCAP Nicolas D. Georganas Best Paper Award
19	The ACM Multimedia Grand Challenge 2011 in a Nutshell
20	Open Source Column: Tribler: P2P Search, Share and Stream
24	MM2012: The Multimedia Grand Challenge goes into its fourth year
25	MPEG Column: 99th MPEG Meeting
26	PhD thesis abstracts
26	Anita Sobe
27	Dominik Kaspar
28	Sree Hari Krishnan Parthasarathi
29	Mohammad Kazemi Varnamkhasti
30	Mona Omidyeganeh
31	Xirong Li
32	Event and publication reports
33	Calls for contributions
35	Free Material
35	ENST-Drums (Videos + Audios)
35	QUASI database for the evaluation of music source separation
35	Job Opportunities
35	Lecturer (Assistant Professor)
35	Back matter



Association for
Computing Machinery

Advancing Computing as a Science & Profession

ISSN 1947-4598
<http://sigmm.org/records>

SIGMM Records

Volume 4, Number 1, March 2012

Editorial

Dear Member of the SIGMM Community,

Welcome to the first issue of the SIGMM Records in 2012.

We provide you with a report from the ACM Multimedia 2011, including links to the digital library for all papers, of course. As part of the conference, the Grand Challenge was arranged. The report from the 2011 Grand Challenge can be found in this issue, as well as the Grand Challenges that have been posed by industry for the next round. You should read the article and attack the 2012 Grand Challenges in time for ACM Multimedia 2012.

It is time to nominate your candidates for three of SIGMM's prestigious awards: the Nicolas Georganas TOMCCAP Award, the SIGMM Award for Outstanding Technical Contributions to Multimedia Computing, Communications and Applications, and the SIGMM Award for Outstanding PhD Thesis in Multimedia Computing, Communications and Applications. You find the calls in this issue.

You can read PhD thesis summaries provided by a total six candidates who have recently passed their doctoral exams. Their varied topics span most of the interest range of SIGMM, you should be certain to find work that interests you.

Christian Timmerer reports from the 99th MPEG meeting, which concluded some important work and is preparing for new challenges. Free material is offered to you in this issue: you can read an article about Tribler, an open source P2P system. You can also find short descriptions of and references to two datasets for the multimedia community: ENST-Drums and QUASI.

Last but not most certainly not least, you find pointers to the latest issues of TOMCCAP and MMSJ, and several announcements from ACM and SIGMM.

We hope that you enjoy this issue of the Records.

The Editors
Stephan Kopf
Viktor Wendel
Lei Zhang
Pradeep Atrey

Christian Timmerer
Pablo Cesar
Carsten Griwodz

ACM Multimedia 2011

Authors: K. Selçuk Candan, Sethuraman Panchanathan, Balakrishnan Prabhakaran, Hari Sundaram, Wu-Chi Feng, Nicu Sebe, Borko Furht, Jin Li, Maria Luisa Sapino
Venue: Scottsdale, Arizona, USA
URL: <http://www.acmmm11.org>
by K. Selçuk Candan et al.

General Chairs

K. Selçuk Candan, Arizona State University, AZ, USA
Sethuraman Panchanathan, Arizona State University, AZ, USA

Balakrishnan Prabhakaran, University of Texas at Dallas, TX, USA

Technical Program Chair

Hari Sundaram, Arizona State University, AZ, USA
Wu-Chi Feng, Portland State University, OR, USA
Nicu Sebe, University of Trento, IT

Workshop Chairs

Borko Furht, Florida Atlantic University, FL, USA
Jin Li, Microsoft Research
Maria Luisa Sapino, University of Torino, IT

Introduction to ACM Multimedia 2011

We are delighted to report on behalf of the entire organizing committee that the 19th ACM International Conference on Multimedia ACM Multimedia 2011 (MM'11) was held between November 28th and December 1st, 2011, in Scottsdale, Arizona, USA, to great success.

ACM Multimedia (MM) is the flagship conference of the Special Interest Group on Multimedia (SIGMM), which profiles cutting-edge scientific developments and showcases innovative industrial multimedia technologies and applications. The conference aims to promote intellectual exchanges and interactions among scientists, engineers, students, multimedia users, and artists through various events, including keynote talks from leaders in the area, oral and poster sessions focused on research challenges and solutions, workshops in up-and-coming key areas of research, technical and industrial demonstrations of

prototypes and commercial products, tutorials, research and industrial panels, doctoral symposium, mentoring events, scientific competitions (including an open source software and a multimedia grand challenge competition), and interactive art exhibits.

Our key motivation while organizing the MM'11 conference was to find innovative ways to design an "inclusive" conference program: lowering the barriers between various MM sub-communities, boosting the cross-fertilization of ideas among the contributors and attendees across the various MM events, and maximizing the return-on-investment for the MM'11 participants. Examples of this new approach include the following:

- New plenary poster sessions, where all contributors (i.e., authors of long and short research papers, of workshop papers, and contributors to all other MM'11 events) are invited to share poster versions of their contributions with the rest of the MM'11 community. These plenary poster sessions allow conference participants to get a quick idea of interesting things happening in the multiple parallel sessions they cannot clone themselves to attend!
- Workshops were aligned with the other MM'11 events, instead of being held on a separate "workshops day" where many workshop participants never got to know the main conference and vice-versa. Our aim was to integrate workshops (which were chosen, in the first place, to represent emerging topics that complement the areas covered by the main technical program) organically with the other conference events and encourage broader participation by registrants in all conference and workshop programs.
- These innovations aimed to eliminate barriers in the program had to be supported by corresponding innovations in the MM'11 registration policies. Thus, we have instituted an "all-in-one" registration fee structure, which covers attendance to all MM'11 events, including presentation and poster sessions, panels, demonstrations, tutorials, and workshops.
- By keeping the overall registration fee lower than recent years and by shaving one day off from the conference program, we also reduced the overall participation cost for most of the MM'11 attendees.

A travel grants program, generously supported by the National Science Foundation (NSF) and SIGMM, also helped us lower the barrier for participation for many students and a full 30% of the registrants to MM'11 were students. A number of mentorship activities was organized at MM'11, including women mentoring event, organized and sponsored by SIGMM, a Doctoral Symposium program, which (in addition to having regular panels and presentations as before) opened up its doors to all student authors who wanted to present

posters at the event, a panel on "Job Opportunities and Career Perspective for Fresh Graduates of the Multimedia Community", and a new "vis-a-vis meeting with researchers" social event where graduate students could meet and exchange ideas and receive guidance with internationally recognized researchers in their research area.

Of course, apart from the above, MM'11 also continued with programs that proved to be extremely successful in the past. We have continued with the well-established and highly successful Open Source Software Competition, with special emphasis this year on instructional open source software designed for educational use in teaching multimedia-related courses at undergraduate or graduate level. Like the previous years, the Multimedia Grand Challenges competition attracted challenges from many leaders of the multimedia industry, including HP, Technicolor, Nokia, Yahoo, Huawei, and 3D Life, and proposals from all over the world. A report form the Grand Challenges can be found in a separate article. Similarly, this year's industrial exhibits program, which complemented the MM technical demonstrations program, focussed on cutting-edge research prototypes, including system and product demonstrations from many industrial leaders, such as IBM, FX Palo Alto Labs, Microsoft Research, Exalead, and Yacast. The panels program emphasized opportunities and challenges faced by researchers, industry, and open-source communities in multimedia and thus covers timely topics, such as "Smart Games", "Towards Synergy Between the Open Source and the Research Multimedia Communities", and "Innovating the Multimedia Experience".

We are enthused to report that MM'11 included three exciting keynote talks by three industry and academic leaders in multimedia research: Alex Pentland, Head of the MIT Human Dynamics Lab, Genevieve Bell, Director of the Interaction and Experience Research at the Intel Labs, and Arnaud Robert, Senior Vice President of Technology at The Walt Disney Studios. We are proud that MM'11 hosts the prestigious SIGMM Technical Achievement Award presentation to Prof Shih-Fu Chang (Columbia University) and his award acceptance speech.

The Award for the Best Paper of MM'11 was presented by the technical program chairs to F. Yu, R. Ji and S. Chang for their paper "Active Query Sensing for Mobile Location Search". After a tough competition, the best student paper award of 2011 was shared by two papers. It was granted to W. Wu, A. Arefin, G. Kurillo, B. Agarwal, K. Nahrstedt and R. Bajcsy for their paper "Color-plus-Depth Level-of-Details in 3D Teleimmersive Video - A Psychophysical Approach" and to R. Garg, A. Varma, M. Wu for their paper "Seeing ENF: Natural Time

Stamp for Digital Video via Optical Sensing and Signal Processing". Also the best technical demo was chosen after long deliberation, and the award was presented to David S Monaghan, James O'Sullivan, and Noel O'Connor for their demo "Low-cost Creation of a 3D Interactive Museum Exhibition".

We would like to acknowledge all who have contributed to the success of MM'11. First of all, we would like to thank all authors who submitted papers to the technical program, various workshops, and other events of MM'11. We also thank the authors of the accepted papers who will present their work in MM'11 and the panelists and keynote speakers who have accepted to participate in the conference to discuss current and future challenges in the field of multimedia and to propose innovative solutions. We are grateful to the members of the various program committees and external reviewers who have helped put together a high-quality program and would like acknowledge members of the various MM'11 organizing committees and many student volunteers for their invaluable help at every step of the process. We would like to thank the staff of ACM and Sheridan for their continuous support and the Conference Management Toolkit Team (CMT) at Microsoft for letting us use CMT for handling the submission and review workflows of MM'11. Finally, we would like to thank our sponsors (as of this writing), Google, IBM, Microsoft, FxPal, Technicolor, Qualcomm, Springer, Yahoo!, Arizona State University, and the University of Texas at Dallas, who have extended their generous support to MM'11. We would also like to thank the National Science Foundation (NSF) and SIGMM for their generous support for the MM'11 student travel award program.

The conference program

We had an exciting technical program at ACM Multimedia 2011. The process to select the technical program included several innovations. These innovations - guided by the recent report by a select SIG Multimedia committee - included the following: MM'11 moved to technical areas instead of tracks, each submitted paper had a primary area and an optional secondary area set by the authors at the time of submission, and an author rebuttal phase.

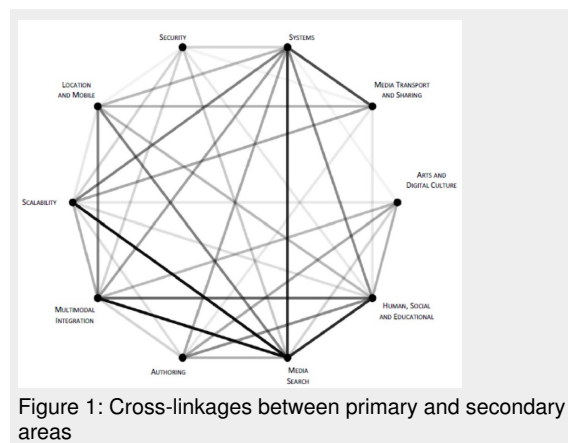


Figure 1: Cross-linkages between primary and secondary areas

The ten areas were as follows:

- Multi-modal Integration and Understanding in the Imperfect World
- Media Analysis and Search
- Scalability in Media Processing, Analysis, and Applications
- Multimedia Systems and Middleware
- Media Transport and Sharing
- Multimedia Security
- Media Authoring and Production
- Location-based and Mobile Multimedia
- Human, Social, and Educational Aspects of Multimedia
- Arts and Contemporary Digital Culture

The areas - chosen in consultation with the broad SIG Multimedia community - reflect core research areas (e.g. multimedia systems and middleware), as well as different multimedia research contexts (e.g. location-based mobile multimedia). Typically, each area had two area chairs managing the review process; media analysis and search was an exception: due to the large number of submissions (39%), we assigned six area chairs to manage this area. Figure 1 shows the relationship between areas in the papers submitted for review - there is an edge between two areas when a paper has both areas specified in the paper. In the figure, edges with higher strength are more opaque.

Many authors specified both primary and secondary areas, resulting in the assignment of two reviewers from the primary area and one reviewer from the secondary area. In the minority of cases when there was no secondary area, we assigned all three reviewers from the primary area. The selection of reviewers from two different areas, allowed for a cross-disciplinary evaluation of the submitted paper. In past conferences a paper submitted to one of the four tracks was exclusively evaluated by reviewers from that track. Additionally, to

increase the responsibility of the chairs as well as to give more credit to their work on their accepted paper, we have indicated the name of the AC that was supervising the reviewing process and which recommended the paper for acceptance.

The conference was highly competitive with low acceptance rates. We received a total of 666 submissions which included 335 long papers and 331 short papers. Subsequent to the initial notification, authors had one week to rebut the criticisms of the papers. After the rebuttal phase, the area chairs led a discussion with the reviewers on the merits of each paper, which included the author rebuttals. In the end, we accepted 58 long papers, with an acceptance rate of 17.3%, and 120 short papers with an acceptance rate of 36.3%. We additionally recommended 52 long papers to appear as short papers. Figure 2 shows the distribution of primary areas for all papers. We selected three papers were selected for the best paper and best student paper session. We solicited nominations from each area, for best student paper and best paper competition. The technical program chairs selected three from the nominations.

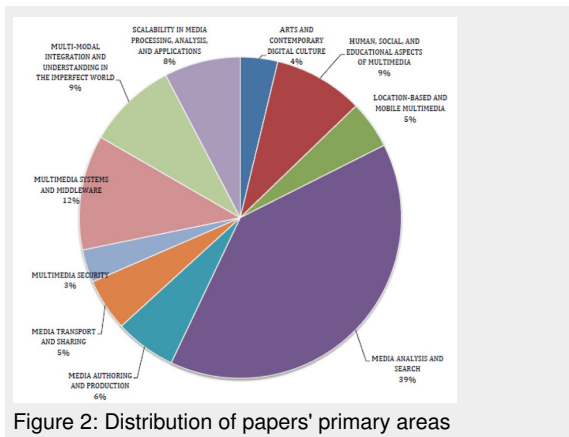


Figure 2: Distribution of papers' primary areas

We would like to thank all of our area chairs and reviewers who volunteered a significant amount of their time to ensure a high quality program.

The workshop

This year, ACM Multimedia experimented with a new workshop format that co-located the workshops with ACM Multimedia and ran in parallel with the conference sessions. The new format influenced the selection process: among the 28 very strong proposals we received, we could select only 11 workshops. We conducted the selection process in accordance with the SIGMM group guidelines, and admitted 11 strong workshops to the program, whose themes are distinct from (and complement) the areas of the

main conference. The gave us the following very rich workshops program, spanning from relatively established topics at ACM MM (this is the case for those workshops which are at their 3rd edition) to topics that are very new in the ACM MM community:

- Workshop on Music Information Retrieval with User-Centered and Multimodal Strategies (MIRUM'11)
- Workshop on Multimedia in Forensics and Intelligence (MiFor'11)
- Workshop on Automated Media Analysis and Production for Novel TV Services (AIEMPro 2011)
- Workshop on Social Media (WSM11)
- Workshop on Social and Behavioral Networked Media Access (SBNMA'11)
- Workshop on Multimedia Technologies for Distance Learning (MTDL'11)
- Workshop on Interactive Multimedia on Mobile and Portable Devices (IMMPD'11)
- Joint Workshop on Modeling and Representing Events (J-MRE'11)
 - Part 1: Workshop on Events in Multimedia (EiMM11)
 - Part 2: Workshop on Sparse Representation for Event Detection in Multimedia (SRED'11)
- Joint Workshop on Human Gesture and Behavior Understanding (J-HGBU'11)
 - Part 1: Workshop on Social Signal Processing (SSPW'11)
 - Part 2: Workshop on Multimedia access to 3D Human Objects (MA3HO'11)
- Workshop on Medical Multimedia Analysis and Retrieval (MMAR)
- Workshop on Ubiquitous Meta User Interfaces (Ubi-MUI'11)

We would like to thank all the organizers who submitted their workshop proposals, and in particular the organizers of the workshops that appeared in the program. We are aware that new workshop format caused a significant amount of synchronization work from the organizers, who were required to align their internal deadlines and schedule with the ones of the main conference. We really appreciated their collaboration, and we reached a very interesting and successful workshop program. We hope that all ACM MM attendees enjoyed our program, and that the new format will increase the appeal of the workshops, and significantly boost intellectual exchange.

Papers

SESSION: Technical achievement award and best paper candidates

- Shih-Fu Chang:
Content based multimedia retrieval: lessons learned from two decades of research
- Felix X. Yu, Rongrong Ji, Shih-Fu Chang:
Active query sensing for mobile location search
- Wanmin Wu, Ahsan Arefin, Gregorij Kurillo, Pooja Agarwal, Klara Nahrstedt, Ruzena Bajcsy:
Color-plus-depth level-of-detail in 3D tele-immersive video: a psychophysical approach
- Ravi Garg, Avinash L. Varna, Min Wu:
"Seeing" ENF: natural time stamp for digital video via optical sensing and signal processing
- Jinfeng Zhuang, Tao Mei, Steven C.H. Hoi, Xian-Sheng Hua, Shipeng Li:
Modeling social strength in social media community via kernel-based learning
- Wei Jiang, Alexander C. Loui:
Audio-visual grouplet: temporal audio-visual interactions for general video concept classification
- Zechao Li, Meng Wang, Jing Liu, Changsheng Xu, Hanqing Lu:
News contextualization with geographic and visual information
- Zhenyong Fu, Horace H.S. Ip, Hongtao Lu, Zhiwu Lu:
Multi-modal constraint propagation for heterogeneous image clustering

SESSION: Human, social, and educational aspects of multimedia

- Ramanathan Subramanian, Victoria Yanulevskaya, Nicu Sebe:
Can computers learn from humans to see better?: inferring scene semantics from viewers' eye movements
- Axel Carlier, Guntur Ravindra, Vincent Charvillat, Wei Tsang Ooi:
Combining content-based analysis and crowdsourcing to improve user interaction with zoomable video
- Lexing Xie, Apostol Natsev, John R. Kender, Matthew Hill, John R. Smith:
Visual memes in social media: tracking real-world news in YouTube videos
- Bo Geng, Linjun Yang, Chao Xu, Xian-Sheng Hua, Shipeng Li:
The role of attractiveness in web image search

SESSION: Location-based and mobile multimedia

- Yang Wang, Tao Mei, Jingdong Wang, Houqiang Li, Shipeng Li:
JIGSAW: interactive mobile visual search with multimodal queries
- An-Jung Cheng, Yan-Ying Chen, Yen-Ta Huang, Winston H. Hsu, Hong-Yuan Mark Liao:
Personalized travel recommendation by mining people attributes from community-contributed photos
- Zhijie Shen, Sakire Arslan Ay, Seon Ho Kim, Roger Zimmermann:
Automatic tag generation and ranking for sensor-rich outdoor videos
- Shu Shi, Cheng-Hsin Hsu, Klara Nahrstedt, Roy Campbell:
Using graphics rendering contexts to enhance the real-time video coding for mobile cloud gaming

SESSION: Multi-modal integration and understanding in the imperfect world

SESSION: Arts and contemporary digital culture

- Wendy Ann Mansilla, Andrew Perkis, Touradj Ebrahimi:
Implicit experiences as a determinant of perceptual quality and aesthetic appreciation
- Edgar Roman-Rangel, Carlos Pallan Gayol, Jean-Marc Odobez, Daniel Gatica-Perez:
Searching the past: an improved shape descriptor to retrieve maya hieroglyphs
- Christian Jacquemin, Georges Gagneré, Benoît Lahoz:
Shedding light on shadow: real-time interactive artworks based on cast shadows or silhouettes

SESSION: Media transport and sharing

- Minhui Zhu, Sebastien Mondet, Géraldine Morin, Wei Tsang Ooi, Wei Cheng:
Towards peer-assisted rendering in networked virtual environments
- Yao Liu, Fei Li, Lei Guo, Yang Guo, Songqing Chen:
BlueStreaming: towards power-efficient internet P2P streaming to mobile devices
- Ke Liang, Roger Zimmermann, Wei Tsang Ooi:
Peer-assisted texture streaming in metaverses
- Yan Huang, Zhenhua Li, Gang Liu, Yafei Dai:
Cloud download: using cloud utilities to achieve high-quality content distribution for unpopular videos

SESSION: Media authoring and production 1

- Alberto Piacenza, Fabrizio Guerrini, Nicola Adami, Riccardo Leonardi, Julie Porteous, Jonathan Teutenberg, Marc Cavazza:
Generating story variants with constrained video recombination
- Xirong Li, Efstratios Gavves, Cees G.M. Snoek, Marcel Worring, Arnold W.M. Smeulders:
Personalizing automated image annotation using cross-entropy

- Song Tan, Chong-Wah Ngo, Hung-Khoon Tan, Lei Pang:
Cross media hyperlinking for search topic browsing
- Chun-Wei Liu, Tz-Huan Huang, Ming-Hsu Chang, Ken-Yi Lee, Chia-Kai Liang, Yung-Yu Chuang:
3D cinematography principles and their applications to stereoscopic media processing

SESSION: Media analysis and search 1

- Xiangyu Chen, Xiaotong Yuan, Shuicheng Yan, Jinhui Tang, Yong Rui, Tat-Seng Chua:
Towards multi-semantic image annotation with graph regularized exclusive group lasso
- Martha Larson, Christoph Kofler, Alan Hanjalic:
Reading between the tags to predict real-world size-class for visually depicted objects in images
- Zhigang Ma, Yi Yang, Feiping Nie, Jasper Uijlings, Nicu Sebe:
Exploiting the entire feature space with sparsity for automatic image annotation
- Ju-Chiang Wang, Yu-Chin Shih, Meng-Sung Wu, Hsin-Min Wang, Shyh-Kang Jeng:
Colorizing tags in tag cloud: a novel query-by-tag music search system

SESSION: Media authoring and production 2

- Rodrigo Laiola Guimarães, Pablo Cesar, Dick C.A. Bulterman, Vilmos Zsombori, Ian Kegel:
Creating personalized memories from social events: community-based support for multi-camera recordings of school concerts
- Frank Nack, Ichiro Ide:
Why did the prime minister resign?: generation of event explanations from large news repositories
- Christophe Lino, Marc Christie, Roberto Ranon, William Bares:
The director's lens: an intelligent assistant for virtual cinematography
- Vivek K. Singh, Jiebo Luo, Dhiraj Joshi, Phoury Lei, Madirakshi Das, Peter Stubler:
Reliving on demand: a total viewer experience

SESSION: Media analysis and search 2

- Sheng-hua Zhong, Yan Liu, Yang Liu:
Bilinear deep learning for image classification
- Dayong Wang, Steven C.H. Hoi, Ying He, Jianke Zhu:
Retrieval-based face annotation by weak label regularized local coordinate coding
- Xinmei Tian, Yijuan Lu, Linjun Yang, Qi Tian:
Learning to judge image search results
- Olivier Le Meur, Thierry Baccino, Aline Roumy:

Prediction of the inter-observer visual congruency (IOVC) and application to image ranking

SESSION: Multimedia systems and middleware 1

- Dao T.P. Quynh, Ying He, Xiaoming Chen, Jiazhi Xia, Qian Sun, Steven C.H. Hoi:
Modeling 3D articulated motions with conformal geometry videos (CGVs)
- Qianqian Xu, Tingting Jiang, Yuan Yao, Qingming Huang, Bowei Yan, Weisi Lin:
Random partial paired comparison for subjective video quality assessment via hodgerank
- Wei Song, Dian Tjondronegoro, Michael Docherty:
Saving bitrate vs. pleasing users: where is the break-even point in mobile video quality?
- Peijia Zheng, Jiwu Huang:
Implementation of the discrete wavelet transform and multiresolution analysis in the encrypted domain

SESSION: Media analysis and search 3

- Jingkuan Song, Yi Yang, Zi Huang, Heng Tao Shen, Richang Hong:
Multiple feature hashing for real-time large scale near-duplicate video retrieval
- Xianming Liu, Hongxun Yao, Rongrong Ji, Pengfei Xu, Xiaoshuai Sun, Qi Tian:
Learning heterogeneous data for hierarchical web video classification
- Xiao-Yong Wei, Zhen-Qun Yang:
Coached active learning for interactive video search
- Jin Yuan, Zheng-Jun Zha, Yao-Tao Zheng, Meng Wang, Xiangdong Zhou, Tat-Seng Chua:
Learning concept bundles for video search with complex queries

SESSION: Multimedia systems and middleware 2

- Pengpeng Ni, Ragnhild Eg, Alexander Eichhorn, Carsten Griwodz, Pål Halvorsen:
Flicker effects in adaptive video streaming to handheld devices
- Yao Liu, Lei Guo, Fei Li, Songqing Chen:
An empirical evaluation of battery power consumption for streaming data transmission to mobile devices
- Jui-Hsin Lai, Chieh-Li Chen, Po-Chen Wu, Chieh-Chi Kao, Shao-Yi Chien:
Tennis real play: an interactive tennis game with models from real videos
- Xiangwen Chen, Minghua Chen, Baochun Li, Yao Zhao, Yunnan Wu, Jin Li:
Celerity: a low-delay multi-party conferencing solution

SESSION: Media analysis and search 4

- Wenbin Tang, Rui Cai, Zhiwei Li, Lei Zhang:
Contextual synonym dictionary for visual object retrieval
- Wenhao Lu, Jingdong Wang, Xian-Sheng Hua, Shengjin Wang, Shipeng Li:
Contextual image search
- Zhang Liu, Chaokun Wang, Yiyuan Bai, Hao Wang, Jianmin Wang:
MUSIZ: a generic framework for music resizing with stretching and cropping
- Nobuyuki Morioka, Jingdong Wang:
Robust visual reranking via sparsity and ranking constraints
- Pablo Cesar, Wei Tsang Ooi, Ben Moskowit, Zohar Babin, Dick Bulterman, Rainer Lienhart, Robert Richter:
Towards synergy between the open source and the research multimedia communities
- Yu-Ru Lin, Vincent Oria, K. Selcuk Candan, Lyndon Kennedy, Dulce Dulce Ponceleon, Hari Sundaram, Rong Yan, Roger Zimmerman:
Job opportunities and career perspective for fresh graduates of the multimedia community
- Khaled El-Maleh, Haohong Wang, Susie Wee, Heather Yu, James D. Johnston, Zhengyou Zhang:
Innovating the multimedia experience

SESSION: Applications

- Troy McDaniel, Morris Goldberg, Daniel Villanueva, Lakshmi Narayan Viswanathan, Sethuraman Panchanathan:
Motor learning using a kinematic-vibrotactile mapping targeting fundamental movements
- Xiaohong Xiang, Mohan S. Kankanhalli:
Affect-based adaptive presentation of home videos
- Naoko Nitta, Noboru Babaguchi:
Example-based video remixing support system
- Rongrong Ji, Ling-Yu Duan, Jie Chen, Hongxun Yao, Yong Rui, Shih-Fu Chang, Wen Gao:
Towards low bit rate mobile visual search with multiple-channel coding

SESSION: Plenary talk sessions

- Alex Pentland:
Honest signals: how social networks shape human behavior
- Genevieve Bell:
"U are happy life": telling the future's stories
- Arnaud Robert:
Digital media distribution: the future

SESSION: Events

- Klara Nahrstedt, Svetha Venkatesh, Nalini Venkatasubramanian, Dulce Ponceleon, Maria Zemankova, Susanne Boll:
Networking of multimedia women event beyond epsilon science: where to look and how to realize new opportunities
- David Tinapple, Todd Ingalls:
ACM multimedia interactive art program: interaction stations

PANEL SESSION: Panels

- Abdulmotaleb El Saddik:

Serious games

- Cynthia C.S. Liem, Meinard Müller, Douglas Eck, George Tzanetakis:
1st international ACM workshop on music information retrieval with user-centered and multimodal strategies (MIRUM)
- Sebastiano Battiato, Sabu Emmanuel, Adrian Ulges, Marcel Worring:
Third ACM international workshop on multimedia in forensics and intelligence (MiFor 2011)
- Sid-Ahmed Berrani, Alberto Messina:
AIEMPro 2011: the 4th international workshop on automated media analysis and production for novel TV services
- Steven Chu-Hong Hoi, Michal Jacovi, Ioannis Kompatsiaris, Jiebo Luo, Konstantinos Tserpes:
WSM2011: third ACM workshop on social media
- Naeem Ramzan, Fei Wang, Charalampos Z. Patrikakis, Peng Cui, Nikolaos Doulamis, Shiqiang Yang, Gordon Sun:
ACM international workshop on social and behavioral networked media access (SBNMA'11)
- Vasileios Mezaris, Ansgar Scherp, Ramesh Jain, Mohan Kankanhalli, Huiyu Zhou, Jianguo Zhang, Liang Wang, Zhengyou Zhang:
Modeling and representing events in multimedia
- Maja Pantic, Alex Pentland, Alessandro Vinciarelli, Rita Cucchiara, Mohamed Daoudi, Alberto Del Bimbo:
Joint ACM workshop on human gesture and behavior understanding: (J-HGBU'11)
- Yu Cao, Jayashree Kalpathy-Cramer, Devrim Ünay:
Medical multimedia analysis and retrieval
- Ali Asghar Nazari Shirehjini, Sahin Albayrak, Abdulsalam Yassine:
Ubi-MUI 2011 ACM workshop summary
- Rynson W.H. Lau, Timothy K. Shih, Frederick W.B. Li, Neil Y. Yen:
The third ACM international workshop on multimedia technologies for distance learning (MTDL 2011)

WORKSHOP SESSION: Workshop overviews

- Jiebo Luo, Caifeng Shan, Ling Shao, Minoru Etoh:
ACM international workshop on interactive multimedia on mobile and portable devices (IMMPD'11)

TUTORIAL SESSION: Tutorial overviews

- Alan Hanjalic, Martha Larson:
Frontiers in multimedia search
- Tao Mei, Ruofei Zhang, Xian-Sheng Hua:
Internet multimedia advertising: techniques and technologies
- Cees G.M. Snoek, Arnold W.M. Smeulders:
Internet video search
- Simone Santini:
Semantic computing in multimedia
- Gaël Richard:
Tutorial on multimedia music signal processing
- Gerald Friedland:
Acoustic and multimodal processing for multimedia content analysis
- Xiao-Ping Zhang, Zhu Liu:
Graphical probabilistic modeling and applications in multimedia content analysis
- Jialie Shen, Meng Wang, Shuicheng Yan, Xian-Sheng Hua:
Multimedia tagging: past, present and future
- Harish Katti, Mohan Kankanhalli:
Eye-tracking methodology and applications to images and video

SESSION: Grand challenge session

- Christoph Kofler, Martha Larson, Alan Hanjalic:
Alice's worlds of wonder: exploiting tags to understand images in terms of size and scale
- Guan-Long Wu, Yu-Chuan Su, Tzu-Hsuan Chiu, Liang-Chi Hsieh, Winston H. Hsu:
Scalable mobile video question-answering system with locally aggregated descriptors and random projection
- Yu-Heng Lei, Yan-Ying Chen, Lime Iida, Bor-Chun Chen, Hsiao-Hang Su, Winston H. Hsu:
Photo search by face positions and facial attributes on touch devices
- Chun Chet Tan, Yu-Gang Jiang, Chong-Wah Ngo:
Towards textually describing complex video contents with audio-visual concept classifiers
- Dimitrios S. Alexiadis, Philip Kelly, Petros Daras, Noel E. O'Connor, Tamy Boubekeur, Maher Ben Moussa:
Evaluating a dancer's performance using kinect-based skeleton tracking
- Tsung-Hung Tsai, Wen-Huang Cheng, Yung-Huan Hsieh:

- *Dynamic social network for narrative video analysis*
- Marc Gowing, Philip Kell, Noel E. O'Connor, Cyril Concolato, Slim Essid, Jean Lefevre, Robin Tournemene, Ebroul Izquierdo, Vlado Kitanovski, Xinyu Lin, Qianni Zhang:
Enhanced visualisation of dance performance from automatically synchronised multimodal recordings
- Bruno do Nascimento Teixeira, Jùlia Epischina Engràcia de Oliveira, Fillipe Dias Moreira de Souza, Tiago Oliveira Cunha, Arnaldo de Albuquerque Araújo, Christiane Okamoto, Lucas Figueiredo, Vinicius de Oliveira Silva, Igor Galil Loures de Oliveira:
News browsing system: multimodal analysis
- Slim Essid, Yves Grenier, Mounira Maazaoui, Gaël Richard, Robin Tournemene:
An audio-driven virtual dance-teaching assistant
- Yoshitaka Ushiku, Tatsuya Harada, Yasuo Kuniyoshi:
Understanding images with natural sentences
- Wanmin Wu, Ahsan Arefin, Gregorij Kurillo, Pooja Agarwal, Klara Nahrstedt, Ruzena Bajcsy:
A psychophysical approach for real-time 3D video processing
- Yin-Hsi Kuo, Wen-Yu Lee, Winston H. Hsu, Wen-Huang Cheng:
Augmenting mobile city-view image retrieval with context-rich user-contributed photos

SESSION: Open source software competition

- Jonathon S. Hare, Sina Samangoeei, David P. Dupplaw:
OpenIMAJ and ImageTerrier: Java libraries and tools for scalable multimedia analysis and indexing of images
- Isaac Esteban, Judith Dijk, Frans C.A. Groen:
From images to 3d models made easy
- Fabien Cazenave, Vincent Quint, Cécile Roisin:
Timesheets.js: tools for web multimedia
- Christopher A. Brooks, Markus Ketterl, Adam Hochman, Josh Holtzman, Judy Stern, Tobias Wunden, Kristofor Amundson, Greg Logan, Kenneth Lui, Adam McKenzie, Denis Meyer, Markus Moormann, Matjaz Rihtar, Ruediger Rolf, Nejc Skofic, Micah Sutton, Ruben Perez Vazquez, Benjamin Wulff:
OpenCast Matterhorn 1.1: reaching new heights
- Sung Hee Park, Andrew Adams, Eino-Ville Talvala:
The FCam API for programmable cameras
- Jérôme Gorin, Hervé Yviquel, Françoise Prêteux, Mickaël Raulet:
Just-in-time adaptive decoder engine: a universal video decoder based on MPEG RVC
- Jean Le Feuvre, Cyril Concolato, Jean-Claude Dufourd, Romain Bouqueau, Jean-Claude Moissinac:

- Experimenting with multimedia advances using GPAC*
- Sherif Halawa, Derek Pang, Ngai-Man Cheung, Bernd Girod:
ClassX: an open source interactive lecture StreamingSystem
 - Christopher Müller, Christian Timmerer:
A VLC media player plugin enabling dynamic adaptive streaming over HTTP
 - Andrés Barrios, Matias Barrios, Daniel De Vera, Pablo Rodríguez-Bocca, Claudia Rostagnol:
GoalBit: a free and open source peer-to-peer streaming network
 - Werner Bailer, Hermann Fürntratt, Peter Schallauer, Georg Thallinger, Werner Haas:
A C++ library for handling MPEG-7 descriptions
 - Mathias Lux:
Content based image retrieval with LIRe
 - Niels Zeilemaker, Mihai Capot#, Arno Bakker, Johan Pouwelse:
Tribler: P2P media search and sharing
 - Jean Bresson, Carlos Agon, Gérard Assayag:
OpenMusic: visual programming environment for music composition, analysis and research
- SESSION: Industrial exhibits 1**
- Julien Law-To, Gregory Grefenstette:
VOVALEAD: a scalable video search engine based on content
 - Raphaël Blouet, Charlotte Juan:
MMSI talk: an applicative use case of quaero media monitoring & social impact
 - Jingdong Wang, Xian-Sheng Hua:
Web-scale image search by color sketch
 - Arthur Lenoir, Rémi Landais:
MuMa: a scalable music search engine based on content analysis
 - Christian Wengert, Tobias Jaeggli, Philippe Messmer, Till Quack, Peter Cech, Cristi Prodan, Tomas Carnecky, Franco Sebregondi, David Wisti:
Kooba interactive posters
- DEMONSTRATION SESSION: Technical demos 1**
- Boqing Gong, Jianzhuang Liu, Xiaogang Wang, Xiaou Tang:
3D object retrieval with semantic attributes
 - Zhou Ren, Jingjing Meng, Junsong Yuan, Zhengyou Zhang:
Robust hand gesture recognition with kinect sensor
 - Zhijie Shen, Sakire Arslan Ay, Seon Ho Kim:
SRV-TaGS: An Automatic TAGging and Search System for Sensor-Rich Outdoor Videos
 - Kazuhiro Otsuka, Kamil Sebastian Mucha, Shiro Kumano, Dan Mikami, Masafumi Matsuda, Junji Yamato:
A system for reconstructing multiparty conversation field based on augmented head motion by dynamic projection
 - Andreas Zingerle, Tyler Freeman:
enabling the VJ as performer with rhythmic wearable interfaces
 - David Sadlier, Paul Ferguson, Dian Zhang, Noel E. O'Connor, Hyowon Lee:
InSPeCT: integrated surveillance for port container traffic
 - Svetha Venkatesh, Stewart Greenhill, Dinh Phung, Brett Adams:
Cognitive intervention in autism using multimedia stimulus
 - Takayuki Yamada, Seiichi Gohshi, Isao Echizen:
iCabinet: stand-alone implementation of a method for preventing illegal recording of displayed content by adding invisible noise signals
 - Jian Dong, Yuzhao Ni, Jiashi Feng, Shuicheng Yan:
Purposive hidden-object game (P-HOG) towards imperceptible human computation
 - Genliang Guan, Zhiyong Wang, Xian-Sheng Hua, Dagan Feng:
StoryImaging: a media-rich presentation system for textual stories
 - Ning Zhang, Tao Mei, Xian-Sheng Hua, Ling Guan, Shipeng Li:
TapTell: understanding visual intents on-the-go
 - Britta Meixner, Johannes Köstler, Harald Kosch:
A mobile player for interactive non-linear video
 - Carmelo Velardo, Jean-Luc Dugelay:
Real time extraction of body soft biometric from 3D videos
 - Ahsan Arefin, Zixia Huang, Raoul Rivas, Shu Shi, Wanmin Wu, Klara Nahrstedt:
Tele-immersive gaming for everybody
 - Diana Siwiak, Nicole Lehrer, Michael Baran, Yinpeng Chen, Margaret Duff, Todd Ingalls, Thanassis Rikakis:
A home-based adaptive mixed reality rehabilitation system
 - Derek Pang, Sherif Halawa, Ngai-Man Cheung, Bernd Girod:
ClassX Mobile: region-of-interest video streaming to mobile devices with multi-touch interaction
 - Changhu Wang, Jun Zhang, Bruce Yang, Lei Zhang:
Sketch2Cartoon: composing cartoon images by sketching
 - Beomjoo Seo, Jia Hao, Guanfeng Wang:
Sensor-rich video exploration on a map interface
 - Andrew Au, Jie Liang:

Ztitch: a mobile phone application for 3D scene creation, navigation, and sharing

- Emi Myodo, Satoshi Ueno, Koichi Takagi, Shigeyuki Sakazawa:
Automatic comic-like image layout system preserving image order and important regions

SESSION: Industrial exhibits 2

- Eric Bouillet, Luca Gasparini, Olivier Verscheure:
Towards a real time public transport awareness system: case study in dublin
- Gene Golovchinsky, Scott Carter, Anthony Dunnigan:
ARA: the active reading application
- Jacob T. Biehl, Thea Turner, William van Melle, Andreas Girgensohn:
myUnity: a new platform to support communication in the modern workplace

SESSION: Technical demos 2

- Lei Pang, Song Tan, Hung Khoon Tan, Chong Wah Ngo:
Galaxy browser: exploratory search of web videos
- Felix X. Yu, Rongrong Ji, Tongtao Zhang, Shih-Fu Chang:
A mobile location search system with active query sensing
- Florian Mehm, Sandro Hardy, Stefan Göbel, Ralf Steinmetz:
Collaborative authoring of serious games for health
- Haojie Li, Lei Yi, Jinhui Tang, Xiaohui Wang:
Capturing a great photo via learning from community-contributed photo collections
- Alberto Piacenza, Fabrizio Guerrini, Nicola Adami, Riccardo Leonardi, Jonathan Teutenberg, Julie Porteous, Marc Cavazza:
Changing video arrangement for constructing alternative stories
- Hervé Goëau, Alexis Joly, Souheil Selmi, Pierre Bonnet, Elise Mouysset, Laurent Joyeux, Jean-François Molino, Philippe Birnbaum, Daniel Bathelemy, Nozha Boujemaa:
Visual-based plant species identification from crowdsourced data
- Vivek K. Singh, Jiebo Luo, Dhiraj Joshi, Madirakshi Das, Phoury Lei, Peter Stubler:
Dynamic media show drivable by semantics
- Steven C.H. Hoi, Pengcheng Wu:
SIRE: a social image retrieval engine
- Paul B. Beskow, Håkon K. Stensland, Håvard Espeland, Espen A. Kristiansen, Preben N. Olsen, Ståle Kristoffersen, Carsten Griwodz, Pål Halvorsen:
Processing of multimedia data using the P2G framework

- Qia Wang, Alex Lobzhanidze, Suman Deb Roy, Wenjun Zeng, Yi Shang:
Positionit: an image-based remote target localization system on smartphones

- David Monaghan, James O'Sullivan, Noel E. O'Connor, Bridget Kelly, Olivier Kazmierczak, Lorraine Comer:
Low-cost creation of a 3D interactive museum exhibition

- Koichi Mori, Rafael Ballagas, Glenda Revelle, Hayes Raffle, Hiroshi Horii, Mirjana Spasojevic:
Interactive rich reading: enhanced book reading experience with a conversational agent

- Klaus Schoeffmann, Manfred del Fabro:
Hierarchical video browsing with a 3D carousel
- Axel Carlier, Arash Shafiei, Julien Badie, Salim Bensiali, Wei Tsang Ooi:
COZI: crowdsourced and content-based zoomable video player

- Christophe Lino, Marc Christie, Roberto Ranon, William Bares:
A smart assistant for shooting virtual cinematography with motion-tracked cameras

- Gerald Friedland, Jaeyoung Choi, Adam Janin:
Video2GPS: a demo of multimodal location estimation on flickr videos

- Alexis Fesnin, Valerie Gouet-Brunet, Scott Kominen, Vincent Oria, Jichao Sun:
Towards a privacy preserving personal photo album manager with semantic classification, indexing and querying capabilities

- Yang Cai, Linjun Yang, Wei Ping, Fei Wang, Tao Mei, Xian-Sheng Hua, Shipeng Li:
Million-scale near-duplicate video retrieval system

- Junfeng He, Tai-Hsu Lin, Jinyuan Feng, Shih-Fu Chang:
Mobile product search with bag of hash bits

SESSION: Oral presentation session

- Stefan Romberg:
From local features to local regions
- Zixia Huang:
Synchronized dissemination framework for supporting high-quality tele-immersive shared activity
- Ahsan Arefin:
Session management of correlated multi-stream 3D tele-immersive environments
- Ivan Anselmo Sipiran Mendoza:
Local features for partial shape matching and retrieval

POSTER SESSION: Posters Session

- Xiaodong Yang:

- Recognizing clothes patterns for blind people by confidence margin based feature combination*
- Shu Shi:
Building low-latency remote rendering systems for interactive 3D graphics rendering on mobile devices
 - Felix X. Yu:
Intelligent query formulation for mobile visual search
 - Xia Li:
Robust image representation for efficient recognition and retrieval
 - Neela Sawant:
Modeling tagged photos for automatic image annotation
 - Raoul Rivas:
Acropolis: operating system for real-time multi-streaming systems
 - Wanmin Wu, Klara Nahrstedt:
Human-centric control of video functions and underlying resources in 3D tele-immersive systems
 - Michele Merler:
Analysis, indexing and visualization of presentation videos
 - Yao Liu:
Towards efficient resource utilization in internet mobile streaming
 - Vivek K. Singh:
From multimedia data to situation detection
 - Troy McDaniel, Sethuraman Panchanathan:
The building blocks of somatic information delivery
 - Richa Tiwari, Chengcui Zhang:
Video genre detection using a multimodality approach
 - Wei-Bang Chen, Chengcui Zhang:
Multiple object retrieval in image databases using hierarchical segmentation tree
 - Shenghua Zhong, Yan Liu, Yang Liu:
Bilinear deep learning for image classification
 - Song Tan:
Beyond video searching: summarization and exploration
 - Xirong Li:
Image search 2.0
- SESSION: Short paper session 1*
- Guifang Duan, Neela Sawant, James Z. Wang, Dean Snow, Danni Ai, Yen-Wei Chen:
Analysis of cyriot icon faces using ICA-enhanced active shape model representation
 - Diogo Cabral, João Valente, João Silva, Urândia Aragão, Carla Fernandes, Nuno Correia:
A creation-tool for contemporary dance using multimodal video annotation
 - Gianluca Monaci, Tommaso Gritti, Fabio Vignoli, Wouter Walmlink, Maarten Hendriks:
Flower power
 - Bauke Freiburg, Jaap Kamps, Cees G.M. Snoek:
Crowdsourcing visual detectors for video search
 - Alexander Reben, Joseph Paradiso:
A mobile interactive robot for gathering structured social video
 - Andreea Danielescu, Ryan P. Spicer, David Tinapple, Aisling Kelliher, Shawn Nikkila, Sean Burdick:
Abstract rendering of human activity in a dynamic distributed learning environment
 - Steve Mann, Ryan Janzen, Jason Huang:
"WaterTouch": an aquatic interactive multimedia sensory table based on total internal reflection in water
 - Abhishek Bhattacharya, Wanmin Wu, Zhenyu Yang:
Quality of experience evaluation of voice communication systems using affect-based approach
 - Mihalis A. Nicolaou, Hatice Gunes, Maja Pantic:
A multi-layer hybrid framework for dimensional emotion classification
 - Catherine H. Vuong, Todd Ingalls, James J. Abbas:
Transforming clinical rehabilitation into interactive multimedia
 - Ramin Tadayon, Ashish Amresh, Winslow Burleson:
Socially relevant simulation games: a design study
 - Ting Yao, Chong-Wah Ngo, Tao Mei:
Context-based friend suggestion in online photo-sharing community
 - Anastasia Gumulia, BartBomiej Puzon, Naoko Kosugi:
Music visualization: predicting the perceived speed of a composition -- misual project --
 - Xiuzhuang Zhou, Junlin Hu, Jiwen Lu, Yuanyuan Shang, Yong Guan:
Kinship verification from facial images under uncontrolled conditions
 - Mitchell J. Morris, John R. Kender:
VastMM-Tag: a semantic tagging browser for unstructured videos
 - Qiyam Tung, Ranjini Swaminathan, Alon Efrat, Kobus Barnard:
Expanding the point: automatic enlargement of presentation video elements
 - Senthil Kumar, Sreedal Menon, Francis Zane:
Sharing rectangular objects in a video conference
 - Hao Ji, Fei Su:
Biased metric learning for person-independent head pose estimation
 - Dong Liu, Shuicheng Yan, Hong-Jiang Zhang:
Next photo please: towards visually consistent sequential photo browsing

- Ngo Quang Minh Khiem, Guntur Ravindra, Wei Tsang Ooi:
Towards understanding user tolerance to network latency in zoomable video streaming
- Marian Ursu, Pedro Torres, Vilmos Zsombori, Michael Franzis, Rene Kaiser:
Socialising through orchestrated video communication
- Fernanda Brandi, Eckehard Steinbach:
Perceptual coding of recorded telemanipulation sessions
- Jacopo Staiano, Bruno Lepri, Ramanathan Subramanian, Nicu Sebe, Fabio Pianesi:
Automatic modeling of personality states in small group interactions
- Si Liu, Qiang Chen, Jian Dong, Shuicheng Yan, Changsheng Xu, Hanqing Lu:
Snap & play: auto-generate personalized find-the-difference mobile game
- Aveek Shankar Brahmachari, Sudeep Sarkar:
Fast detection of noisy GPS and magnetometer tags in wide-baseline multi-views
- Linjun Yang, Yang Cai, Alan Hanjalic, Xian-Sheng Hua, Shipeng Li:
Video-based image retrieval
- Roberto Yus, Eduardo Mena, Jorge Bernad, Sergio Ilarri, Arantza Illarramendi:
Location-aware system based on a dynamic 3D model to help in live broadcasting of sport events
- Ziyang Tang, Orkun Ozbek, Xiaohu Guo:
Real-time 3D interaction with deformable model on mobile devices
- Jia Hao, Guanfeng Wang, Beomjoo Seo, Roger Zimmermann:
Keyframe presentation for browsing of user-generated videos on map interfaces
- Xunyi Yu, Aura Ganz:
Detecting and identifying people in mobile videos
- Fang-Erh Lin, Yin-Hsi Kuo, Winston H. Hsu:
Multiple object localization by context-aware adaptive window search and search-based object recognition
- Matthew L. Cooper:
Clustering geo-tagged photo collections using dynamic programming
- Sam S. Tsai, David Chen, Huizhong Chen, Cheng-Hsin Hsu, Kyu-Han Kim, Jatinder P. Singh, Bernd Girod:
Combining image and text features: a hybrid approach to mobile book spine recognition
- Michael E. Houle, Vincent Oria, Shin'ichi Satoh, Jichao Sun:
Knowledge propagation in large image databases using neighborhood information
- Qiang Zhou, Shifeng Chen, Jianzhuang Liu, Xiaoou Tang:
Edge-preserving single image super-resolution
- Chen Cao, Shifeng Chen, Wei Zhang, Xiaoou Tang:
Automatic motion-guided video stylization and personalization
- Shiai Zhu, Chong-Wah Ngo, Yu-Gang Jiang:
On the pooling of positive examples with ontology for visual concept learning
- Yuming Fang, Zhenzhong Chen, Weisi Lin, Chia-Wen Lin:
Saliency-based image retargeting in the compressed domain
- Jian Yi, Yuxin Peng, Jianguo Xiao:
Mining concept relationship in temporal context for effective video annotation
- Xiangmin Zhou, Lei Chen, Xiaofang Zhou:
Structure tensor series-based matching for near-duplicate video retrieval
- Junfeng Jiang, Xiao-Ping Zhang:
A smart video player with content-based fast-forward playback
- Hongyuan Cai, Jiang Yu Zheng:
Video anatomy: cutting video volume for profile
- Harlyn Baker, Nelson L. Chang, Arun Paruchuri:
Capture and display for live immersive 3D entertainment
- Junhao Shi, Mingmin Zhang, Zhigeng Pan:
A real-time bimanual 3D interaction method based on bare-hand tracking
- Xinming Zhang, Zheng-Jun Zha, Changsheng Xu:
Learning "verb-object" concepts for semantic image annotation
- Yongqing Sun, Akira Kojima:
A novel method for semantic video concept learning using web images
- Jun Imura, Teppei Fujisawa, Tatsuya Harada, Yasuo Kuniyoshi:
Efficient multi-modal retrieval in conceptual space
- Xiangyu Wang, Yong Rui, Mohan S. Kankanhalli:
Up-fusion: an evolving multimedia decision fusion method
- Zhou Ren, Junsong Yuan, Zhengyou Zhang:
Robust hand gesture recognition based on finger-earth mover's distance with a commodity depth camera
- Xiaodong Yang, Shuai Yuan, YingLi Tian:
Recognizing clothes patterns for blind people by confidence margin based feature combination
- Milan Redzic, Conor Brennan, Noel E. O'Connor:
Dual-sensor fusion for indoor user localisation
- Zhonghua Li, Bingjun Zhang, Ye Wang:

Document dependent fusion in multimodal music retrieval

- Stevan Rudinac, Alan Hanjalic, Martha Larson:
Finding representative and diverse community contributed images to create visual summaries of geographic areas
- Luca Del Pero, Philip Lee, James Magahern, Emily Hartley, Kobus Barnard, Ping Wang, Atul Kanaujia, Niels Haering:
Fusing object detection and region appearance for image-text alignment
- Biao Han, Hao Zhu, Youdong Ding:
Bottom-up saliency based on weighted sparse coding residual
- Yuzhao Ni, Jian Dong, Jiashi Feng, Shuicheng Yan:
Purposive hidden-object-game: embedding human computation in popular game
- Mayank Bansal, Harpreet S. Sawhney, Hui Cheng, Kostas Daniilidis:
Geo-localization of street views with aerial image databases

SESSION: Short papers session 2

- Jitao Sang, Jing Liu, Changsheng Xu:
Exploiting user information for image tag refinement
- Kong-Wah Wan, Yan-Tao Zheng, Lekha Chaisorn:
Known-item video search via query-to-modality mapping
- Yang Yang, Yi Yang, Zi Huang, Heng Tao Shen:
Transfer tagging from image to video
- Srinivasan H. Sengamedu, Subhajit Sanyal, Sriram Satish:
Detection of pornographic content in internet images
- Chunlei Yang, Jialie Shen, Jianping Fan:
Effective summarization of large-scale web images
- Gang Yu, Junsong Yuan, Zicheng Liu:
Real-time human action search using random forest based hough voting
- Xin-Shun Xu, Xiangyang Xue, Zhi-Hua Zhou:
Ensemble multi-instance multi-label learning approach for video annotation task
- Ce Li, Jianru Xue, Nanning Zheng, Zhiqiang Tian:
Nonparametric bottom-up saliency detection using hypercomplex spectral contrast
- LiMin Wang, Yirui Wu, Tong Lu, Kang Chen:
Multiclass object detection by combining local appearances and context
- Yong Luo, Dacheng Tao, Bo Geng, Chao Xu, Stephen Maybank:
Shared feature extraction for semi-supervised image classification
- Yangxi Li, Bo Geng, Zheng-Jun Zha, Dacheng Tao, Linjun Yang, Chao Xu:

Difficulty guided image retrieval using linear multiview embedding

- Tianlong Chen, Shuqiang Jiang, Lingyang Chu, Qingming Huang:
Detection and location of near-duplicate video sub-clips by finding dense subgraphs
- Yingfei Li, Bo Geng, Zheng-jun Zha, Yangxi Li, Dacheng Tao, Chao Xu:
Query expansion by spatial co-occurrence for image retrieval
- Aixin Sun, Sourav S. Bhowmick, Jun-An Chong:
Social image tag recommendation by concept matching
- Wei Zhang, Yao Lu, Xiangyang Xue, Jianping Fan:
Automatic image annotation with weakly labeled dataset
- Wei Zhang, Ke Gao, Yongdong Zhang, Jintao Li:
Efficient approximate nearest neighbor search with integrated binary codes
- Peng Yang, Hui Li, Qingshan Liu, Lin Zhong, Dimitris Metaxas:
Content quality based image retrieval with multiple instance boost ranking
- Ying Zheng, Steve Gu, Carlo Tomasi:
Detecting motion synchrony by video tubes
- Teresa Bracamonte, Barbara Poblete:
Automatic image tagging through information propagation in a query log based graph structure
- Lyndon Kennedy, Malcolm Slaney:
Identifying authoritative sources of multimedia content: mining specificity and expertise from large-scale multimedia databases
- Kuiyuan Yang, Lei Zhang, Meng Wang, Hong-Jiang Zhang:
Semantic point detector
- Hsiao-Hang Su, Tse-Wei Chen, Chieh-Chi Kao, Winston H. Hsu, Shao-Yi Chien:
Scenic photo quality assessment with bag of aesthetics-preserving features
- Hao Xu, Jingdong Wang, Xian-Sheng Hua, Shipeng Li:
Hybrid image summarization
- Xavier Anguera, Juan Manuel Barrios, Tomasz Adamek, Nuria Oliver:
Multimodal fusion for video copy detection
- Wang Junqiang, Huadong Ma:
Pedestrian detection with geometric context from a single image
- Di Niu, Hong Xu, Baochun Li, Shuqiao Zhao:
Risk management for video-on-demand servers leveraging demand forecast
- Zhi Wang, Lifeng Sun, Shiqiang Yang, Wenwu Zhu:

- Prefetching strategy in peer-assisted social video streaming*
- Dan Miao, Wenwu Zhu, Chong Luo, Chang Wen Chen:
Resource allocation for cloud-based free viewpoint video rendering for mobile phones
 - Zhen Wei Zhao, Wei Tsang Ooi:
APRICOD: a distributed caching middleware for fast content discovery of non-continuous media access
 - Sebastiano Battiato, Giovanni Maria Farinella, Enrico Messina, Giovanni Puglisi:
Robust image registration and tampering localization exploiting bag of features based forensic signature
 - Xiangyang Xue, Wei Li, Yue Yin:
Towards content-based audio fragment authentication
 - Rui Min, Jean-Luc Dugelay:
Cap detection for moving people in entrance surveillance
 - Han-Ping Cheng, Yun-Chung Shen, Ja-Ling Wu, Kiyoharu Aizawa:
High efficient distributed video coding with parallelized design for cloud computing
 - Tse-Chung Su, Yun-Chung Shen, Ja-Ling Wu:
Real-time decoding for LDPC based distributed video coding
 - Shingo Uchihashi, Tsutomu Tanzawa:
Mixing remote locations using shared screen as virtual stage
 - Kuan-Ta Chen, Yu-Chun Chang, Po-Han Tseng, Chun-Ying Huang, Chin-Laung Lei:
Measuring the latency of cloud gaming systems
 - Hsiao-Yun Tseng, Yun-Chung Shen, Ja-Ling Wu:
Distributed video coding with compressive measurements
 - Lei Huang, Tian Xia, Ji Wan, Yongdong Zhang, Shouxun Lin:
Personalized portraits ranking
 - Jong-Seok Lee, Lutz Goldmann, Touradj Ebrahimi:
A new analysis method for paired comparison and its application to 3D quality assessment
 - Tommaso Gritti, Gianluca Monaci:
ImagiLight: a vision approach to lighting scene setting
 - Florian Schweiger, Georg Schroth, Michael Eichhorn, Eckehard Steinbach, Michael Fahrmaier:
Consensus-based cross-correlation
 - Junyong You, Touradj Ebrahimi, Andrew Perkis:
Modeling motion visual perception for video quality assessment
 - Zhiding Yu, Chunjing Xu, Jianzhuang Liu, Oscar C. Au, Xiaoou Tang:
Automatic object segmentation from large scale 3D urban point clouds through manifold embedded mode seeking
 - Zhiqiang Tian, Jianru Xue, Xuguang Lan, Ce Li, Nanning Zheng:
Key object-based static video summarization
 - Nick C. Tang, Chiou-Ting Hsu, Tsung-Yi Lin, Hong-Yuan Mark Liao:
Example-based human motion extrapolation based on manifold learning
 - Jiangbo Lu, Viet Anh Nguyen, Zeping Niu, Bhavdeep Singh, Zhiping Luo, Minh N. Do:
CuteChat: a lightweight tele-immersive video chat system
 - Haiyang Ma, Deepak Gangadharan, Nalini Venkatasubramanian, Roger Zimmermann:
Energy-aware complexity adaptation for mobile video calls
 - Yanjie Li, Lifeng Sun, Tianfan Xue:
Fast frame-rate up-conversion of depth video via video coding
 - Mu Mu, Johnathan Ishmael, Keith Mitchell, Nicholas Race, Andreas Mauthe:
Multimodal QoE evaluation in P2P-based IPTV systems
 - Steve Mann, Jason Huang, Ryan Janzen, Raymond Lo, Valmiki Rampersad, Alexander Chen, Taqveer Doha:
Blind navigation with a wearable range camera and vibrotactile helmet
 - Snehasis Mukherjee, Sujoy Kumar Biswas, Dipti Prasad Mukherjee:
Recognizing interaction between human performers using 'key pose doublet'
 - Huayou Su, Chunyuan Zhang, Jun Chai, Mei Wen, Nan Wu, Ju Ren:
High-efficient software parallel CAVLC encoder based on programmable stream processor
 - Pengjie Wang, Rynson W.H Lau, Mingmin Zhang, Jiang Wang, Haiyu Song, Zhigeng Pan:
A real-time database architecture for motion capture data
 - Jay Geagan, Dulce Ponceleon:
Once upon a time, i bought a movie and it played everywhere in my home
- SESSION: Short papers session 3*
- Hua Wang, Feiping Nie, Heng Huang, Yi Yang:
Learning frame relevance for video classification
 - Wengang Zhou, Houqiang Li, Yijuan Lu, Qi Tian:
Large scale image search with geometric coding
 - Xianwang Wang, Tong Zhang:
Clothes search in consumer photos via color matching and attribute learning
 - Nakamasa Inoue, Koichi Shinoda:

- A fast MAP adaptation technique for gmm-supervector-based video semantic indexing systems*
- Shen-Fu Tsai, Liangliang Cao, Feng Tang, Thomas S. Huang:
Compositional object pattern: a new model for album event recognition
 - Linjun Yang, Alan Hanjalic:
Learning from search engine and human supervision for web image search
 - Bor-Chun Chen, Yin-Hsi Kuo, Yan-Ying Chen, Kuan-Yu Chu, Winston Hsu:
Semi-supervised face image retrieval using sparse coding with identity constraint
 - Hong Lu, Renzhong Wei, Yanran Shen, Xiangyang Xue:
Level influence of spatial pyramid matching in object classification
 - Xin-Shun Xu, Yuan Jiang, Liang Peng, Xiangyang Xue, Zhi-Hua Zhou:
Ensemble approach based on conditional random field for multi-label image and video annotation
 - Yingbin Zheng, Renzhong Wei, Hong Lu, Xiangyang Xue:
Refining local descriptors by embedding semantic information for visual categorization
 - Hongtao Xie, Ke Gao, Yongdong Zhang, Jintao Li, Huamin Ren:
Common visual pattern discovery via graph matching
 - Feng Su, Li Yang, Tong Lu, Gongyou Wang:
Environmental sound classification for scene recognition using local discriminant bases and HMM
 - Yang Liu, Yan Liu, Shenghua Zhong, Keith C.C. Chan:
Semi-supervised manifold ordinal regression for image ranking
 - Bolan Su, Shijian Lu, Chew Lim Tan:
Blurred image region detection and classification
 - Zhongwei Cheng, Lei Qin, Qingming Huang, Shuqiang Jiang, Shuicheng Yan, Qi Tian:
Human group activity analysis with fusion of motion and appearance information
 - Lingqiao Liu, Lei Wang:
Exploring latent class information for image retrieval using the bag-of-feature model
 - Zhiwu Lu, Yuxin Peng:
Combining latent semantic learning and reduced hypergraph learning for semi-supervised image categorization
 - Shayok Chakraborty, Vineeth Balasubramanian, Sethuraman Panchanathan:
Optimal batch selection for active learning in multi-label classification
 - Yuta Nakashima, Noboru Babaguchi:
Extracting intentionally captured regions using point trajectories
 - Chih-Fan Chen, Yu-Chiang Frank Wang:
Exploring self-similarities of bag-of-features for image classification
 - Pengjie Li, Huadong Ma, Anlong Ming:
Non-rigid 3D model retrieval using multi-scale local features
 - Miriam Redi, Bernard Merialdo:
Marginal-based visual alphabets for local image descriptors aggregation
 - Christian Beecks, Anca Maria Ivanescu, Steffen Kirchhoff, Thomas Seidl:
Modeling multimedia contents through probabilistic feature signatures
 - Christian Wengert, Matthijs Douze, Hervé Jégou:
Bag-of-colors for improved image search
 - Mihir Jain, Hervé Jégou, Patrick Gros:
Asymmetric hamming embedding: taking the best of our bits for large scale image search
 - Xiangang Cheng, Liang-Tien Chia:
Spatially-coherent pyramid matching based on max-pooling
 - Daan T.J. Vreeswijk, Bouke Huurnink, Arnold W.M. Smeulders:
Text and image subject classifiers: dense works better
 - Damian Borth, Adrian Ulges, Thomas Michael Breuel:
Automatic concept-to-query mapping for web-based concept detector training
 - Yueting Zhuang, Yang Liu, Fei Wu, Yin Zhang, Jian Shao:
Hypergraph spectral hashing for similarity search of social image
 - Michele Merler, John R. Kender:
Selecting the best faces to index presentation videos
 - Sheng He, Junwei Han, Xintao Hu, Ming Xu, Lei Guo, Tianming Liu:
A biologically inspired computational model for image saliency detection
 - Xiaoshuai Sun, Hongxun Yao, Rongrong Ji, Xianming Liu, Pengfei Xu:
Unsupervised fast anomaly detection in crowds
 - Sicheng Zhao, Hongxun Yao, Xiaoshuai Sun, Pengfei Xu, Xianming Liu, Rongrong Ji:
Video indexing and recommendation based on affective analysis of viewers
 - Brett Adams, Dinh Phung, Svetha Venkatesh:
Eventscapes: visualizing events over time with emotive facets
 - Minh-Son Dao, Duc-Tien Dang-Nguyen, Francesco G.B. De Natale:
Signature-image-based event analysis for personal photo albums

Award opportunities

- Lin Pang, Juan Cao, Yongdong Zhang, Shouxun Lin: *Leveraging collective wisdom for web video retrieval through heterogeneous community discovery*
- Keiichiro Hoashi, Chihiro Ono, Daisuke Ishii, Hiroshi Watanabe: *Automatic preview generation of comic episodes for digitized comic search*
- Xiangqian Yu, Vincent Oria, Pierre Gouton, Geneviève Jomier: *2D geon based generic object recognition*
- Ying Yuan, Fei Wu, Yueting Zhuang, Jian Shao: *Image annotation by composite kernel learning with group structure*
- Xiaofeng Zhu, Zi Huang, Heng Tao Shen: *Video-to-shot tag allocation by weighted sparse group lasso*
- Zheshen Wang, Mrityunjay Kumar, Jiebo Luo, Baoxin Li: *Extracting key frames from consumer videos using bi-layer group sparsity*
- Xia Li, Yan Song, Yijuan Lu, Qi Tian: *Spatial pooling for transformation invariant image representation*
- Rui Zhang, Lei Zhang, Xin-Jing Wang, Ling Guan: *Multi-feature pLSA for combining visual features in image annotation*
- Yue Gao, Meng Wang, Huanbo Luan, Jialie Shen, Shuicheng Yan, Dacheng Tao: *Tag-based social image search with visual-text joint hypergraph learning*
- Xiaojian Zhao, Guanda Li, Meng Wang, Jin Yuan, Zheng-Jun Zha, Zhoujun Li, Tat-Seng Chua: *Integrating rich information for video recommendation with multi-task rank aggregation*
- David S. Monaghan, Philip Kelly, Noel O'Connor: *Quantifying human reconstruction accuracy for voxelcarving in a sporting environment*
- Vladislavs Dovgalecs, Rémi Mégret, Yannick Berthoumieu: *Time-aware co-training for indoors localization in visual lifelogs*
- Yoshitaka Ushiku, Tatsuya Harada, Yasuo Kuniyoshi: *Automatic sentence generation from images*
- Vasant Manohar, Stavros Tsakalidis, Pradeep Natarajan, Rohit Prasad, Prem Natarajan: *Audio-visual fusion using bayesian model combination for web video retrieval*
- Lamberto Ballan, Marco Bertini, Alberto Del Bimbo, Giuseppe Serra: *Enriching and localizing semantic tags in internet videos*
- Kazuki Sawai, Tomokazu Takahashi, Daisuke Deguchi, Ichiro Ide, Hiroshi Murase: *Scene segmentation of wedding party videos by scenario-based matching with example videos*
- Aibo Tian, Xuemei Zhang, Daniel R. Tretter: *Content-aware photo-on-photo composition for consumer photos*
- Cheng-Te Li, Hsun-Ping Hsieh, Shou-De Lin: *PhotoFeel: feeling your photo collection with graph-based audiovisual flocking*
- Minwoo Park, Jiebo Luo, Andrew Gallagher, Majid Rabbani: *Learning to produce 3D media from a captured 2D video*
- Andreas Girgensohn, Frank Shipman, Lynn Wilcox, Qiong Liu, Chunyuan Liao, Yuichi Oneda: *A tool for authoring unambiguous links from printed content to digital media*
- Jung-Yu Yeh, Min-Chun Hu, Wen-Huang Cheng, Ja-Ling Wu: *Interactive digital scrapbook generation for travel photos based on design principles of typography*
- Patricia Wang, Xiaofeng Tong, Yangzhou Du, Jianguo Li, Wei Hu, Yimin Zhang: *Augmented makeover based on 3D morphable model*
- Yingbo Li, Bernard Merialdo, Mickael Rouvier, Georges Linares: *Static and dynamic video summaries*

Award opportunities

SIGMM Award for Outstanding PhD Thesis in Multimedia Computing, Communications and Applications

More informational: <http://sigmm.org/Awards/thesisaward>

Award Description

This award will be presented at most once per year to a researcher whose PhD thesis has the potential of very high impact in multimedia computing, communication and applications, or gives direct evidence of such impact. A selection committee will evaluate contributions towards advances in multimedia including multimedia processing, multimedia systems, multimedia network protocols and services, multimedia applications and interfaces. The award will recognize members of the SIGMM community and their research contributions in

their PhD theses as well as the potential of impact of their PhD theses in multimedia area. The selection committee will focus on candidates' contributions as judged by innovative ideas and potential impact resulting from their PhD work.

The award includes a US\$500 honorarium, an award certificate of recognition, and an invitation for the recipient to receive the award at a current year's SIGMM-sponsored conference, the ACM International Conference on Multimedia (ACM Multimedia). A public citation for the award will be placed on the SIGMM website, in the SIGMM Records e-newsletter as well as in the ACM e-newsletter.

Funding

The award honorarium, the award plaque of recognition and travel expenses to the ACM International Conference on Multimedia will be fully sponsored by the SIGMM budget.

Nomination Applications

Nominations will be solicited by the *1st May 2012* with an award decision to be made by August 30. This timing will allow a recipient to prepare for an award presentation at ACM Multimedia in that Fall (October/November).

The initial nomination for a PhD thesis must relate to a dissertation deposited at the nominee's Academic Institution between January and December of the year previous to the nomination. As discussed below, *some dissertations may be held for up to three years by the selection committee for reconsideration*. If the original thesis is not in English, a full English translation must be provided with the submission. Nominations for the award must include:

- PhD thesis (upload here)
- A statement summarizing the candidate's PhD thesis contributions and potential impact, and justification of the nomination (two pages maximum);
- Curriculum Vitae of the nominee
- Three endorsement letters supporting the nomination including the significant PhD thesis contributions of the candidate. Each endorsement should be no longer than 500 words with clear specification of nominee PhD thesis contributions and potential impact on the multimedia field.
- A concise statement (one sentence) of the PhD thesis contribution for which the award is being given. This statement will appear on the award certificate and on the website.

The nomination rules are:

- The nominee can be any member of the scientific community.
- The nominator must be a SIGMM member.
- No self-nomination is allowed.

If a particular thesis is considered to be of exceptional merit but not selected for the award in a given year, the selection committee (at its sole discretion) may elect to retain the submission for consideration in at most two following years. The candidate will be invited to resubmit his/her work in these years.

A thesis is considered to be outstanding if:

- Theoretical contributions are significant and application to multimedia is demonstrated.
- Applications to multimedia is outstanding, techniques are backed by solid theory with clear demonstration that algorithms can be applied in new domains - e.g., algorithms must be demonstrably scalable in application in terms of robustness, convergence and complexity.

The submission process of nominations will be preceded by the call for nominations. The call of nominations will be widely publicized by the SIGMM awards committee and by the SIGMM Executive Board at the different SIGMM venues, such as during the SIGMM premier ACM Multimedia conference (at the SIGMM Business Meeting) on the SIGMM web site, via SIGMM mailing list, and via SIGMM e-newsletter between September and December of the previous year.

Submission Process

- Register an account at <https://cmt.research.microsoft.com/SIGMM2012/> and upload one copy of the nominated PhD thesis. The nominee will receive a Paper ID after the submission.
- The nominator must then collate other materials detailed in the previous section and upload them as supplementary materials, *except the endorsement letters, which must be emailed separately as detailed below*.
- Contact your referees and ask them to send all endorsement letters to sigmmaward@gmail.com with the title: "PhD Thesis Award Endorsement Letter for [YourName]". The web administrator will acknowledge the receipt and the submission CMT website will reflect the status of uploaded documents and endorsement letters.

It is the responsibility of the nominator to follow the process and make sure documentation is complete. Thesis with incomplete documentation will be considered invalid.

Selection Committee

For the period 2010-2012, the award selection committee consists of:

- Professor Svetha Venkatesh (svetha.venkatesh@deakin.edu.au) - Chair (Deakin University, Australia)
- Professor Dick Bulterman (Dick.Bulterman@cwi.nl) (CWI, The Netherlands)
- Professor Abed El Saddik (abed@mcrlab.uottawa.ca) University of Ottawa, Canada

SIGMM Award for Outstanding Technical Contributions to Multimedia Computing, Communications and Applications

More informational: <mailto:lienhart@informatik.uni-augsburg.de>

2012 Call for Nominations of Multimedia Researchers for the SIGMM Award for Outstanding Technical Contributions to Multimedia Computing, Communications and Applications

Deadline for Nominations: May 31, 2012

Submit Nomination Material to Awards Committee Chair: Rainer Lienhart (lienhart@informatik.uni-augsburg.de)

Award Description

This award is presented every year to a researcher who has made significant and lasting contributions to multimedia computing, communication and applications. Outstanding technical contributions through research and practice are recognized. Towards this goal, contributions are considered from academia and industry that focus on major advances in multimedia including multimedia processing, multimedia content analysis, multimedia systems, multimedia network protocols and services, and multimedia applications and interfaces. The award recognizes members of the community for long-term technical accomplishments or those who have made a notable impact through a significant technical innovation. The selection committee focuses on candidates' contributions as judged by innovative ideas, influence in the community, and/or the technical/social impact resulting from their

work. The award includes a \$1000 honorarium, an award certificate of recognition, and an invitation for the recipient to present a keynote talk at a current year's SIGMM-sponsored conference, the ACM International Conference on Multimedia (ACM Multimedia). A public citation for the award will be placed on the SIGMM website.

Funding

The award honorarium, the award certificate of recognition and travel expenses to the ACM International Conference on Multimedia is fully sponsored by the SIGMM budget.

Nomination Process

Nominations are solicited by May 31, 2012 with decision made by July 30 2012, in time to allow the above recognition and award presentation at ACM Multimedia 2012. Nominations for the award must include:

1. A statement summarizing the candidate's accomplishments, description of the significance of the work, and justification of the nomination (two pages maximum);
2. Curriculum Vitae of the nominee;
3. Three endorsement letters supporting the nomination including the significant contributions of the candidate. Each endorsement should be no longer than 500 words with clear specification of nominee contributions and impact on the multimedia field;
4. A concise statement (one sentence) of the achievement(s) for which the award is being given. This statement will appear on the award certificate and on the website.

The nomination rules are:

1. The nominee can be any member of the scientific community.
2. The nominator must be a SIGMM member.
3. No self-nomination is allowed.
4. Nominations that do not result in an award can be resubmitted.
5. The SIGMM elected officers as well as members of the Awards Selection Committee are not eligible.

Previous Recipients

- 2011: Shi-Fu Chang for pioneering research and inspiring contributions in multimedia analysis and retrieval"
- 2010: Ramesh Jain

"for pioneering research and inspiring leadership that transformed multimedia information processing to enhance the quality of life and visionary leadership of the multimedia community"

- 2009: Lawrence A. Rowe
"for pioneering research in continuous media software systems and visionary leadership of the multimedia research community"
- 2008: Ralf Steinmetz
"for pioneering work in multimedia communications and the fundamentals of multimedia synchronization"

ACM TOMCCAP Nicolas D. Georganas Best Paper Award

More informational: <http://tomccap.acm.org/>

Call for Nominations: ACM Transactions on Multimedia Computing, Communications and Applications Nicolas D. Georganas Best Paper Award

The Editor in Chief of ACM TOMCCAP invites you to nominate candidates for the ACM Transactions on Multimedia Computing, Communications and Applications Nicolas D. Georganas Best Paper.

The award is given annually to the author(s) of an outstanding paper published in ACM TOMCCAP within the previous legal year from January 1 until December 31. The award carries a plaque as well as travel funds to the ACM MM conference where the awardee(s) will be honored. Nominations for the award must include the following:

- A statement describing the technical contributions of the nominated paper and a description of the significance of the paper. The statement should not exceed 500 words. No self-nomination is accepted.
- Two additional supporting statements by recognized experts in the field regarding the technical contribution of the paper and its significance to the respective field.

Nominations will be reviewed by the Selection Committee and the winning paper will finally be voted by the TOMCCAP Editorial Board.

Deadline for nominations of papers published in 2011 (Volume 7) is the 15th of June 2012. Only papers published in regular issues (no Special Issues) can be nominated.

The ACM Multimedia Grand Challenge 2011 in a Nutshell

By Gerald Friedland and Yohan Jin

Gerald Friedland is a senior research scientist at the International Computer Science Institute and the 2011 chair of the ACM MM Grand Challenge. Contact him at fractor@icsi.berkeley.edu.

Yohan Jin is a Senior Data Scientist at Tapjoy and the 2011 chair of the ACM MM Grand Challenge. Contact him at yohanjin@gmail.com.

The 2011 ACM Multimedia Grand Challenge proved to be the most competitive so far. This article provides a quick summary of the competition. More comprehensive coverage can be found in an IEEE Multimedia Magazine report.

When the ACM Multimedia Grand Challenge was started in 2009 it was a radical idea: instead of submitting conference papers on solutions to self-imposed problems, scientists from the multimedia community were encouraged to engage in problems formulated by industry sponsors, called the Grand Challenges. In continuation of what now has become a tradition, the 2011 Grand Challenge consisted of six challenges sponsored by five industry sponsors: HP, Yahoo, Technicolor, Nokia, and Huawei/3DLife. The 2011 challenges, of which some are likely to be continued in 2012 can be found at the original website.

The initial submissions were pared down to 12 finalists in a two-round peer review with two industry and two academic reviewers. Each finalist submitted a four-page short paper, published in the ACM Multimedia 2011 proceedings.

The 12 finalists then presented a 180-seconds elevator pitch at ACM Multimedia in Scottsdale, Arizona, in a quick succession followed by a 120-minute question grilling by the audience and a jury. The jury consisted of representatives of the companies sponsoring the challenges, namely Qian Lin (HP), Christophe Diot (Technicolor), and Timo Pekka Pylvanainen (Nokia). After the audience had enjoyed the fast-paced conference session with the 12 innovative presentations, they sat down with the chairs to reach a decision based on concrete selection criteria formulated as part of the call. A major focus of the criteria was on the solution's workability and generalizability as well as on the reproducibility of the results by other researchers.

The first prize was the only grading that the jury was guaranteed to award. It was nominated with a cash award of \$500 and was reserved for the finalists who stood out in addressing the selection criteria. The second prize, for the runner up, was optional. Two special prizes optionally recognized the projects that showed the Best Integration of Multiple Media and the Best Application of a Theory Framework.

Here are the winners:

- The first prize was awarded to "Photo Search by Face Positions and Facial Attributes on Touch Devices" by Yu-Heng Lei, Yan-Ying Chen, Lime Iida, Bor-Chun Chen, Hsiao-Hang Su, and Winston H. Hsu from National Taiwan University, Taipei, Taiwan.
- The second prize was given to "Towards Textually Describing Complex Video Contents with Audio-Visual Concept Classifiers" by Chunk Chet Tan, Yu-Gang Jiang, Chong-Wah Ngo. (the first and third authors are with the Department of Computer Science, City University of Hong Kong, Kowloon, Hong Kong, and the second author is with the School of Computer Science, Fudan University, Shanghai, China).
- The Best Application of a Theory Framework special award was given to "Understanding Images with Natural Sentences" by Yoshitaka Ushiku, Tatsuya Harada, and Yasuo Kuniyoshi, who are all at the University of Tokyo's Graduate School of Information Science and Technology. Tatsuya Harada is also at JST PRESTO.
- The Best Integration of Multiple Media special award was given to "News Browsing System: Multimodal Analysis" by Bruno do Nascimento Teixeira, Júlia Epischina Engrácia de Oliveira, Fillipe Dias Moreira de Souza, Tiago Oliveira Cunha, Igor Calil Loures de Oliveira, Chrisitane Okamoto, Lucas Figueiredo Gonçalves, Vinícius de Oliveira Silva, and Arnaldo de Albuquerque Araújo from the Department of Computer Science at the Federal University of Minas Gerais Belo Horizonte, Brazil.

The ACM Multimedia Grand Challenge is a continuing event at ACM Multimedia, and the 2012 conference is carrying the tradition into the fourth year. The new call can be found in this issue of the SIGMM Records and on the ACM Multimedia 2012 web pages.

We encourage everybody to consider participation as the competition has not only become prestigious but participants also report to have had a lot of fun!

References

1. Y.-H. Lei et al., "Photo Search by Face Positions and Facial Attributes on Touch Devices," Proc. ACM Multimedia, ACM Press, 2011, pp. 651-654.

2. C. Chet Tan et al., "Towards Textually Describing Complex Video Contents with Audio-Visual Concept Classifiers by C," Proc. ACM Multimedia, ACM Press, 2011, pp. 655-658.
3. Y. Ushiku, T. Harada, and Y. Kuniyoshi, "Understanding Images with Natural Sentences," Proc. ACM Multimedia, ACM Press, 2011, pp. 679-682.
4. B. do Nascimento Teixeira et al., "News Browsing System: Multimodal Analysis," Proc. ACM Multimedia, ACM Press, 2011, pp. 671-674.
5. G. Friedland, Y. Jin, Industry Dares You: The ACM Multimedia Grand Challenge 2011," IEEE Multimedia Magazine, 19(1), Jan 2012.

Open Source Column: Tribler: P2P Search, Share and Stream

Authors: Niels Zeilemaker and Johan Pouwelse
URL: <http://www.tribler.org>
by Niels Zeilemaker and Johan Pouwelse

Niels Zeilemaker, Delft University of Technology, The Netherlands niels@zeilemaker.nl

Johan Pouwelse, Delft University of Technology, The Netherlands peer2peer@gmail.com

Six years ago, we created a new open source P2P file sharing program called Tribler. During this time over one million users have used it, and three generations of Ph.D. students tested their algorithms in the real world.

Tribler is built around BitTorrent. Introduced in 2001, BitTorrent revolutionized the P2P world because of its unprecedented efficiency. However, some problems are not properly addressed in BitTorrent. First, it does not specify how to search the network, relying instead on central websites. These websites allow users to find and download small metadata files called torrents. A torrent describes the content and is required for downloading to start. Second, BitTorrent's unique method for downloading files is incompatible for streaming. This is due to the fact that it is optimized for speedy and reliable downloading, not providing a method for quick buffering.

Tribler is the first client which continuously tries to improve upon the basic BitTorrent implementation by addressing some of the flaws described above. It implements, amongst others, remote search, streaming, channels and reputation-management. All these

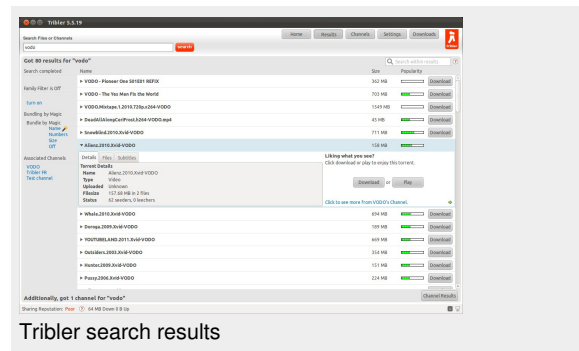
features are implemented in a completely distributed manner, not relying on any centralized component. Still, Tribler manages to remain fully backwards compatible with BitTorrent. Work on Tribler was initiated in 2005 and has been supported by multiple European grants.

In order to maximize the resource contribution of peers (other computers downloading/uploading the same file), BitTorrent splits a file into small pieces. This way, downloaders (called leechers) can upload completed pieces to other leechers, without the need to have the complete file first. Furthermore, uploading is encouraged by the tit-for-tat incentive mechanism: a leecher will rank its connected peers by their upload speed and will upload only to the fastest uploaders. Peers which have the complete file can help others by sending them pieces for free, these peers are called seeders.

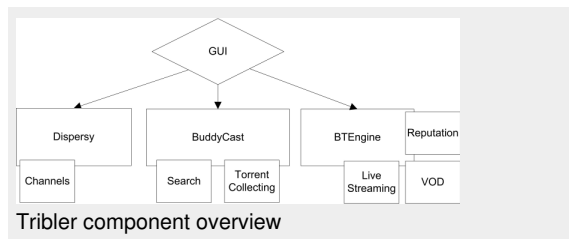
Before being able to download a file, a peer first has to obtain a torrent. This torrent describes the content of the file and includes the SHA-1 cryptographic hash per piece. This mechanism protects against transfer errors and malicious modifications of the file.

to this swarm and communicate with each other using the BuddyCast protocol. Peers are discovered by connecting to SuperPeers. Identical to 'normal' peers, but are considered to always be online. BuddyCast connects to a new peer every 15s and will exchange its preference list. This list contains the last downloads of a peer. By collecting them, we can calculate which peers are most similar to a given user. Those peers, taste buddies, are then used during search. While performing the BuddyCast handshake the user will exchange his current connections as well, allowing it to hillclimb towards finding his most similar peers and discovering new peers.

Search



Tribler Design and Features



A basic overview of Tribler is shown in the figure on the right, it consists of four distinct components.

- GUI: build using wxWidgets in order to be platform independent
- BTengine: a BitTorrent engine, which has been altered to allow for our customizations
- BuddyCast: our custom BitTorrent overlay, which is slowly being phased out
- Dispersy: our new custom protocol build with NAT traversal and distributed permissions

Tribler is built around PermlDs, permanent identities that allow us to identify the actions of users. PermlDs are stored as a public/private keypair and are used in Tribler to sign every message.

Communication between peers is established by using a 'special' BitTorrent swarm. All Tribler peers connect

Performing remote search in a decentralized manner has been a problem for many years. An early P2P protocol, Gnutella, used to send a message to all of its neighbors, which was then forwarded until a TTL of 7 was reached. Such an implementation is called flooding, as it causes a search-query to be sent to almost all peers in the network. Flooding a network is very quick, but it consumes huge amounts of bandwidth.

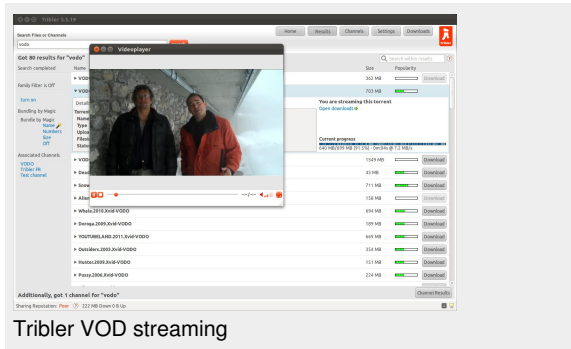
In contrast, Tribler uses a TTL of 1 (i.e., it only uses its neighbors to perform a remote search). Using the connections of the user to similar peers, we can obtain good results. Using only taste buddies, hitrates over 60% are possible. This figure is further improved by local caches deployed at every peer. The caches contain information for up to 50 000 torrents, which are used for improving search. Tribler connects to up to 10 taste buddies and to 10 random peers, thus allowing the user to search within up to 1050000 torrents.

A torrent is collected when our algorithms deem it interesting for a peer. This is calculated by using the same user-item matrix that is used to find similar peers. The user-item matrix is constructed by storing the BuddyCast preferences. Collaborative filtering allows us to 'recommend' torrents to be collected. Collected torrents are thus tailored for every user, resulting in

quicker search results, as we can display in the GUI the locally cached results before receiving any response from a peer.

More details are available in our papers [1,2].

Streaming



Tribler VOD streaming

Tribler supports two distinct types of streaming: Video-On-Demand and Live-Steaming. Both of them extend BitTorrent, by replacing one aspect of it. VOD requires a different approach for downloading pieces. The default policy of BitTorrent is to download the rarest piece first, ensuring the health of all pieces in the swarm. In contrast, in VOD we want to download the first few pieces as soon as possible to commence playback as soon as possible.

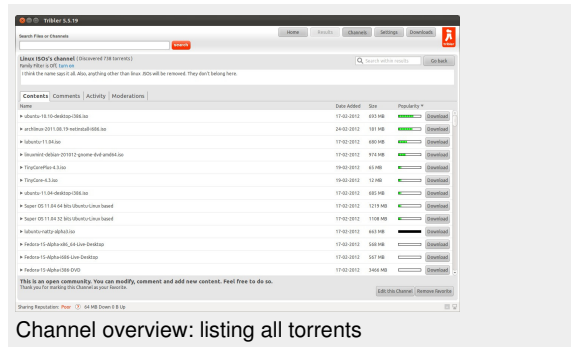
To allow for this we have defined three priorities (high, mid and low). Priorities are assigned to pieces, based on the current playback position. High priority pieces are downloaded in-order, thus allowing to start playback quickly. After all high priority pieces are downloaded, we start downloading mid priority pieces. Those are downloaded rarest-first, this to ensure we maintain the overall health of the swarm. Because mid priority pieces are only a subset of all available pieces, we still ensure that the playback-buffer remains stable. After downloading all high and mid priority pieces, we start downloading the low priority pieces, rarest-first. Because the playback position is moving forward, the priority of the pieces will be continuously modified.

Furthermore, we replaced the default BitTorrent incentive mechanism (tit-for-tat) with Give-to-Get. This incentive mechanism will rank peers according to their forwarding rank. A metric describing how well a peer is sending pieces to other peers. Full details are available in our paper [5].

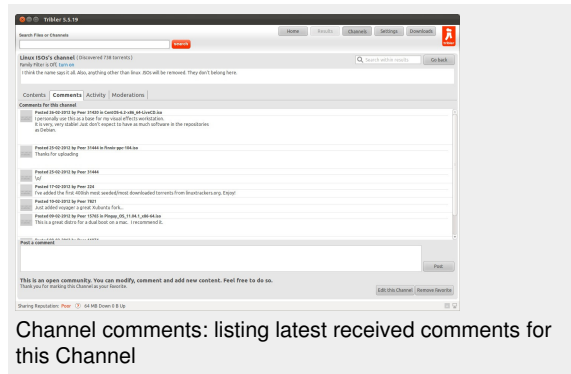
For live streaming we had to modify the actual torrent file. Because in live streaming pieces are not known beforehand, we cannot include their SHA-1 hashes. We

replaced the verification scheme by specifying the public key of the original source in the torrent file. Using the public key, every peer can then verify the validity of the pieces. Because live streaming may have an indefinite duration, we keep pieces that are at most 15 minutes old relative to our playback position.

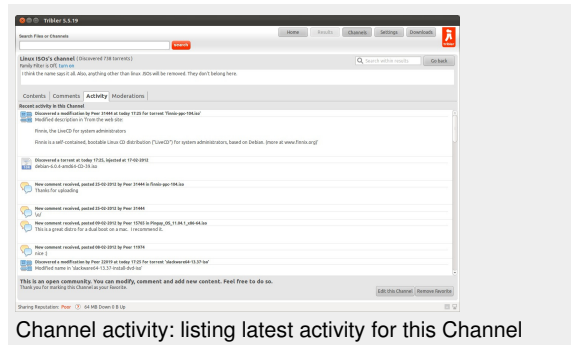
Channels



Channel overview: listing all torrents



Channel comments: listing latest received comments for this Channel



Channel activity: listing latest activity for this Channel

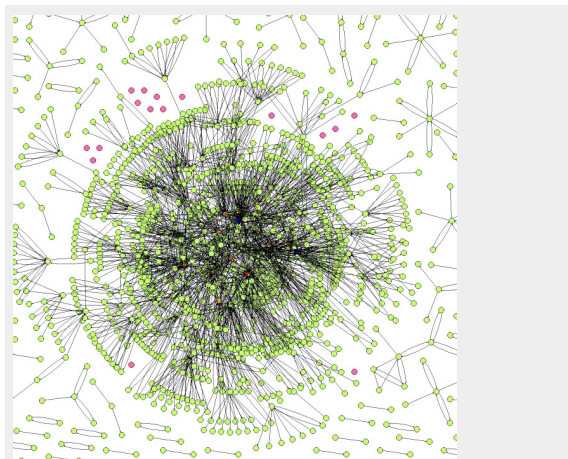
Since December 2011 we are evaluating, in the wild, the performance of our new transport protocol, Dispersy. Dispersy is the successor of BuddyCast, and it is focused on NAT-traversal. Instead of using TCP, Dispersy uses UDP. Furthermore, while BuddyCast implemented one global overlay to connect all Tribler-

peers using a 'special' BitTorrent swarm, Dispersy creates a separate overlay per protocol.

Using Dispersy we implemented Channels. These are created by users and consist of a list of torrents they like. Channels are implemented as separate Dispersy overlays and are discovered through one special overlay to which all peers connect. Channels have evolved from a simple list of torrents to a community in which users can comment on torrents, modify their name and description and organize playlists. Modifications are publicly visible, in a system that resembles Wikipedia. By allowing everyone to edit/improve the metadata of torrents, we hope to get a similar quality of experience as in Wikipedia. If a channel owner (the user creating the channel) does not want other users to interfere with his channel, he can limit which messages other users are able to send. This is enabled by using the decentralized permission system build into Dispersy, and it allows for a flexible configuration of channels.

By voting on a Channel, Dispersy will start to collect its contents. Before voting on a Channel only a snapshot of its content is available. More details regarding the voting is described in our paper [6]. Currently, popular channels have well over 30 000 torrents. Furthermore, our users have currently casted over 60 000 votes.

Reputation



BarterCast graph: showing data transfers between peers

A feature lacking in BitTorrent is the cross-swarm identification of peers. While downloading a file, peers have an incentive to upload to others due to tit-for-tat. But after completing a download, no incentives are in place to motivate a peer to keep uploading a file.

In order to address this, Tribler uses its PermlDs to identify Tribler peers in other swarms. Additionally, we employ a mechanism called BarterCast, which builds a

history of upload and download traffic between Tribler peers. We can then build a graph consisting of the download behavior of peers, scoring them accordingly. A peer which has shown to upload more than others is rewarded by being able to download at a faster rate, while lacking peers can be prevented from downloading at all. More details are available in our papers [3,4].

Acknowledgments

Since the start of the project in 2005 many many people have contributed to the project. Amongst others are A. Bakker, J.J.D. Mol, J. Yang, L. d'Acunto, J.A. Pouwelse, J. Wang, P. Garbacki, A. Iosup, J. Doumen, J. Roozenburg, Y. Yuan, M. ten Brinke, L. Musat, F. Zindel, F. van der Werf, M. Meulpolder, J. Taal, R. Rahman, B. Schoon and N.S.M. Zeilemaker. Tribler is a project which continues to evolve with the help of its community. Currently we have an active userbase who are commenting and suggesting features in the forums and continues to innovate together with our european partners.

If you are interested by the text above and want to try out Tribler, you can download it from our website <http://www.tribler.org>. Furthermore, the website has even more documentation of feature Tribler has and had, feel free to look around and leave a comment in the forums.

References

1. Pouwelse JA, Garbacki P, Wang J, et al. Tribler: A social-based peer-to-peer system. *Concurrency and Computation: Practice and Experience* 2008. Available at: <http://www3.interscience.wiley.com/journal/114219988/abstract>
2. Zeilemaker N, Capota M, Bakker A. Tribler: P2P media search and sharing. *Proceedings of the 19th ACM international conference on Multimedia (ACM MM)* 2011. Available at: <http://dl.acm.org/citation.cfm?id=2072433>
3. Meulpolder M, Pouwelse J. Bartercast: A practical approach to prevent lazy freeriding in p2p networks. *Sixth Int'l Workshop on Hot Topics in P2P Systems (HoT-P2P)* 2009. Available at: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5160954
4. Delaviz R, Andrade N, Pouwelse JA. Improving accuracy and coverage in an internet-deployed reputation mechanism. *IEEE Tenth International Conference on Peer-to-Peer Computing (IEEE P2P)* 2010. Available at: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5569965
5. Mol J, Pouwelse J, Meulpolder M, Epema D, Sips H. Give-to-get: Free-riding-resilient video-

on-demand in p2p systems. Fifteenth Annual Multimedia Computing and Networking (MMCN) 2008. Available at: <http://www.pds.ewi.tudelft.nl/pubs/papers/mmcn2008.pdf>

6. Rahman R, Hales D, Meulpolder M, Heinink V, Pouwelse JA and Sips H. Robust vote sampling in a P2P media distribution system. In: Proceedings of IPDPS 2009. Available at: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5160946
7. Tribler Protocol Specification. Available at: <http://svn.tribler.org/bt2-design/proto-spec-unified/trunk/proto-spec-current.pdf>

MM2012: The Multimedia Grand Challenge goes into its fourth year

Authors: Yushi Jin, Go Irie, Marcel Worring
URL: <http://www.acmmm12.org/call-for-multimedia-grand-challenge-solutions>
by Marcel Worring

Submission Deadline: May 14, 2012

Date: 29 Oct - 2 Nov, 2012

Location: Nara, Japan

MM2012: The Multimedia Grand Challenge goes into its fourth year

What problems do Google, HP, NHK, NTT, Technicolor, and other companies see driving the future of multimedia? The Multimedia Grand Challenge is a set of problems and issues from these industry leaders, geared to engage the Multimedia research community towards solving relevant, interesting, and challenging questions in the multimedia industry's 2-5 year horizon. The Grand Challenge was first presented as part of ACM Multimedia 2009 and in 2012 is having its fourth edition. Researchers are encouraged to submit working systems in response to these challenges to win the grand Challenge competition!

This year's challenges are partially continuation of earlier challenges so progress can be measured, while there are also a number of brand new challenges. The set of current challenges is as follows:

Google Challenge

Automatic Music Video Generation.

This challenge considers the automatic generation of cool and entertaining soundtracks for user-generated videos and automatic generation of interesting music videos.

3DLife / Huawei Challenge

Realistic Interaction In Online Virtual Environments.

This challenge calls for demonstrations of technologies that support real-time realistic interaction between humans in online virtual environments. A dataset related to a scenario in which an expert salsa teacher is performing over the Internet is provided. Motion data should be used to animate a realistic avatar of the teacher in an online virtual ballet studio.

HP challenge

Understanding the Emotional Impact of Images and Videos.

This challenge considers methods akin to the way professionals enhance the emotional value of imagery when creating websites, magazine covers, and printed advertisement with a special focus on understanding emotional impact and how this comes about.

NHK Challenge

Where is beauty? Video Segment Extraction Based on Aesthetic Quality Assessment.

This challenge asks for solutions to automatically recognize beautiful scenes in broadcast programs where the solution should both define beauty as well as providing the means to find it in video data.

NTT Docomo Challenge

Event Understanding through Social Media and its Text-Visual Summarization.

This challenge is seeking innovative techniques for data-mining on social media to retrieve, summarize, and visualize events for a selected topic such as a local event.

Technicolor Challenge

Audiovisual Recognition of Specific Events.

This challenge aims at exploring tools to search multimedia databases based on audio-visual queries in non-professional environments with wildly unstructured repositories by exploiting and combining several information channels.

Submission and awards

We encourage you to consider the challenges and submit your contribution to the ACM Multimedia 2012 Grand Challenge track.

The submissions must:

- Significantly address one of the industry challenges posted on the Grand Challenge web site.
- Depict working, presentable systems or demos.
- Describe why the system presents a novel and interesting solution.

Submit contributions or solutions described in a 4-page paper to the ACM Multimedia 2012 Grand Challenge track by May 14 through <https://cmt.research.microsoft.com/ACMMM2012/>

A number of solutions will be selected as finalists and invited to describe their work, demonstrate their solution and argue for the paper's success in the Grand Challenge Session in Nara. Each finalist will have several minutes to present their case. This year, every finalist team will receive a cash award equivalent to one student registration fee. Final winner(s) will be chosen by industry scientists, engineers and business luminaries based on the presentations.

See you in Nara!

MPEG Column: 99th MPEG Meeting

Authors: Christian Timmerer
URL: <http://multimediacommunication.blogspot.com/2012/03/mpeg-news-report-from-99th-meeting-san.html>
by Christian Timmerer

Christian Timmerer an assistant professor in the Department of Information Technology (ITEC), Multimedia Communication Group (MMC), Klagenfurt University, Austria. He has participated in ISO/MPEG work for several years, notably in the area of MPEG-21, MPEG-M, and MPEG-V.

MPEG news: a report from the 99th meeting, San Jose, CA, USA

The official press release is available here and I'd like to highlight two topics from MPEGs' 99th meeting in San Jose, CA, USA:

- HEVC advances to Committee Draft (CD)

- Public workshop on MPEG-H 3D Audio

High-Efficiency Video Coding reaches first formal milestone towards completion

As described in the official press release "*ISO/IEC's Moving Picture Experts Group (MPEG) is pleased to announce the completion of the ISO/IEC committee draft of the High Efficiency Video Coding (HEVC) standard developed by the Joint Collaborative Team on Video Coding (JCT-VC), a joint team between MPEG and the ITU-T's Video Coding Experts Group (VCEG)*". For those who are not familiar with the ISO/IEC standardization process, committee draft (CD) means that the standard is not yet finalized but entering the committee stage which enables national bodies to comment on the standard. That is, changes to HEVC can be only made through national body comments which needs to be registered in due time.

In terms of *performance of HEVC* one can conclude that the *mission is accomplished*. Preliminary HM5 vs. AVC subjective performance comparison looks impressive, i.e. > 50% bitrate reduction overall, specifically 67% in HD and 49% for WVGA sequences. Please note that these results are not validated through official verification tests which are usually conducted in a later stage of the standardization process.

From a deployment perspective *currently one profile is foreseen* which is preliminarily referred to as the "main" profile with a largest coding unit (LCU) between 16x16 and 64x64 and a max. pictures storage capacity always 6 (compared to AVC which is max. 16) among others.

Research issues: in my last report I wrote "the ultimate goal to have a performance gain of more than 50% compared to the predecessor which is AVC". It seems this has been achieved so one might wonder what else needs to be done. In practice, however, there is always space for improvement, right?

The next step in audio coding: MPEG-H 3D Audio

The MPEG-H 3D Audio Workshop attracted more than 100 attendees which followed presentations covering three areas of 3D audio.

1. ATSC 3.0 and the Future of Broadcast Television (FoBTV)
2. 22.2 multichannel sound for Ultra High Definition TV (UHDTV), Next Generation Broadcast Television,

and New Heights in Multichannel Sound: Explorations and Considerations

3. Realistic audio representation technologies for UHDTV, backward-compatible 3D audio coding, and innovating beyond 5.1.

The *presentations are publicly available* here within a single ZIP file. MPEG established an AhG on 3D Audio (and Audio Maintenance) with the following mandates (among others):

- Progress possible use cases, requirements and evaluation methods for 3D Audio
- Identify test material appropriate for 3D Audio work and a process to make the material available to interested MPEG delegates.

Subscription to the reflector is open to everyone. A possible timeline for part 3 of MPEG-H could mean to have a Call for Proposals (CfP) in July 2012 followed by the evaluation in January 2013, all preliminary, no guarantee.

Finally, the next meeting will be MPEGs' 100th meeting which will include a social event with participation of representatives from ITU, ISO, IEC, and others.

In this thesis the non-sequential delivery of media in dynamic networks is investigated. Consider a scenario where people participate at a social event. With the increased popularity of smart phones and tablet computers people produce more and more multimedia content. They share their content and consume it on popular web platforms. The production and the consumption of such media are, however, different from the typical sequential movie pattern: we call this non-sequential media access. If the infrastructure is not available, visitors cannot share their content with other visitors during the event. The idea is to connect the devices directly, which is further robust even if people move during the event (dynamic networks). Non-sequential media access in combination with dynamic networks brings new challenges for the whole multimedia life cycle. A formalism called Video Notation helps to define the single parts of the life cycle with a simple and short notation. New measures for transport are needed as well. A caching technique is introduced that allows for evaluating the goodness of content for being cached based on its popularity in different user groups. However, this cache does not cope with the dynamic network requirement, because such a delivery has to be robust, adaptive and scalable. Therefore, we concentrate on self-organizing algorithms that provide these characteristics. In this thesis the implemented algorithm is inspired by the endocrine system of higher mammals. A client can express its demands by creating hormones that will be released to the network. The corresponding resources are attracted by this hormone and travel towards a higher hormone concentration. This leads to a placement of content near to the users. Furthermore, the robustness and service quality is increased by placing replicas of the traveling content along the transport path. Unused replicas are automatically removed from the nodes, to ensure storage balancing. Finally, we show with a use case that a middleware based on the hormone-based delivery including well-defined interfaces to the user and to the network can be used for content delivery. For such a general application recommendations on possible configurations are made.

PhD thesis abstracts

Anita Sobe

Self-Organizing Multimedia Delivery



Advisor(s): Laszlo Böszörményi (supervisor, 1st reviewer), Pascal Felber (2nd reviewer)

Distributed Multimedia Systems Group

<http://www.aau.at/tewi/inf/itec/dms/>

Current research topics:

- Self-organizing Content Delivery
- Interactive Image and Video Search
- Multimedia Content Visualization

- Social Aspects of Multimedia Information Systems
- User-centered Multimedia Information Retrieval
- Creating Summaries and Stories out of Large Social Events
- Applications in the Medical Domain (Endoscopy) and in Traffic Surveillance

While our main interest lies in basic research, we aim to actively participate in the international scientific community and strive to apply our results in close cooperation with industry.

Dominik Kaspar

Multipath Aggregation of Heterogeneous Access Networks



The explosive deployment of wired and wireless communication infrastructure has recently enabled many novel applications and sparked new research problems. One of the unsolved issues in today's Internet - the main topic of this thesis - is the goal of increasing data transfer speeds of end hosts by aggregating and simultaneously using multiple network interfaces. This objective is most interesting when the Internet is accessible through several, relatively slow and variable (typically wireless) networks, which are unable to single-handedly provide the required data rate for resource-intensive applications, such as bulk file transfers and high-definition multimedia streaming.

Communication devices equipped with multiple network interfaces are now commonplace. Smartphones and

laptops are often shipped with built-in network adapters of different wireless technologies, typically enabling them to connect to wireless local area networks and cellular data networks (such as WLAN and HSPA). At the same time, wireless network coverage has become so widespread that mobile devices are often located in overlapping coverage areas of independent access networks. However, even if multiple interfaces are successfully connected to the Internet, operating systems typically use only a single default interface for data transmission, leaving secondary interfaces idle. This technical restriction is based on the fact that the majority of current Internet traffic is conveyed by transport protocols (TCP and UDP) that do not support multiple IP addresses per endpoint. Another crucial factor is that path heterogeneity introduces packet reordering, which can negatively affect the performance of transport protocols and strain the buffer requirements of applications.

This thesis explores the problem of multipath aggregation with the attempt to find solutions that achieve increased data throughput by concurrently utilizing heterogeneous and dynamic access networks. We strive for approaches that support existing applications without requiring significant modifications to the current infrastructure. The exploration starts at the IP layer with a study of the IP packet reordering that is caused by the use of heterogeneous paths. Our practical experiments confirm that multipath reordering exceeds the typical packet reordering in the Internet by an extent that renders the usual reordering metric useless. The outcome of our network-layer study is a proxy-based, adaptive multipath scheduler that is able to mitigate packet reordering while transparently forwarding a single transport connection over multiple paths.

After introducing a novel metric for quantifying the benefit of path aggregation, our analysis continues on the transport layer, where we investigate TCP's resilience to multipath-inflicted packet reordering. While IP packet reordering is known for its destructive effect on TCP's performance, our practical experiments indicate that a modern implementation is significantly more robust to packet reordering than standard TCP and achieves a substantial aggregation benefit even in certain cases of extreme multipath heterogeneity. In addition, we run a large set of emulation-based multipath experiments and identify several TCP parameters that lead to improved multipath performance when correctly tuned.

Finally, we present an application-layer solution that builds upon the idea of logical file segmentation for streaming a single video to multihomed clients. The novelty of this approach lies in diverting standard

protocol features (i.e., HTTP pipelining and range retrieval requests) from their intended purpose and using them for scheduling video segments over different paths. Interoperable with existing server infrastructure, our proposed solution can be deployed in a lightweight and purely client-based manner. We validate the proposed algorithms by implementing them into an existing video streaming platform.

Advisor(s): Supervisors: Paal Engelstad, Audun F. Hansen, Carsten Griwodz, Pål Halvorsen, Opponents: Anna Brunström, Olivier Bonaventure
 ISBN number: ISSN 1501-7710 (No. 1155)
<http://urn.nb.no/URN:NBN:no-30595>

Media Performance Group

<http://simula.no/departement/media>

The interaction with digital media pervades most people everyday lives, both at home and at work. A person may easily understand that he interacts with digital media, or be completely oblivious to it. Application ideas appear limitless, yet so appear inhibitors of new uses. Inhibitors range from unchangeable laws of physics through scalability problems faced in processing digital media to a lack of understanding for people's requirements to appreciate a particular application.

The Media department investigates the means of overcoming or evading inhibitors for the use of time-dependent digital media in distributed systems. The department finds solutions by exploring, understanding and improving on a particular inhibitor in an application context. Improvements are found in better operating systems mechanisms, programming tools, protocols, distributed architectures, digital media formats or a better understanding of people's perception of media in a context. The department takes a multimedia systems approach, whereby successful research leads to quantifiable improvements and success is proven experimentally. The essential results of the research are algorithms, methods, tools or prototypes that provide solutions for overcoming a particular set of challenges in using time-dependent digital media in distributed systems that are practical and realistic today or in the near future.

Sree Hari Krishnan Parthasarathi

Privacy-Sensitive Audio Features for Conversational Speech Processing



The work described in this thesis takes place in the context of capturing real-life audio for the analysis of spontaneous social interactions. Towards this goal, we wish to capture conversational and ambient sounds using portable audio recorders. Analysis of conversations can then proceed by modeling the speaker turns and durations produced by speaker diarization. However, a key factor against the ubiquitous capture of real-life audio is privacy. Particularly, recording and storing raw audio would breach the privacy of people whose consent has not been explicitly obtained.

In this thesis, we study audio features instead -- for recording and storage -- that can respect privacy by minimizing the amount of linguistic information, while achieving state-of-the-art performance in conversational speech processing tasks. Indeed, the main contributions of this thesis are the achievement of state-of-the-art performances in speech/nonspeech detection and speaker diarization tasks using such features, which we refer to, as privacy-sensitive. Besides this, we provide a comprehensive analysis of these features for the two tasks in a variety of conditions, such as indoor (predominantly) and outdoor audio. To objectively evaluate the notion of privacy, we propose the use of human and automatic speech recognition tests, with higher accuracy in either being interpreted as yielding lower privacy.

For the speech/nonspeech detection (SND) task, this thesis investigates three different approaches to privacy-sensitive features. These approaches are based on simple, instantaneous, feature extraction methods, excitation source information based methods, and feature obfuscation methods. These approaches are benchmarked against Perceptual Linear Prediction (PLP) features under many conditions on a large meeting dataset of nearly 450 hours. Additionally, automatic speech (phoneme) recognition studies on

TIMIT showed that the proposed features yield low phoneme recognition accuracies, implying higher privacy.

For the speaker diarization task, we interpret the extraction of privacy-sensitive features as an objective that maximizes the mutual information (MI) with speakers while minimizing the MI with phonemes. The source-filter model arises naturally out of this formulation. We then investigate two different approaches for extracting excitation source based features, namely Linear Prediction (LP) residual and deep neural networks. Diarization experiments on the single and multiple distant microphone scenarios from the NIST rich text evaluation datasets show that these features yield a performance close to the Mel Frequency Cepstral coefficients (MFCC) features. Furthermore, listening tests support the proposed approaches in terms of yielding low intelligibility in comparison with MFCC features.

The last part of the thesis studies the application of our methods to SND and diarization in outdoor settings. While our diarization study was more preliminary in nature, our study on SND brings about the conclusion that privacy-sensitive features trained on outdoor audio yield performance comparable to that of PLP features trained on outdoor audio. Lastly, we explored the suitability of using SND models trained on indoor conditions for the outdoor audio. Such an acoustic mismatch caused a large drop in performance, which could not be compensated even by combining indoor models.

Thesis Advisors: Dr. Daniel Gatica-Perez, Prof. Herve Bourlard; Thesis Jury President: Dr. Jean-Marc Vesin; Thesis Reviewers: Prof. Daniel P.W. Ellis, Prof. Simon King, Prof. Jean-Philippe Thiran
<http://library.epfl.ch/theses/?nr=5234>

Social computing group

<http://www.idiap.ch/~gatica/research-team.html>

Mohammad Kazemi Varnamkhasti

Multiple Description Video Coding Based on Base and Enhancement Layers of SVC and Channel Adaptive Optimization



Multiple distortion coding (MDC) is a promising solution for video transmission over lossy channels. In MDC, multiple descriptions of a source are generated which are dependently decodable and mutually refinable. When all descriptions are available, the corresponding quality is called central quality; otherwise it is called side quality. Generally, there exists a trade-off between side and central quality in all MDC schemes. MDC methods which provide better central-side quality trade-off are of more interest to designers.

In this thesis a new MDC scheme is introduced which has better trade-off between side and central quality compared to existing schemes. In other words, for the same central quality, it provides higher side quality; or equivalently for the same side quality, it has higher central quality. This method is based on the mixing of the base and enhancement layers of Coarse-Grain Scalable (CGS) coding and hence is called Mixed Layer MDC (MLMDC). At the central decoder the layers are separated and we have two-layer quality, such as in the CGS decoder. At the side decoder, some descriptions are not available and hence we cannot separate the layers, directly. We propose to use estimation for this purpose. MLMDC for two-description coding and four-description coding is implemented in JM16.0, H.264/AVC reference software. The experimental results show that for videos which have dynamic enough content (texture and motion activity), MLMDC in comparison to the conventional MDCs provides higher side quality for the same central quality. In addition, for video transmission over channels with packet loss (such as the Internet), MLMDC provides higher average video quality, in particular for four-description coding.

In error prone environments, we need to have higher side quality while in less noisy conditions higher central quality is more important. Therefore, in order to have the best quality in different channel conditions, optimization is needed. For this purpose, a new model for end-to-end distortion is introduced which takes into account both quantization and transmission distortions for predicting the quality at the receiver side. The transmission distortion is the result of error propagation which in turn originated from the mismatch between side and central decoder outputs. The derived model is applicable for all DCT-domain MDCs. With experimental results, the model is verified first and then used for MDCs optimization. The results show the performance of the optimizer and also MLMDC higher video quality compared to that of conventional MDCs when they are designed optimally.

While the thesis is in Farsi, its first part is available in English in the following paper:

Kazemi, M.; Sadeghi, K.H.; Shirmohammadi, S., "A Mixed Layer Multiple Description Video Coding Scheme", IEEE Transactions on Circuits and Systems for Video Technology, Volume: 22 Issue:2, Feb. 2012, pp. 202-215, <http://dx.doi.org/10.1109/TCSVT.2011.2159431>

Advisor(s): Khosro Haj Sadeghi (supervisor), Shervin Shirmohammadi (co-supervisor)
<http://site.uottawa.ca/~shervin/theses/2012-MohammadKazemi.pdf>

Mona Omidyeganeh

Parametric Analysis and Modeling of Video Signals



Video modeling and analysis have been of great interest in the video research community, due to their essential contribution to systematic improvements concerned in a wide range of video processing techniques. Parametric modeling and analysis of video provides appropriate means for processing the signal and the necessary mining of information for efficient representation of the signal. Video comparison, human action recognition, video retrieval, video abstraction, video transmission, and video clustering are some applications that can benefit from video modeling and analysis.

In this thesis, the parametric analysis and modeling of the video signal is studied through two schemes. In the first scheme, spatial parameters are first extracted from video frames and temporal evolution of these spatial parameters is investigated. Spatial parameters are selected based on the statistics of the 2D wavelet transform of the video frames, where wavelet transform provides a sparse representation of the signals and structurally conforms to the frequency sensitivity distribution of the human visual system. To analyze the temporal relations and progress of these spatial parameters, three methods are considered: inter-frame distance measurement, temporal decomposition, and Autoregressive (AR) modeling. In the first method, employing the Kullback-Leibler (KL) distance between spatial parameters as the similarity measure, the temporal evolution of the spatial features is studied. This analysis is used to determine shot boundaries, segment shots into clusters and select keyframes properly based on both similarity and dissimilarity criteria, within and outside the corresponding cluster, respectively. In the second method, the video signal is assumed to be a sequence of overlapping independent visual components called events, which typically are temporally overlapping compact functions that describe temporal evolution of a given set of the spatial parameters of the video signal. This event-based temporal decomposition technique is used for video abstraction, where no shot boundary detection or clustering is required. In the third method, the video signal is assumed to be a combination of spatial feature time series that are temporally approximated by the AR model. The AR model describes each spatial feature vector as a linear combination of the previous vectors within a reasonable time interval. Shot boundaries are well detected based on the AR prediction errors, and then at least one keyframe is extracted from each shot. To evaluate these models, subjective and objective tests, on TRECVID and Hollywood2 datasets, are conducted and simulation results indicate high accuracy and effectiveness of these techniques.

In the second scheme, video spatio-temporal parameters are extracted from 3D wavelet transform of the natural video signal based on the statistical

characteristics analysis of this transform. Joint and marginal statistics are studied and the extracted parameters are utilized for human action recognition and video activity level detection. Subjective and objective test results, on the popular Hollywood2 and KTH datasets, confirm high efficiency of this analysis method, as compared to the current techniques.

While the thesis is written in Farsi, the following English papers encompass some of the main technical aspects from the thesis:

1. M. Omidyeganeh, S. Ghaemmaghami, and S. Shirmohammadi, "Video Keyframe Analysis Using a Segment-Based Statistical Metric in a Visually Sensitive Parametric Space," IEEE Transactions on Image Processing, vol. 20, issue 10, pp. 2730:2737, 2011.
2. M. Omidyeganeh, S. Ghaemmaghami, and S. Shirmohammadi, "Group Based Spatio-Temporal Video Analysis and Abstraction Using Wavelet Parameters," Signal, Image and Video Processing, Springer, Accepted 2011, to appear.
3. M. Omidyeganeh, S. Ghaemmaghami, and S. Shirmohammadi, "Application of 3D-Wavelet Statistics to Video Analysis," Multimedia Tools and Applications, Springer, Accepted 2012, to appear.

Advisor(s): Supervisor: Shahrokh Ghaemmaghami, Co-supervisor: Shervin Shirmohammadi
<http://www.site.uottawa.ca/~shervin/theses/2012-MonaOmidyeganeh.pdf>

Distributed and Collaborative Virtual Environment Research Lab (DISCOVER Lab), University of Ottawa, Canada

<http://www.discover.uottawa.ca/>

Research at the DISCOVER Lab is directed towards the enhancement of next generation human-to-human communication through advanced multimedia technology and virtual environments. Through our many projects, we are developing new ideas and technology that will make easy-to-use virtual environments and mobile computing a reality. Research projects at the DISCOVER lab typically fall into the following categories:

- Networked Games and Collaborative Virtual Environments

- Mobile, 3D, and Multiview Video
- Haptics and Teleoperation
- Multimedia Systems and Applications
- 3D Physical Modelling and Animation
- Intelligent Sensor Networks and Ubiquitous Computing
- Multimedia-Assisted Rehabilitation Engineering

Xirong Li

Content-Based Visual Search Learned from Social Media



In a world with increasing amounts of digital pictures, content-based visual search is an important and scientifically challenging problem in ICT research. This thesis tackles the problem by learning from social media. The fundamental question addressed in this thesis is: what is the value of socially tagged images for visual search?

To that end, we propose the neighbor voting algorithm (Chapter 2) and its multi-feature variant (Chapter 3) to verify whether what people spontaneously say about an image is factually in the pictorial content. The two algorithms are used to find high-quality positive examples for learning automated image taggers. To obtain negative training examples without manual verification, we go beyond the classical random sampling approach by introducing informative negative bootstrapping (Chapter 4). For answering complex visual searches, we introduce the notion of bi-concepts as a retrieval method for unlabeled images in which two concepts are co-occurring (Chapter 5). Finally, as users

have their own associations with image semantics, we propose personalized image tagging by jointly exploiting personal tagging history and content-based analysis, optimized through Monte Carlo sampling (Chapter 6).

On the basis of the reported theories, algorithms, and experiments, this thesis has revealed the value of socially tagged images for content-based visual search, providing a basis for uncovering universal knowledge on images and semantics. With the methodologies established, this thesis opens up promising avenues for image search engines which provide access to the semantics of the visual content, but without the need of manual labeling.

Promotor: Arnold W.M. Smeulders; Co-Promotors: Marcel Worring, Cees G.M. Snoek; Opponents: James Z. Wang, Maarten de Rijke, Guus Schreiber, Alan Hanjalic
ISBN number: 9789461820822
<http://dare.uva.nl/en/record/410580>

Intelligent Systems Lab Amsterdam

<http://isla.science.uva.nl/>

The Intelligent Systems Lab Amsterdam ISLA at the University of Amsterdam performs fundamental, applied and spin-off research. We define intelligence as observing and learning; observing the world by video, still pictures, signals and text and abstracting knowledge or decisions to act from these observations.

Event and publication reports

TOMCCAP, Volume 8, Issue 1, January 2012

Editor-in-Chief: Ralf Steinmetz
URL: <http://tomccap.acm.org/>
Sponsored by ACM SIG Multimedia
Published: January 2012

Papers

- Ralf Steinmetz:
Editorial note and call for nominations: Nicolas D. Georganas best paper award
- Georghita Ghinea, Oluwakemi Ademoye:

The sweet smell of success: Enhancing multimedia applications with olfaction

- Mohamed Hefeeda, Cheng-Hsin Hsu:
Design and evaluation of a testbed for mobile TV networks
- Yu-Ru Lin, Hari Sundaram, Munmun De Choudhury, Aisling Kelliher:
Discovering multirelational structure in social media streams
- Xu Cheng, Jiangchuan Liu:
Exploring interest correlation for peer-to-peer socialized video sharing
- Tao Mei, Lusong Li, Xian-Sheng Hua, Shipeng Li:
ImageSense: Towards contextual image advertising
- Lauro Snidaro, Ingrid Visentini, Gian Luca Foresti:
Fusing multiple video sensors for surveillance
- Jiun-Long Huang, Shih-Chuan Chiu, Man-Kwan Shan:
Towards an automatic music arrangement framework using score reduction

TOMCCAP, Volume 8S, Issue 1, February 2012

Editor-in-Chief: Ralf Steinmetz
URL: <http://tomccap.acm.org/>
Sponsored by ACM SIG Multimedia
Published: February 2012

Papers

- Ralf Steinmetz:
Editorial note
- Dongyu Liu, Fei Li, Bo Shen, Songqing Chen:
Building an efficient transcoding overlay for P2P streaming to heterogeneous devices
- Zhijie Shen, Roger Zimmermann:
ISP-friendly P2P live streaming: A roadmap to realization
- Xiaosong Lou, Kai Hwang:
Quality of data delivery in peer-to-peer video streaming
- Chuan Wu, Baochun Li, Shuqiao Zhao:
Diagnosing network-wide P2P live streaming inefficiencies
- Chuan Wu, Zongpeng Li, Xuanjia Qiu, Francis C. M. Lau:
Auction-based P2P VoD streaming: Incentives and optimal scheduling
- Tieying Zhang, Xueqi Cheng, Jianming Lv, Zhenhua Li, Weisong Shi:
Providing hierarchical lookup service for P2P-VoD systems

MMSJ, Volume 18, Issue 1, February 2012

Special Issue on "SI on Mobile and Ubiquitous
Multimedia"
Guest Editors: Marios C. Angelides
Editor-in-Chief: Thomas Plagemann
URL: <http://www.springer.de/>
Published: February 2012

Papers

- Marios Angelides:
Editorial: Mobile and Ubiquitous Multimedia
- Dineshbabu Jayagopi, Taemie Kim, Alex Pentland
and Daniel Gatica-Perez:
*Privacy-sensitive recognition of group conversational
context with sociometers*
- Andreas Löcken, Tobias Hesselmann, Martin Pielot,
Niels Henze and Susanne Boll:
*User-centred process for the definition of free-hand
gestures applied to controlling music playback*
- Matthias Rehm and Karin Leichtenstern:
*Gesture-based mobile training of intercultural
behavior*
- Yefeng Liu, Vili Lehdonvirta, Todorka Alexandrova
and Tatsuo Nakajima:
*Drawing on mobile crowds via social media; Case
UbiAsk: image based mobile social search across
languages*
- Damon Daylamani Zad, Marios C. Angelides and
Harry Agius:
*Personalise your massively multiplayer online game
(MMOG) with Artemis*

MMSJ, Volume 18, Issue 2, March 2012

Special Issue on "Privacy-Aware Multimedia Surveillance
Systems"
Guest Editors: Pradeep K. Atrey, Sabu Emmanuel,
Sharad Mehrotra and Mohan S. Kankanhalli
Editor-in-Chief: Thomas Plagemann
URL: <http://www.springer.de/>
Published: March 2012

Papers

- Pradeep K. Atrey, Sabu Emmanuel, Sharad Mehrotra
and Mohan S. Kankanhalli:
*Guest editorial: Privacy-aware multimedia
surveillance systems*
- Thomas Winkler and Bernhard Rinner:
User-centric privacy awareness in video surveillance

- Alexander J. Aved and Kien A. Hua:
*A general framework for managing and processing
live video data with privacy protection*
- Sk. Md. Mizanur Rahman, M. Anwar Hossain,
Hussein Mouftah, Abdulmotaleb El Saddik and Eiji
Okamoto:
*Chaos-cryptography based privacy preservation
technique for video surveillance*
- Yuta Nakashima, Noboru Babaguchi and Jianping
Fan:
*Intended human object detection for automatically
protecting privacy in mobile video surveillance*
- Peng Zhang, Tony Thomas and Sabu Emmanuel:
*Privacy enabled video surveillance using a two state
Markov tracking algorithm*

Calls for contributions

Calls for SIGMM Sponsored and Co-sponsored Events

ACM Multimedia

Full paper Deadline: April 12, 2012
Event location: Nara, Japan
Event date: Oct 29 - Nov 2, 2012
URL: <http://www.acmmm12.org/>

ACM Multimedia is the worldwide premier multimedia conference and a key event to display scientific achievements and innovative industrial products. The Conference offers to scientists and practitioners in the area of Multimedia plenary scientific and technical sessions, tutorials, panels, and discussion meetings on relevant and challenging questions for the next years horizon of multimedia.

The Multimedia Grand Challenge @ ACM MM 2012

Full paper Deadline: May 14, 2012
Event location: Nara, Japan
Event date: Oct 29 - Nov 2, 2012
URL: <http://www.acmmm12.org/call-for-multimedia-grand-challenge-solutions/>

The Multimedia Grand Challenge is a set of problems proposed by industry leaders, geared to engage the Multimedia research community in solving relevant, interesting, and challenging questions about the industry's 3-5 year horizon for multimedia. The Grand Challenge was first presented as part of ACM

Multimedia 2009 and has established itself as a prestigious competition in the multimedia community.

Open Source Software Competition @ ACM MM 2012

Full paper Deadline: May 14, 2012
Event location: Nara, Japan
Event date: Oct 29 - Nov 2, 2012
URL: <http://www.acmmm12.org/call-for-open-source-software-competition/>

After several years of success, the Open-Source Software Competition is now an integral part of the ACM Multimedia program. It is mainly intended to celebrate and encourage the contribution of researchers, software developers and educators to advance the field by providing the community with implementations of codecs, middleware, frameworks, toolkits, libraries, multimedia players, applications, authoring tools, and other multimedia software.

Video Program @ ACM MM 2012

Full paper Deadline: May 14, 2012
Event location: Nara, Japan
Event date: Oct 29 - Nov 2, 2012
URL: <http://www.acmmm12.org/call-for-video-program/>

ACM Multimedia 2012 Video Program is intended to allow researchers and video producers to demonstrate their research results or video productions without having to bring the equipment for a live demo. Videos should present innovative research results, like for example merging augmented reality and real world entities, showing the use of multiple cameras, camera-arrays or video-sensor networks or communicate complex ideas through novel means and solutions or showing creative applications or MM-arts in new contexts.

Calls for Events held in cooperation with SIGMM

The 14th ACM Workshop on Multimedia and Security

Full paper Deadline: April 20, 2012
Event location: Coventry, UK
Event date: Sept. 6-7, 2012
URL: <http://www.mmsec12.com/>

Since its inauguration in 1998, MMSec has been intended for disseminating knowledge in the areas of multimedia security such as data protection, multimedia

forensics, hidden communications and biometrics, as well as topics in privacy, policy and legal issues in relation to multimedia security and forensics in real world applications.

10th International Conference on Advances in Mobile Computing and Multimedia

Full paper Deadline: July 15, 2012
Event location: Bali, Indonesia
Event date: December 3-5, 2012
URL: <http://www.iivas.org/conferences/momm2012/>

MoMM is a leading international conference for researchers and industry practitioners to share their new ideas, original research results and practical development experiences from all mobile computing and multimedia related areas.

14th International Conference on Multimodal Interfaces

Full paper Deadline: May 4, 2012
Event location: Santa Monica, CA
Event date: October 22-26th, 2012
URL: <http://www.acm.org/icmi/2012/>

ICMI is the premier international forum for multidisciplinary research on multimodal human-human and human-computer interaction, interfaces, and system development. The conference focuses on theoretical and empirical foundations, component technologies, and combined multimodal processing techniques that define the field of multimodal interaction analysis, interface design, and system development.

Other multimedia-related Events

6th International Conference on Distributed Smart Cameras

Full paper Deadline: June 10, 2012
Event location: Hong Kong
Event date: Oct 30 - Nov 2, 2012
URL: <http://www.icdsc.org/>

Technological developments in imaging, processing, and networking have created an opportunity for multi-disciplinary approaches to applications based on vision. The extensive availability and use of cameras in various application domains calls for the study of new embedded processing systems and algorithms.

6th IEEE International Conference on Semantic Computing

Full paper Deadline: May 4, 2012
 Event location: Palermo, Italy
 Event date: September 19-21, 2012
 URL: <http://icsc2012.pa.icar.cnr.it/>

The conference builds on the success of the past ICSC conferences as an international forum for researchers and practitioners to present research that advances the state of the art and practice of Semantic Computing, as well as identifying emerging research topics and defining the future of the field.

Free Material

ENST-Drums (Videos + Audios)

URL: <http://www.tsi.telecom-paristech.fr/aa0/?p=152>

The ENST-Drums database (only audio was previously distributed) is a varied research database for automatic drum transcription and processing:

- Three professional drummers specialized in different music genres were recorded.
- Total duration of audio material recorded per drummer is around 75 minutes.
- Each drummer played his own drum kit.
- Each sequence used either sticks, rods, brushes or mallets to increase the diversity of drum sounds.
- The drum kits themselves are varied, ranging from a small, portable, kit with two toms and 2 cymbals, suitable for jazz and latin music ; to a larger rock drum set with 4 toms and 5 cymbals.

Each sequence is recorded on 8 individual audio channels, is filmed from two angles, and is fully annotated. In this release, the public dataset is accessible both for audio and video content.

Olivier Gillet, Gaël Richard

QUASI database for the evaluation of music source separation

URL: <http://www.tsi.telecom-paristech.fr/aa0/?p=605>

This database is composed of 11 multitrack songs, each mixed in several ways by a professional sound engineer. Both the mixes and the original sources are provided under Creative Commons licences. These data have been used for the "Professionally produced music recordings" task of the SiSEC 2010 and 2011 evaluations.

Thomas Fillon and Gaël Richard, Télécom ParisTech
 Emmanuel Vincent, INRIA

Job Opportunities

Lecturer (Assistant Professor)

Employer: University College London

Valid until: 3 May 2012

More info: https://atsv7.wcn.co.uk/search_engine/jobs.cgi?owner=5041178&ownertype=fair&jcode=1244921

UCL Computer Science (<http://www.cs.ucl.ac.uk/>) is looking for a world-class researcher to strengthen and augment existing expertise in information retrieval, collaborative filtering, peer-to-peer architectures for retrieval and dissemination of information, (big) data mining, analytics, multimedia content analysis, web science, and analysis of social networks.

UCL is among the world's top universities, as reflected in performance in a range of rankings and tables. 21 Nobel prizewinners have come from the UCL community.

Back matter

Notice to Contributing Authors to SIG Newsletters

By submitting your article for distribution in this Special Interest Group publication, you hereby grant to ACM the following non-exclusive, perpetual, worldwide rights:

- to publish in print on condition of acceptance by the editor
- to digitize and post your article in the electronic version of this publication
- to include the article in the ACM Digital Library and in any Digital Library related services
- to allow users to copy and distribute the article for noncommercial, educational or research purposes

However, as a contributing author, you retain copyright to your article and ACM will refer requests for republication directly to you.

Impressum

Editor-in-Chief Carsten Griwodz, Simula Research Laboratory
Editors Stephan Kopf, University of Mannheim
 Viktor Wendel, Darmstadt University of Technology
 Lei Zhang, Microsoft Research Asia
 Pradeep Atrey, University of Winnipeg
 Christian Timmerer, Klagenfurt University
 Pablo Cesar, CWI