Volume 3, Number 3
September 2011

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

SIGMM RECORDS

Table of Contents

Association for Computing Machinery

Advancing Computing as a Science & Profession

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

SIGMM Records

Volume 3, Number 3, September 2011

Editorial

Carsten Griwodz

Dear Member of the SIGMM Community,

Welcome to the September issue of the SIGMM Records in 2011.

We announce this year's winner of the SIGMM Award for Outstanding Technical Contributions to Multimedia Computing, Communications and Applications. The award will be presented at the ACM Multimedia Conference 2011.

Christian Timmerer has newly joined the SIGMM Records as a columnist who will regularly inform you about the developments in MPEG standardization. Read his report from the 97th MPEG meeting.

You can read about the GameDays 2011, where scientists meet business people to discuss demands and learn about the latest research in serious games.

In the SGIMM education column, our new columnist Pradeep K. Atrey presents the course "Multimedia Systems Design" that is taught by Klara Nahrstedt at UIUC.

You can also read the PhD thesis summaries provided by four candidates who have recently passed their doctoral exams, you can find out about the top 10 most downloaded papers of SIGMM-sponsored events, and find shortcuts to the latest journal publications in TOMCCAP and MMSJ.

The Editors
Stephan Kopf
Viktor Wendel
Lei Zhang
Pradeep Atrey
Christian Timmerer

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

Authors: Rainer Lienhart
by Rainer Lienhart

The 2011 winner of the prestigious ACM Special Interest Group on Multimedia (SIGMM) award for Outstanding Technical Contributions to Multimedia Computing, Communications and Applications is Dr. Shih-Fu Chang. Dr. Shih-Fu Chang is currently professor in the department of electrical engineering and computer sciences at the Columbia University, NY. The ACM SIGMM Technical Achievement award, given in recognition of outstanding contributions over a researcher's career, cites Dr. Chang's "pioneering research and inspiring contributions in multimedia analysis and retrieval". The SIGMM award will be presented at the ACM International Conference on Multimedia 2011 that will be held Nov 28 - Dec 1 2011 in Scottsdale, Arizona, USA.

Prof. Dr. Shih-Fu Chang

Prof. Shih-Fu Chang has made significant contributions that shape directions in many key areas of multimedia, including multimedia search, video summarization, compressed-domain manipulation, and trustworthy media. His works have been highly influential, and with prestigious recognition and broad impact across research, education, and practical applications. In the 1990's, he and his students developed several of the first image/video search engines, such as VisualSEEk, VideoQ, and WebSEEk. He has also been recognized with technical awards and best paper awards for inventing novel systems that combine content analysis, adaptive mobile communication, and multimedia summarization. Other significant contributions include large-scale concept-based video search engines (e.g., CuZero), a widely used library of image classification models (e.g., Columbia374), international multimedia indexing and communication standards (e.g., MPEG-7 and MPEG-21), and large multimedia ontologies (e.g., LSCOM). His group demonstrated the best multimedia indexing performance in international benchmarking forums such as TRECVID (2008 and 2010). Chang also led the ADVENT university-industry research consortium with the participation of more than 25 industry sponsors. Many video indexing technologies developed by his group have been licensed to companies.

ACM is the professional society of computer scientists, and SIGMM is the special interest group on multimedia.

MPEG Column

URL: http://multimediacommunication.blogspot.com/2011/07/mpeg-news-report-from-97th-meeting.html
by Christian Timmerer

MPEG news: a report from the 97th meeting, Torino, Italy

The 97th MPEG meeting in Torino brought a few interesting news which I'd like to report here briefly. Of course, as usual, there is the official press release, however, I'd like to report on some interesting topics as follows:

- MPEG *Unified Speech and Audio Coding (USAC)* reached FDIS status
- Call for Proposals: *Compact Descriptors for Visual Search (CDVS)*
- Call for Proposals: *Internet Video Coding (IVC)*
- DIS on *MPEG Dynamic Adaptive Streaming over HTTP (DASH)*

MPEG Unified Speech and Audio Coding (USAC) reached FDIS status

ISO/IEC 23003-3 aka Unified Speech and Audio Coding (USAC) reached FDIS status and soon will be an International Standard. The FDIS itself won't be publicly available but the Unified Speech and Audio Coding Verification Test Report in September 2011.

Call for Proposals: Compact Descriptors for Visual Search (CDVS)

I reported previously about that and here comes the final CfP including the evaluation framework.

MPEG is planning standardizing technologies that will enable efficient and interoperable design of visual search applications. In particular we are seeking technologies for visual content matching in images or

video. Visual content matching includes matching of views of objects, landmarks, and printed documents that is robust to partial occlusions as well as changes in vantage point, camera parameters, and lighting conditions.

There are a number of component technologies that are useful for visual search, including format of visual descriptors, descriptor extraction process, as well as indexing, and matching algorithms. As a minimum, the format of descriptors as well as parts of their extraction process should be defined to ensure interoperability.

It is envisioned that a standard for compact descriptors will:

- ensure interoperability of visual search applications and databases,
- enable high level of performance of implementations conformant to the standard,
- simplify design of descriptor extraction and matching for visual search applications,
- enable hardware support for descriptor extraction and matching in mobile devices,
- reduce load on wireless networks carrying visual search-related information.

It is envisioned that such standard will provide a complementary tool to the suite of existing MPEG standards, such as MPEG-7 Visual Descriptors. To build full visual search application this standard may be used jointly with other existing standards, such as MPEG Query Format, HTTP, XML, JPEG, JPSec, and JPSearch.

The Call for Proposals and the Evaluation Framework is publicly available. From a research perspective, it would be interesting to see how technologies submitted as an answer to the CfP compete with existing approaches and applications/services.

In this context, it is probably worth looking at IEEE Multimedia Jul.-Sep. 2011 issue which is dedicated to visual content: identification and search including an overview about this new MPEG standard.

Call for Proposals: Internet Video Coding (IVC)

I reported previously about that and the final CfP for Internet Video Coding Technologies is available here. The requirements reveal some interesting issues the call is about:

- Real-time communications, video chat, video conferencing,
- Mobile streaming, broadcast and communications,
- Mobile devices and Internet connected embedded devices
- Internet broadcast streaming, downloads
- Content sharing.

Requirements fall into the following major categories:

- IPR requirements
- Technical requirements
- Implementation complexity requirements

Clearly, this work item has an optimization towards IPR but others are not excluded. In particular,

It is anticipated that any patent declaration associated with the Baseline Profile of this standard will indicate that the patent owner is prepared to grant a free of charge license to an unrestricted number of applicants on a worldwide, non-discriminatory basis and under other reasonable terms and conditions to make, use, and sell implementations of the Baseline Profile of this standard in accordance with the ITU-T/ITU-R/ISO/IEC Common Patent Policy.

MPEG Dynamic Adaptive Streaming over HTTP (MPEG-DASH)

For all DASH enthusiast, the latest - and probably almost final - version of DASH-related standards can be found here. Please note that DASH has been reorganized into MPEG-DASH referred to as ISO/IEC DIS 23009-1.2, Part 1: Media presentation description and segment formats. Additionally, you might be interested in the following draft:

- ISO/IEC 14496-12:2008/DAM 3, Part 12: ISO base media file format, AMENDMENT 3: DASH support and RTP reception hint track processing
- ISO/IEC FDIS 23001-7, Part 7: Common encryption format for ISO base media file format

All these DASH-related documents are publicly available here. In terms of implementation, the interested reader might check out the ITEC-DASH VLC-based implementation and GPAC (which provides basic support for DASH) respectively.

Further information you may find at the MPEG Web site, specifically under the hot news section and the press release. Working documents of any MPEG standard so far can be found here. If you want to join any of these activities, the list of Ad-hoc Groups (AhG) is available here including the information how to join their reflectors.

GameDays 2011

Authors: Viktor Wendel
URL: http://gamedays2011.de/
by Viktor Wendel

On behalf of the GameDays organizers, we wish to provide a report of the seventh GameDays, which have been held at September 12-13 at the Darmstadtium, the congress center of Darmstadt.

The GameDays are initiated and organized by Dr. Stefan Göbel, head of the Serious Gaming at the Multimedia Communications Lab at TU Darmstadt and take place as "Science meets Business" event in the field of Serious Games on an annual basis since 2005 in cooperation with Hessen-IT and other partners from science and industry.

Similar to 2010, the first day has been held as Int'l Conference focusing on "Serious Games - Theory, Technology & Practice" and was co-chaired by Prof. Dr. Josef Wiemeyer (TU Darmstadt, Institute for Sport Science). Altogether, 22 paper contributions have been submitted to the Int'l Scientific Committee and have been reviewed by at least five reviewers per paper. 11 papers have been accepted for the Conference Proceedings as full paper, another 6 as short, demo or poster presentation. Further, 6 papers (acceptance rate < 30%) have been selected for additional publishing as special issue on Serious Games within the Int'l Journal of Computer Science and Sports. The contributions and presentations have been clustered into the sessions "Educational Games", "Training Programs and Games for Health", and "Theory, Technology and Best Practice". Key notes have been provided by Patrick Felicia, Waterford Institute of Technology, Ireland about *Game-based Learning - Status Quo and Quo Vadis* and Roman Schönsee, Ranj Serious Games about *Remedial Games - Tools for Therapy*. The Conference brought together more than 100 participants - around 80% scientists - from Germany, Austria, Switzerland, The Netherlands, France, United Kingdom, Denmark and Hong Kong.

Stefan Goebel, Ruth Lemmen, Jörg Müller-Lietzkow, Georg Matzner, and Florian Stadlbauer; not in picture: Anja Herdel, Ralf Steinmetz

Highlights of the second day with about 150 participants included presentations about best practice examples for Serious Games in the fields of training and simulation (police training "ViPol" by TriCat) and games for health (AD(H)S game "Dr. Bonneys Zappelix Zaubert" by MultiMediaManufaktur), both games have been awarded with the Serious Games Award in silver at the Serious Games Conferences in 2010 resp. 2011. Further, the status quo of Serious Games has been discussed at a podium discussion with members of academia (Prof. Jörg Müller-Lietzkow, University of Paderborn and Prof. Ralf Steinmetz, TU Darmstadt), industry (Florian Stadlbauer, Deck13, Anja Herdel, Darmstadt Marketing and Ruth Lemmen, BIU - the german association for interactive entertainment industries) and state government (Georg Matzner, Hessian Ministry for economy, transport and state development).

In addition to the Science meets Business talks, the Games Park providing >30 demos of Serious Games (ranging from commercial titles to research prototypes and a few poster presentations) addressing the societal relevant domains education, learning and training as well as sports and health plus two examples the use of Serious Games for - scientifically sound - marketing and advertisement.

GameDays visitors playing one of the games on exhibition, a game for playful rehabilitation

Further info about the GameDays 2011, its program, exhibits etc. is available at http://www.gamedays2011.de.

The GameDays 2012 will take place at TU Darmstadt at September 18-20.

SIGMM Education Column

Authors: Klara Nahrstedt
URL: http://www.sigmm.org/Education
by Pradeep K. Atrey

Welcome to the SIGMM Educational Column. In this column, we report interesting and notable multimedia educational activities around the SIGMM community. This issue's column brings forward the Multimedia

Systems course taught in the Department of Computer Science at the University of Illinois, Urbana-Champaign, by Professor Klara Nahrstedt. In rest of this column, you will find the detail of the course in Prof. Nahrstedt's own words.

- Pradeep K. Atrey

Reflection and Discussion of the CS414 Multimedia Systems Design Course in the Department of Computer Science at the University of Illinois, Urbana-Champaign

The history of the course

I have designed and started to teach the Multimedia Systems course, CS414, in Spring 1996, based on the textbook Steinmetz & Nahrstedt: "Multimedia Computing, Communications and Applications", published in Prentice Hall 1995. Since then it has been offered every year.

The CS414 multimedia system course and the overall multimedia area have been started as a response to a large NSF (National Science Foundation) infrastructure grant that the Computer Science Department at University of Illinois, Urbana-Champaign received 1994-1995. As part of the grant missions, systematic teaching of multimedia systems design became one of the educational goals.

Three instructors taught this course over the last 15 years: Klara Nahrstedt, Karrie Karahalios and Brian Bailey.

The rationale for offering the course

The rationale for offering the course is to teach students an integrated view of the overall multimedia system. As we know, since multimedia data is time-sensitive, if one component in the end-to-end performance does not work right, the overall system performance suffers. For example, if networking component does not have a bounded delay control among distributed nodes, a conversational application such as video teleconference, Skype, collaborative WebEx, will have interactivity problems and usage issues. If video servers are not properly designed, we can see long starting delays to playback movies or overprovisioning to cache movies for fast upload and playback. So the main rationale of the course was and is to show students integration and interdependencies among media characteristics, compression, networking, operating system, storage, applications and user interfaces.

The target students for the course

The course is targeted for undergraduate students and first year graduate students. It is interesting that over the years, the makeup of the class changed. When I started to offer this course in 1996, 90-95% of the class was graduate students. Only very few undergraduate students had the background to take this course. The reason was that undergraduate students learned many important system concepts only in their senior year and so they run out of time to take the multimedia systems class which assumed that students had had already some systems course (operating system or distributed system or networking class). Hence, only graduate students who finished their undergraduate studies were able to take this class.

During the last 5-6 years, I have seen a great change and now 80-90% of students are undergraduate students. The reason is that I have designed a sophomore/junior class in 2003, called "systems programming", where undergraduate students get exposed to many basic system concepts such as threads, synchronization, file systems, memory management, socket programming. This course is now mandatory for each computer science undergraduate. With these systems skills under their belt, the undergraduate students are comfortable to take the multimedia systems design course and succeed.

The educational objective

The educational objective of the multimedia systems course is two-folded: (a) the students should be aware of the most important multimedia systems concepts that have been developed over the last 20 years in the multimedia systems community, and of their integrating aspect; (b) the students need to get hand-on experiences and design/implement a multimedia system to understand concepts and the integrating needs.

The topics covered

The class covers all major components of an end-to-end multimedia system, such as a conversational interactive system or an on-demand system. The course starts with discussion of multimedia characteristics, especially audio, images and video, and then goes into some depth of coding and compression, especially JPEG, MPEG-X/ MP3 and H.26X standards families. The compression topic is then followed with multimedia networking component, especially concepts such as Quality of Service aspects, protocols such as RSVP, IPv6, RTP/ RTCP, SIP, RTSP, streaming, peer-to-peer streaming, buffering, adaptation. The networking is followed by video servers and content distribution networks for VOD services, synchronization issues, and user interfaces. The course ends with applications discussions such as

Skype, retrieval /VOD systems (e.g., Flickr, PPLive), and others. All components are updated every year, but especially the applications module changes as new applications arise.

The mode of instruction (lecture, paper reading, etc.)

The mode of instruction is lectures, but since the course material is being regularly updated (multimedia systems area is fast changing area), for many lectures, I recommend supplementary paper reading. Even though I lecture, since the material is not in the textbook, the students need to read papers.

The course meets three times a week for 50 minutes lecturing. In addition to lectures, we offer tutorials on a need-to-know basis, exposing students to practical tools to be used in their projects. These tutorials and implementation Q&A sessions are being run by the teaching assistant. Usually these sessions are 3-4 times in one semester. Furthermore, the students have access to the multimedia laboratory where computers are equipped with cameras and microphones for their projects.

The course has usually 3-4 machine problems and through these machine problems, the students develop a small multimedia systems in a group, i.e., the students form groups at the beginning of the semester and then stay together throughout the semester to build the required system. Each machine problem (MP) has a deadline. The students are requested to do a careful system design and software engineering since they reuse their MPs. For example, in the MP1 they build a video and audio recorder and player (on a local machine). The goal is that they get familiar with the audio and video formats (e.g., MPEG2,/MPEG4) and some of the timing aspects of real-time audio and video when recording multimedia or playing them. In the MP2, the students experiment with control and streaming protocols between the sender(s) and the receiver(s) to enable captured multimedia to send over the Internet and play the multimedia in real-time at the receiver. In MP3, the students explore synchronization concepts between audio and video since they need to play the streamed video/audio flows together in synchronized way. In MP4, they put additional user interface and other application services together to demonstrate an interesting video-on-demand system, a teleconferencing system or other multimedia system.

Assignments/projects that is a result from the course

Every year the class builds a different multimedia system to expose students to the dynamics and changes that happen in this space. Over the years the students built (a) multimedia mailing system, (b) multi-party video-conferencing system, (c) authoring system, (d) peer-to-peer streaming system, (e) peer-to-peer video server, and others. In Spring 2011, the students built a video chat system between two Android phones.

Over the last five years, we have had industrial sponsors for the multimedia systems class who sponsored competition among the students in terms of the final demonstration project. Since each group in the class builds the same system, it is very easy to compare the results of each group. Over the years, the following companies sponsored the cs414 class competition: Pavlov, Technicolor, Google Inc., Vodafone and Qualcomm.

Feedback from students

The students always appreciate the hands-on experience with the machine problems. They would comment that the concepts are interesting, but having built a real system benefits them on the job market and in the future. Often they comment that MPs have been most helpful in understanding the material. One student commented "I have learned more in this class than any other and all of it seems very useful." Often students would come back years after they graduate and provide me a very positive feedback about this course, the breath of the integrated multimedia system material they learned, the system design skills they acquired, and the software engineering methods they employed. It is a hard and time consuming class, but often this class represents the jumping board for many students to get a unique job offer as a multimedia system designer in the industry which is difficult to get without this class.

Other interesting information that might be useful for others who might want to offer similar Multimedia Systems Design courses

There are several issues to consider when offering this type of course:

1. *What is the makeup of the class in terms of students?* I have to admit that the class projects are getting simpler and will even get simpler as more and more students taking the class are undergraduate students younger than seniors. The reason is that the undergraduate students have much higher load than graduate students. Usually an undergraduate student takes 5-6 courses per semester, where a graduate student has only 2-3 courses per semester. Hence, the undergraduate students do not have the time to work on complex multimedia systems. When I started this course and all of the students were

graduate students (1st year MS/PhD students), the groups built extensive multi-party video conferencing systems and multimedia mailing systems. This is not possible with the current undergraduate class. Even more, as the undergraduate students, taking the class, are at the junior or even sophomore level (after taking the systems programming class), the undergraduate students do not have the experience in building larger scale systems. This means that (a) the instructions for building the systems must be much more detailed and directed (with graduate students, more open-ended projects were possible), and (b) the system complexity must be much lower.

2. *What textbook should the class use?* The multimedia systems field is evolving very fast. This also causes that it is not possible to write up-to-date textbook fast enough, unless one does it full time. I use currently Steinmetz & Nahrstedt, "Multimedia Systems" book from Springer Verlag, published in 2004, but I heavily supplement this book with papers from the premier system and networking conferences and journals. This also means that the slides are more dense than usual since the slides serve as a reading material. Since there is much less trend to write textbook in this area, it would be good to have some written educational material in form of surveys and tutorial modules. This might be an interesting effort to undertake by the SIGMM educational committee. If we split the load and each professor in his/her area prepares a good set of slides as an educational module, this could be used by others and provide much more updated curriculum.
3. *What software, programming language, operating system should the class use?* This is a difficult question since it depends on each educational institution. It also depends on the comfort and background of the students from the particular institution.
   a. It would be good if we would have a database of good libraries/tools for teaching multimedia concepts, as well as strong documentation for the various tools. Few years ago, I used ffmpeg as a tool for getting students started on accessing video driver and also playback of video. At the beginning the ffmpeg documentation was very sparse and that made it very difficult to use this tool. However, ffmpeg requires a very steep learning curve which is not possible for the undergraduate students, hence using easier libraries for multimedia are being deployed (e.g., including Java libraries for multimedia support) currently.
   b. If students are taught Java programming language, this might be a good option for the machine problems. However, with Java, often some of the multimedia system interfaces are hidden and students do not learn as much the internals of multimedia devices and behaviors. Hence, one needs to carefully analyze what concepts to teach via Java in the multimedia systems design. If one uses C/C++ for the machine problems in the course, the students get much better insights into the system issues, but the multimedia tools require much longer learning curve (e.g., if using ffmpeg for MPs).
   c. Similar considerations are with Windows OS vs Linux OS. I use for my MPs Linux since multimedia devices (cameras, microphones, etc) are now having good device drivers. But in the past, drivers on the Linux platform were non-existent (e.g, for cameras). Solaris OS was an excellent platform for teaching multimedia systems, it had elaborate libraries for compression functions, etc. but Solaris labs were replaced with Linux labs and this meant to change the platforms for cs414 machine problems.
4. *What hardware should the class use?* Establishing a laboratory, where computing platforms have multimedia I/O requires additional funds. I was lucky that we had and have a multimedia laboratory where machines are equipped with cameras, microphones, speakers and nice displays. However, as students use the I/O devices, they get faulty. This also means to have a system staff personal to check on the devices, and replace the faulty I/O. The faulty hardware (e.g., sound cards do not work) becomes a sore point for students as they test their MPs. This requires institutional commitment.
5. *What assistance one needs for the course?* Over the years, I felt that a very important part of the overall course was a knowledge-able teaching assistant, who had strong background in building multimedia systems. This also led me to always take a student as a teaching assistant who passed through the course with very strong final results and enjoyed the class. Another important component is to have a strong system staff support of the general hardware and software infrastructure. Assistance is needed with faulty cameras, sound boards, phones, and other computing platforms, and constant help is needed throughout the semester.

Klara Nahrstedt

SIGMM FY'11 Annual Report

Authors: Klara Nahrstedt, SIGMM Chair
URL: http://sigmm.org/news/sigmm-annual_report_fy11
June 2010 - June 2011

Awards

Over the last year 20010-2011, we have given out two SIGMM-wide awards, the SIGMM Technical Achievement Award 2010, and the SIGMM Best PhD Thesis Award 2010.

SIGMM Technical Achievement Award 2010: At the ACM Multimedia 2010, held in Florence Italy, we have presented our 3rd SIGMM Award for Outstanding Technical Contributions to Multimedia Computing, Communications and Applications (shortly Technical Achievement Award) to Professor Ramesh Jain from University of California at Irvine for his technical achievements in the area of multimedia processing, retrieval and applications. The awardee for the SIGMM Technical Achievement Award 2010 was selected by the SIGMM awards committee Dr. Hong-Jiang Zhang (chair of the committee) from Microsoft Research, China, Prof. Rainer Lienhart (SIGMM officer) from University of Augsburg, Germany and Prof. Nicolas Georganas from University of Ottawa.

Note: After 2010, both Dr. Zhang and Prof. Georganas stepped down from the awards committee and for the 2011-2013 period, two new members were asked to serve as the members of the awards committee, Dr. Lawrence Rowe from FXPal, and Prof. Tat-Seng Chua from National University Singapore who graciously agreed to serve on this awards committee.

Prof. Ramesh Jain gave an interesting presentation "Life = Experiences (Events) + Vision" on Wednesday Morning October 27, 2010 and his presentation can be found at http://videolectures.net/acmmm2010_jain_taa/ .

SIGMM Best PhD Thesis Award 2010: At ACM Multimedia 2010, Florence Italy, we have presented our first SIGMM PhD Thesis Award 2010 to Dr. Effrosyni Kokiopoulou, now a postdoctoral fellow at ETH Zurich. She did her PhD at the Signal Processing Laboratory of the Swiss Federal Institute of Technology, Lausanne, Switzerland, under the supervision of Prof. Pascal Frossard. Dr. Kokiopoulou received the SIGMM Best PhD Thesis Award for her outstanding PhD thesis "Geometry-Aware Analysis of High-Dimensional Visual Information Sets" and she gave an interesting presentation on Thursday Morning, October 28, 2010 to and her presentation can be found at http://videolectures.net/acmmm2010_kokiopoulou_gaa/.

The SIGMM Best PhD Thesis Awards committee consisted of Prof. Svetha Venkatesh (chair) from Curtin University of Technology, Australia, Prof. Dick Bulterman from CWI, Netherlands, and Prof. Abed El Saddik from University of Ottawa. All three awards committee members will serve also in the following year 2011-2012 (their final year since we have a three year appointment for this award's committee).

Besides SIGMM-wide awards, our flagship journal, ACM Transactions on Multimedia Computing, Communications and Applications (TOMCCAP) acknowledges the work of the associate editors and gives out the *best TOMCCAP associate editor* award to an editor who provides most excellent services to authors and the community. In 2010, Prof. Shervin Shirmohammadi from University of Ottawa was named the best associate editor of ACM TOMCCAP, second time in a row since he received this award also in 2009. The new editor-in- chief, Prof. Ralf Steinmetz from Technical University of Darmstadt, Germany, presented the award at the banquet of the ACM Multimedia 2010 conference in Florence, Italy.

The editor-in-chief(EiC) of TOMCCAP, Ralf Steinmetz, and the ACM SIGMM chair, Klara Nahrstedt, are currently in process of establishing the *ACM TOMCCAP Nicolas Georganas named award* for best paper published in the journal to honor the memory and outstanding work of Prof. Nicolas Georganas, the founder and first EiC of TOMCCAP, who passed away in July 2010.

At our SIGMM-sponsored conferences, we have given out various conference-specific awards as follows:

1. ACM International Conference on Multimedia (ACM Multimedia) 2010: This is the SIGMM flagship conference, and it was held in Florence, Italy October 25-29, 2010.
   a. *Best Paper Award:* We had four papers that competed for the Best Paper award. This award was given out to Richang Hong, Meng Wang, Mengdi Xu, Shuicheng Yan, Tat-Seng Chua, "Dynamic Captioning: Video Accessibility Enhancement for Hearing Impairment" , a team from School of Computing, National University of Singapore. The presentations of all for papers can be found at http://www.acmmm10.org/.
   b. *Best Student Paper Award:* We have awarded Best student paper award: Nikhil Rasiwasia, Jose Costa Pereira, Emanuele Coviello, Gabe Doyle, Gert Lanckriet, Roger Levy, Nuno Vasconcelos, A New Approach To Cross-Modal Multimedia Retrieval
   c. *Best Demo Award:* Yang Cao, Hai Wang, Changhu Wang, Zhiwei Li, Liqing Zhang, Lei Zhang, MindFinder: Interactive Sketch-based Image Search on Millions of Images and Sabine Susstrunk, Clent Fredembach, Daniel Tamburrino, Automatic Skin Enhancement with Visible and Near-Infrared Image Fusion

d. *Open Source Software Competition 2010:* We are very proud to announce the winners of the ACM Open Source Software Competition 2010:
   i. Andrea Vedaldi, Brian Fulkerson, VLFeat - An open and portable library of computer vision algorithms - VLFeat
   ii. Rob Hess, An Open-Source SIFT Library - Open-Source SIFT
   iii. Florian Eyben, Martin Woellmer, Bjoern Schuller, openSMILE - The Munich Versatile and Fast Open-Source Audio Feature Extractor - openSMILE

e. *Multimedia Grand Challenge 2010:* We had a very interesting and inspiring competition for the multimedia grand challenge. The winners of this grand challenge were
   i. Jana Machajdik, Allan Hanbury, Julian Stöttinger: Understanding Affect in Images.
   ii. Wei Song, Dian Tjondronegoro, Ivan Himawan: ROI-based Content Adaptation for Mobile Device Usage of Video Conferencing.
   iii. Julien Law-To, Gregory Grefenstette, Jean-Luc Gauvain, Guillaume Gravier, Lori Lamel, Julien Despres: Introducing topic segmentation and segmented-based browsing tools into a content based video retrieval system.

2. 2011 ACM International Conference on Multimedia Retrieval (ICMR 2011): This is the first time this conference was held under this name in April 17-20, 2011 in Trento, Italy. This conference emerged from ACM CIVR and ACM MIR conferences, both SIGMM sponsored events, The organizers of these two venues came together and merged the two venues into one excellent and highly visible SIGMM-sponsored venue. This conference awarded best paper and best demonstration awards at their social event.

Significant papers

There were several papers on new areas published in SIGMM-sponsored proceedings:

1. *ACM Multimedia 2010* (http://www.acmmm10.org/conference/general-info/):
   a. We had papers/presentations in the brave new topics session that showed topics such as (1) a strong connection between contextual information such as locations, time-stamps, movement and multimedia; (2) technologies for enriching social situational awareness in remote interactions; (3) social media, (4) video genetics and others.
   b. We had short and long research papers in interactive art, including cultural heritage papers. These papers were connected with a very interesting multimedia art exhibition in Galleria Palazzo Medici Riccardi, Florence, Italy.
   c. We had four major tracks (interfaces, applications, content and systems) and in each track there were interesting sessions/papers.
      i. In content track, major advances have been notices in bringing contextual information into multimedia for classification purposes, annotations, searching, coding, etc.
      ii. In systems track coding on smart phones, 3D streaming and introduction to control new media such as scent, and other topics have been discussed.
      iii. In human-centered track, interesting papers included video accessibility interfaces for hearing impaired users, multimedia interfaces in cars, viewing of photo systems, using immersion for browsing and visualizing surveillance video and others.
      iv. In applications track visual aesthetics played a role for photo-quality assessment, multimedia applications for travel routes planning, using multimedia to inform people about their carbon footprint and others.

2. *ACM NOSSDAV 2010* (http://www.cs.columbia.edu/~hgs/nossdav/2010/): This workshop/working conference is under SIGMM sponsorship and it was held June 2-4, 2010 in Amsterdam, Netherlands.
   - Significant papers continued to come out in the areas of 3D Immersive Interactive Systems, Understanding and Improving User Experiences by the multimedia systems and networking designers, i.e., finding efficient quality assessment methods. Another interesting trend we see in multimedia systems and networks is to understand traffic patterns of social networks, e.g., video/photo sharing, and utilize the understanding in better resource management.

3. *ACM Multimedia Systems 2011* (http://www.mmsys.org/?q=node/41):
   - This year, the organizers invited two categories of papers, the traditional multimedia systems/networking papers, and data set papers. In the category of traditional multimedia systems papers, mobile multimedia and large scale storage and transport for multimedia dominated. The dataset track presented various traces such as network traces of virtual worlds, real-time traffic to residential users, mobile visual search datasets, world of warcraft avatar history dataset and corpus of data for actor level emotion magnitude detection. This was a very nice extension to the program.

4. *ACM ICMR 2011* (http://www.icmr2011.org/):

- The topics in papers that dominated the discussion were in automatic tagging, geo-tagging in video, interaction aspects and social media retrieval.

5. *ACM MM&Sec 2010* (http://www.mmsec10.com/):
   - In this SIGMM-sponsored workshop, held in Rome, Italy, September 9-10, the researchers discussed aspects of biometrics in travel documents, forensics approaches, watermarking and biometric smart card devices, and privacy- preserving approaches.

Significant programs

Throughout the SIGMM-sponsored conferences we had several significant programs that provided a springboard for future technical efforts:

1. *ACM Multimedia 2010:*
   a. Highlight of this conference was again the Multimedia Grand Challenge program organized for the second time as a part of the conference program. Multimedia Grand Challenge is a set of problems and issues from a number of industry leaders geared to engage the multimedia research community in solving relevant, interesting and challenging questions about the industry's 2-5 year horizon for multimedia. Researchers were encouraged to submit working systems in response to the challenge. A large number of submissions were received for this first edition of the competition. We will be continuing with this challenge in ACM Multimedia 2011.
   b. The two keynote talks by Mubarak Shah from University of Central Florida "Visual Crowd Surveillance is Like Hydrodynamics" and Duncan Watts from Yahoo! Research "Using the Web to do Social Science" were very well received.
   c. The multimedia art exhibition in the Medici Palace was outstanding and very well attended.
2. *ACM MMSys 2011:*
   a. There were two important keynote talks talking about next generation challenges for future multimedia systems and network: (a) Alain Fiocco from Cisco and (b) Mark Watson from Netflix that drove a lot of the discussion at the conference and beyond.
   b. A new feature was the dataset track which was very well received by the attendees.
   c. Since this event was held in Cisco company, the attendees saw the telepresence system demonstration as well as the IPTV laboratory in Cisco.
3. *ACM ICMR 2011:*
   a. The conference introduced "Practitioner Day" bringing in industry researchers and discussing problems faced by the practitioners.

Innovative programs

Several SIGMM-sponsored conferences had innovative programs which provided service to technical community:

1. *ACM Multimedia 2010:*
   a. Open Source Competition brings major service to technical community since software is then released to the community with corresponding agreements in place.
   b. SIGMM and Simula sponsored "Women Research Lunch" event which was a lunch for female multimedia researchers to encourage networking of female students with senior female researchers. This lunch also served to give feedback to SIGMM officers what can be done to increase participation of female researchers in multimedia area. We plan a more organized event at ACM Multimedia 2011 for female researchers.
   c. Discussion room was set aside and it was an experiment for making ACM Multimedia 2010 conference even more interactive and stimulate free interactivity beyond the traditional Speaker-Attendees scheme.
2. *ACM MMSys 2011:*
   a. There was a special session on modern media transport "Dynamic Adaptive Steaming over HTTP (DASH)". This session presented novel contributions and breaking results on all aspects of DASH and Modern Media Transport.
   b. Cisco-held conference allowed the MMSys organizers to offer remote participation in the conference via WebEx which was very much used by the community and increased the visibility of the conference.
3. *ACM ICMR 2011:*
   a. The organizers invited attendees to an interesting special session for innovative and frontier topic in the field of multimedia retrieval: "Automatic Tagging and Geo-Tagging in Video Collections and Communities",
   b. Multiple user/industrial sessions happened during the "Practitioner Day" which included industrial talks such as "Attribute-based Object Retrievals" from IBM Research Center, "Combining Analysis with Innovative User Interfaces" from FXPal, "Multimedia Mining for Real-World Applications" from CEA LIST, "Recent Research Activities in KDDI R&D Labs", and "Videntifier Forensic - Automatic Video Identification for Police Authorities" from Videntifier company.

Brief summary for the key issues that the memberships of SIGMM will have to deal with in the next 2 years

The key issues are:

1. Prepare special events to celebrate 20th anniversary of ACM Multimedia conference in 2012 in Nara, Japan.
   - Discussions with the organizing committee of ACM Multimedia 2012 are ongoing.
2. Come up with a sustainable funding model for the multimedia art community within the SIGMM community and their participation at our premier ACM Multimedia conference.
3. Expand SIGMM presence in various social networks.
   - At ACM Multimedia 2010, we had used Twitter to collect opinions, Facebook, the SIGMM website, and other social networks (see website) to increase presence of the conference and other SIGMM events, but more needs to be done.
4. Increase industry participation in SIGMM activities to strengthen ties and increase impact between industry and academia.
   a. ACM ICMR 2011 did a very good job with the "Practitioner Day".
   b. ACM MMSys 2011 brought two keynote speakers from Cisco and NetFlix that provided very strong industrial relevance to the multimedia systems research
5. Automate process for talks content, web, other SIGMM material preservation at SIGMM venues.
   a. ACM Multimedia 2010 did a very nice job preserving the various talks of the conference and putting them on the ACM Multimedia 2010 website.
   b. ACM MMSys 2011 did a very good job collecting all slides of presenters and shared the slides on the conference website.
6. Increase SIGMM participation of female researchers.
   - We have started this task with a lunch at ACM Multimedia 2010.
7. Build up the next generation of SIGMM volunteers to serve as SIGMM officers, chairs, leaders of various SIGMM sponsored activities and venues.
   - We have a very active group of volunteers that drive very diverse activities of ACM SIGMM, but we need to bring new members in.
8. Consider additional SIGMM-wide award(s) to recognize wider multimedia community achievements such as service, education, mid-level research achievements, etc. This will be a discussion point at the next SIGMM business meeting, held at ACM Multimedia 2011, Scottsdale, Arizona.

Other Highlights in SIGMM activities

1. Prof. Mohan Kankanhalli, the SIGMM Director of Conferences, put together a review committee to review the efficiency and organization of our premier ACM Multimedia conference. The chair of the committee was Prof. Tat-Seng Chua. The committee reviewed two aspects of ACM Multimedia and related conferences: (a) the conference organization and (b) the procedures for the management and review of papers for the SIGMM-sponsored conferences. The report is available on the SIGMM website. In January 2011, the review committee came back with a set of key recommendations that are being now implemented for the first time at ACM Multimedia 2011 conference, held in Scottsdale, Arizona.
2. We have put together two SIGMM-sponsored retrieval conferences CIVR and MIR into one outstanding conference, the International Conference on Multimedia Retrieval (ICMR), which was held for the first time in Trento Italy.
3. We had the second SIGMM-sponsored ACM Multimedia Systems conference (MMSys 2011), this year, held in February 2011, San Jose, California. It was held within Cisco company and it was a very successful event with over hundred participants local and remote.
4. We have used our formalized and streamlined process for the ACM Multimedia conference location bidding. This year at ACM Multimedia 2010, we have decided that ACM Multimedia 2012 will be in Nara, Japan, and ACM Multimedia 2013 will be in Barcelona. Spain.
5. Our SIGMM e-newsletter has new articles on multimedia education where various educational activities are featured.
6. We have made significant progress in SIGMM preservation efforts via the preservation committee, led by Dr. Mohamed Hefeeda, who set up a website to preserve past SIGMM-sponsored venues as well as establish processes towards presentation of SIGMM-sponsored venues and their websites, proceedings, etc.
7. The new Editor-in-Chief for ACM Transactions on Multimedia Computing, Communications and Applications, Dr. Ralf Steinmetz, works with ACM towards a new ACM TOMCCAP Nicolas Georganas named award for "Best Paper of the Year" published in ACM Transactions on Multimedia Computing, Communications, and Applications.
8. The SIGMM-specific educational committee compiles and keeps up-to-date educational material in the area of multimedia computing, communications, and applications. This effort is led by Dr. Wei Tsang Ooi. This committee now has an editor on the

SIGMM e-newsletter editorial board to bring articles on multimedia education to a broader community.

9. The SIGMM chapter in China is flourishing. The chapter's own conference, 2nd International Conference on Internet Multimedia Computing and Communication (ICIMCS 2010) had its second event December 30-31, 2010 in Harbin, China, and the 3rd ICIMCS 2011 will be August 5-7, 2011 in Chengdu, China.
10. All SIGMM-sponsored events had a very strong government and industry sponsorship and/or industry participation via talks, papers, demonstrations, including National Science Foundation, and companies such as Microsoft Research, FXPal, Simula, Yahoo!, Google, Cisco, HP Labs, MICC, Manning, O'Reilly, IBM Research, RICOH Innovations, Callas, 3D Life, Technicolor, CEWE, Deutsche Institute - Florence, British Columbia Arts Council, Provincia Di Firenze, Ente Cassa Di Risparmio Di Firenze, and others.
11. SIGMM had a number of conferences/workshops in-cooperation such as the International Conference on Distributed Smart Cameras (ICDSC 2010), International Working Conference on Advanced Visual Interfaces (AVI) 2010, 9th Annual Workshop on Network and Systems Support for Games (NetGames) 2010, and others. Note that 21 workshops have been associated with ACM Multimedia 2010, i.e, co-sponsored and in-cooperation with SIGMM organization.
12. We had the following attendance and paper acceptances for SIGMM-sponsored conferences:
    a. ACM MMSys 2011: attendance was 68 external participants (it was 40 in 2010), 82 Cisco in-room participants, 150 WebEx participants and 184 Cisco TV participants; the acceptance rate for the main track was 37% with 41 papers submitted to main track and 15 papers accepted, for the special media transport track 45% with 20 papers submitted and 5 full papers accepted, and for the new dataset track 42% with 12 papers submitted and 5 accepted.
    b. ACM Multimedia 2010: attendance was 635 participants for the main event, 847 (short and long) papers were submitted and 218 (short and long) papers accepted. We had 357 papers submitted as full papers and 61 papers accepted, which is 17.09% acceptance rate. We had 490 short papers accepted and 157 short papers accepted which is 32.04% acceptance rate. We had 58 papers submitted to interactive art program and the acceptance rate was 29.31% with 16 papers accepted. We had 10 video submissions, 67 technical demonstration submissions and 22 industrial exhibition submissions. The acceptance was 8 videos accepted, 43 technical demonstrations, and 19 industrial exhibits were accepted.
    c. ACM ICMR 2011: acceptance rate for oral presentations (long papers) was 16.4%, acceptance rate for poster presentations (short papers) was 35%.
    d. ACM NOSSDAV 2010: attendance was 47 attendees, 51 papers submitted and 21 papers accepted, which is 41.4%.
    e. ACM MM&Sec 2010: attendance was 52 attendees, 53 papers submitted and 32 accepted, which is 60.3%.

Top 10 ACM SIGMM Downloads

by

Here we present the top downloaded ACM SIGMM articles from the ACM Digital Library, from July 2010 to June 2011. We are hoping that this list gives a much deserved exposure to the ACM SIGMM's best articles.

1. Guo-Jun Qi, Xian-Sheng Hua, Yong Rui, Jinhui Tang, Tao Mei, Meng Wang, Hong-Jiang Zhang. *Correlative multilabel video annotation with temporal kernels.* In *ACM Trans. Multimedia Comput. Commun. Appl.* 5(1), 2008
2. Michael S. Lew, Nicu Sebe, Chabane Djeraba, and Ramesh Jain. *Content-based multimedia information retrieval: State of the art and challenges.* In *ACM Trans. Multimedia Comput. Commun. Appl.* 2(1), 2006
3. Ba Tu Truong, Svetha Venkatesh. *Video abstraction: A systematic review and classification.* In *ACM Trans. Multimedia Comput. Commun. Appl.* 3(1), 2007
4. Yu-Fei Ma, Hong-Jiang Zhang. *Contrast-based image attention analysis by using fuzzy growing.* In *ACM Multimedia 2003*
5. Simon Tong and Edward Chang. *Support vector machine active learning for image retrieval.* In *ACM Multimedia 2001*
6. J.-P. Courtiat, R. Cruz de Oliveira, L. F. Rust da Costa Carmo. *Towards a new multimedia synchronization mechanism and its formal definition.* In *ACM Multimedia 1994*
7. Gabriel Takacs, Vijay Chandrasekhar, Natasha Gelfand, Yingen Xiong, Wei-Chao Chen, Thanos Bismpigiannis, Radek Grzeszczuk, Kari Pulli, Bernd Girod. *Outdoors augmented reality on mobile phone using loxel-based visual feature organization.* In *ACM SIGMM MIR 2008*
8. Jiajun Bu, Shulong Tan, Chun Chen, Can Wang, Hao Wu, Lijun Zhang, Xiaofei He. *Music recommendation*

by unified hypergraph: combining social media information and music content. In *ACM Multimedia 2010*

9. Hina Keval, M. Angela Sasse. *To catch a thief -- you need at least 8 frames per second: the impact of frame rates on user performance in a CCTV detection task.* In *ACM Multimedia 2008*
10. Mathias Lux, Savvas A. Chatzichristofis. *Lire: lucene image retrieval: an extensible java CBIR library.* In *ACM Multimedia 2008*

PhD thesis abstracts

Doreen Boehnstedt

Semantic Tagging for Managing Web-based Learning Resources: Models, Methods and a Plattform for Supporting Resource-based Learning

The knowledge explosion, changing circumstances due to new forms of work and many technical developments determine that the knowledge acquired in education is not sufficient throughout life. Therefore, self-directed learning in the workplace is becoming increasingly important. This is a form of learning where a current information need is met by the self-directed interaction with a wide range of digital resources. Therefore, this learning is called Resource-based Learning. Increasingly, the importance of the Web as an information source grows because it provides many resources that can be used for learning purposes. However, self-directed Resource-based Learning also poses many challenges to learners. First, digital resources on the Web are usually not didactically prepared and therefore are not intended to be used as learning materials. In addition, the relevant information is often distributed across many different websites. Further, there is already a very large but still rapidly increasing amount of information available on the Web, which can lead to information overload. In the scenario of self-directed learning considered here, there is no teacher who structures the learning process. Therefore, learners have to independently determine their information needs and plan their proceeding. They have to identify, annotate and organize relevant resources for future use. This makes an appropriate management of resources necessary. However, the majority of learners is unsatisfied with the currently available possibilities for the organization of Web resources. The goal of this thesis is therefore the design and development of a tool to support learners in Resource-based Learning. In particular, the management of resources should be supported and hence challenges mentioned above are addressed. Self-directed Resource-based Learning requires a personal information and knowledge management by the learners. In literature, several models for managing information and knowledge in organizational and personal scopes exist. For self-directed Resource-based Learning such a model is missing so far. Therefore, a model for Resource-based Learning is developed based on the existing models and on a questionnaire survey conducted in the context of this thesis. This model encompasses several process steps that should be supported by the tool. The management of resources necessitates the learners to appropriately store the resources, such as based on topic of interest or task to be executed. Tagging is a simple and accepted way to manage any resource on the Web, but its power of expression is restricted. Other forms of resource management can be found in the area of formal knowledge organization (e.g. modeling of a semantic network), however, expert knowledge is usually required to build a semantic network. As a basis for the tool that is developed in the context of this work, therefore, a combination of both forms is proposed, i.e. a semantic network that is created and expanded by the learners using tagging. Core components of this network are resources and tags. Additionally, a learner is able to assign a type to each tag. Therefore, the information whether the tag is e.g. a topic or task can be stored. As part of this thesis, the types of tags that are necessary for the scenario of Resource-based Learning have been analyzed and evaluated. Furthermore, an algorithm for automatic detection of these tag types is presented, as such an algorithm can reduce the manual maintenance effort for the management of resources. The evaluation of various corpora shows that the knowledge-based algorithm can classify a tag already

during the tagging process with an accuracy which is sufficient for the scenario. Based on the developed model of Resource-based Learning and its requirements for the management of resources, different tools and systems are analyzed with regard to their support of Resource-based Learning. None of the related tools fulfill the requirements appropriately. Therefore, on the basis of the model's process steps and the derived functional requirements a concept for a supporting tool is developed. Based on the technical requirements, a system is designed, consisting of a browser add-on, a backend for the management of the knowledge networks and a web-based frontend. The tool is implemented and evaluated in user studies eventually. The user studies conducted in this work show that the extended form of tagging, based on tag types, is well accepted and allows for appropriate management of resources. Furthermore, the studies show that the implemented tool addresses the challenges of self-directed Resource-based Learning adequately. The present work thus creates a basis for optimizing the approach to self-directed interaction with resources in order to meet an information need.

Advisor(s): Ralf Steinmetz, Ulrik Schroeder
SIG MM member(s): Ralf Steinmetz
http://tuprints.ulb.tu-darmstadt.de/2729/

Multimedia Communications Lab

http://www.kom.tu-darmstadt.de/

Jun Wang

Semi-Supervised Learning for Scalable and Robust Visual Search

Unlike textual document retrieval, searching of visual data is still far from satisfactory. There exist major gaps between the available solutions and practical needs in both accuracy and computational cost. This thesis aims at the development of robust and scalable solutions for visual search and retrieval. Specifically, we investigate two classes of approaches: graph-based semi-supervised learning and hashing techniques. The graph-based approaches are used to improve accuracy, while hashing approaches are used to improve efficiency and cope with large-scale applications. A common theme shared between these two subareas of our work is the focus on semi-supervised learning paradigm, in which a small set of labeled data is complemented with large unlabeled datasets.

Graph-based approaches have emerged as methods of choice for general semi-supervised tasks when no parametric information is available about the data distribution. It treats both labeled and unlabeled samples as vertices in a graph and then instantiates pairwise edges between these vertices to capture affinity between the corresponding samples. A quadratic regularization framework has been widely used for label prediction over such graphs. However, most of the existing graph-based semi-supervised learning methods are sensitive to the graph construction process and the initial labels. We propose a new bivariate graph transduction formulation and an efficient solution via an alternating minimization procedure. Based on this bivariate framework, we also develop new methods to filter unreliable and noisy labels. Extensive experiments over diverse benchmark datasets demonstrate the superior performance of our proposed methods. However, graph-based approaches suffer from the critical bottleneck in scalability since graph construction requires a quadratic complexity and the inference procedure costs even more. The widely used graph construction method relies on nearest neighbor search, which is prohibitive for large-scale applications. In addition, most large-scale visual search problems involve handling high-dimensional visual descriptors, thereby causing another challenge in excessive storage requirement. To handle the scalability issue of both computation and storage, the second part of the thesis focuses on efficient techniques for conducting approximate nearest neighbor (ANN) search, which is key to many machine learning algorithms, including graph-based semi-supervised learning and clustering. Specifically, we propose Semi-Supervised Hashing (SSH) methods that leverage semantic similarity over a small set of labeled data while preventing overfitting. We derive a rigorous formulation in which a supervised term minimizes the empirical errors on the labeled data and an unsupervised term provides effective regularization by maximizing variance and

independence of individual bits. Experiments on several large datasets demonstrate the clear performance gain over several state-of-the-art methods without significant increase of the computational cost. The main contributions of the thesis include the following. 1) a bivariate formulation for graph-based semi-supervised learning with an efficient solution by alternating optimization; b) theoretic analysis from the view of graph cut for the bivariate optimization procedure; c) novel applications of the proposed techniques, such as interactive image retrieval, automatic re-ranking for text based image search, and a brain computer interface (BCI) for image retrieval. 2) a rigorous semi-supervised paradigm for hash functions learning with a tradeoff between empirical fitness on pair-wise label consistence and an information-theoretic regularizer; b) several efficient solutions for deriving semi-supervised hash functions, including an orthogonal solution using eigen-decomposition, a revised strategy for learning non-orthogonal hash functions, a sequential learning algorithm to derive boosted hash functions, and an extension to unsupervised cases by using pseudo labels. Two parts of the thesis - bivariate graph transduction and semi-supervised hashing - are complimentary and can be combined to achieve significant performance improvement in both speed and accuracy. Hash methods can help build sparse graphs in a linear time fashion and greatly reduce the data size, but they lack sufficient accuracy. Graph-based methods provide unique capabilities to handle non-linear data structures with noisy labels but suffer from high computational complexity. The synergistic combination of the two offers great potential for advancing the state-of-the-art in large-scale visual search and many other applications.

Advisor(s): Shih-Fu Chang
SIG MM member(s): Shih-Fu Chang
http://www.ee.columbia.edu/ln/dvmm/publications/PhD_theses/jwang_thesis.pdf

DVMM Lab

http://www.ee.columbia.edu/ln/dvmm/

The DVMM Lab at Columbia University is dedicated to research of new theories, algorithms, and systems for multimedia content analysis, search, communication, and forensics, with a primary focus on digital video. It hosts faculty, students, and visiting researchers, conducting research as well as development of multimedia technologies, testbeds, and standards.
Our current research activities focus on five areas: multimedia search and retrieval, pervasive media and mobile communication, machine learning and object recognition, media security and forensics, multimedia standard, testbed, and benchmarking.

Philipp Scholl

Semantic and Structural Analysis of Web-based Learning Resources - Supporting Self-directed Resource-based Learning

In the knowledge-based society, the maintenance and acquisition of new knowledge are vital for each individual. Changed living and working conditions and the rapid development of technology cause the half-life of knowledge to decrease. Therefore, the knowledge that is acquired in educational institutions is no longer sufficient for an entire lifetime. Thus, self-directed learning at the workplace and in private life is becoming more and more important. At the same time, the Web has become a very important source for knowledge acquisition, as it provides a huge amount of resources containing information that can be utilized for learning purposes. This form of self-directed learning that often involves learning with web resources is commonly referred to as Resource-Based Learning. In particular, it is characterized by a high degree of freedom in choice of resources and execution of the learning process. When utilizing web resources as learning materials, learners face novel challenges: First, relevant information that covers the specific information need of a learner is often distributed over several web resources. This challenge can be addressed by providing adequate retrieval strategies where retrieval is not only restricted

to a web search but also involves content that learners have already considered to be relevant. However, the so-called vocabulary gap - the fact that information can be expressed in completely different terminology, e.g. in technical terms or colloquial language - makes retrieval difficult. Further, in contrast to Learning Objects that are often used in educational institutions, web resources rarely include well-structured metadata. As Resource-Based Learning using web resources requires learners to handle and organize a large number of web resources efficiently, the availability of relevant metadata is vital. Eventually, in the majority of self-directed learning settings, the role of the teacher or tutor does not exist. These authorities usually set learning goals according to a curriculum, structure the learning process and assess the learning result. In self-directed learning, the learner has to take over these tasks which would otherwise have been accomplished by the teacher. This thesis examines this form of Resource-Based Learning and derives adequate mechanisms to support this kind of learning. The requirements of supporting Resource-Based Learning are deduced and, based on these requirements, the design and the implementation of a tool called ELWMS.KOM is presented. ELWMS.KOM is a tool that enables learners to organize their self-directed learning process and the contributing learning resources in a personal knowledge network by applying semantically typed tags. In particular, web resources are focused. Web resources are primarily not intended to be used for learning and thus, are rarely didactically adapted to learning scenarios. Further, they infrequently expose metadata that are relevant for learners. ELWMS.KOM is designed to attenuate these short-comings and the resulting challenges for learners by providing an appropriate level of support. The contributions of this thesis comprise of the derivation and implementation of paradigms and technologies that enable such a supporting functionality in ELWMS.KOM. Based on an examination of Learning Objects that are commonly used in learning scenarios in educational institutions, the peculiarities and differences to self-directed learning paradigms are analysed and design decisions for ELWMS.KOM are inferred. These design decisions represent a foundation for the supporting functionalities that are proposed in this thesis. Firstly, the technologies are presented that enable ELWMS.KOM to recommend tags and learning resources to the learner based on a semantic representation of their content. A user study based on ELWMS.KOM shows the need to support monolingual as well as cross-lingual approaches to recommend semantically related tags and resources. An analysis of the approach that has been chosen to determine semantic relatedness is presented. Based on this analysis, several strategies are compared that show potential to reduce the computational complexity of this approach without considerably reducing its quality. Additionally, several extensions to improve the quality this approach that incorporate supplementary semantic properties of a reference corpus are presented and evaluated. Furthermore, this thesis presents an approach to automatically segment web resources in order to support learners in the selection of relevant fragments of a web resource. This segmentation is based on a structural and visual analysis of web resources and yields a set of coherent segments. A user study confirms the quality of this approach. In addition, an approach is introduced that supports learners in the consistent creation of their tagging vocabulary in ELWMS.KOM for the semantic tag type Type. This approach automatically recognizes the web genre of a web resource and is language-independent. Novel features have been developed that allow a reliable classification of web genres. Several evaluations using different feature sets and corpora are presented. Finally, this thesis introduces the tag type Goal that supports learners to plan, execute and evaluate their overall learning process. This support feature has been derived from the theory of Self-Regulated Learning and has been implemented accordingly in ELWMS.KOM. The benefits are shown in two large-scale user studies that have been executed with ELWMS.KOM and the implemented goal setting mechanisms.

Advisor(s): Ralf Steinmetz, Wolfgang Effelsberg
SIG MM member(s): Ralf Steinmetz, Wolfgang Effelsberg
http://tuprints.ulb.tu-darmstadt.de/2644/

Multimedia Communications Lab

http://www.kom.tu-darmstadt.de/

Wanmin Wu

Human-centric Control of Video Functions and Underlying Resources in 3D Tele-immersive Systems

3D tele-immersion (3DTI) has the potential of enabling virtual-reality-like interaction among remote people with real-time 3D video. However, today's 3DTI systems still suffer from various performance issues, limiting their broader deployment, due to the enormous demand on temporal (computing) and spatial (networking) resources. Past research focused on system-centric approaches for technical optimization, without taking human users into the loop. We argue that human factors (including user preferences, semantics, limitations, etc.) are an important and integral part of the cyber-physical 3DTI systems, and should not be neglected.
This thesis proposes a novel, comprehensive, human-centric framework for improving the qualities of 3DTI throughout its video function pipeline. We make three major contributions at different phases of the pipeline. At the sending side, we develop an intra-stream data adaptation scheme that reduces level-of-details within each stream without users being aware of it. This human-centric approach exploits limitations of human vision, and excludes details that are imperceptible. It effectively alleviates the data load for computation-intensive operations, thus improves the temporal efficiency of the systems. Yet even with intra-stream data reduced, spatial efficiency is still a problem due to the multi-stream/multi-site nature of 3DTI collaboration. We thus develop an inter-stream data adaptation scheme at the networking phase to reduce the number of streams with minimal disruption to the visual quality. This human-centric approach prioritizes streams based on user views and excludes less important streams from transmission. It considerably reduces the data load for networking, and thus enhances the spatial resource efficiency. The above two approaches (level-of-details reduction within a video stream and view-based differentiation among streams) work seamlessly together to bring both temporal and spatial resource demands under control, and prove to improve various qualities of the systems. Finally, at the receiving side, we take a holistic approach to study the ``quality'' concept in 3DTI environments. Our human-centric quality framework focuses on the Quality-of-Experience (QoE) concept that models user's perceptions, emotions, performances, etc. It investigates how the traditional Quality-of-Service (QoS) impacts QoE, and reveals how QoS should be improved for the best user experience. This thesis essentially demonstrates the importance of bringing human-awareness into the design, execution, and evaluation of the complex resource-constrained 3DTI environments.

Advisor(s): Klara Nahrstedt (Advisor)
SIG MM member(s): Wanmin Wu
https://www.ideals.illinois.edu/handle/2142/26126

Multimedia Operating Systems and Networking (MONET)

http://monet.web.cs.illinois.edu/

Research in the MONET research group focuses on system software issues to provide services and protocols for end-to-end Quality of Service (QoS) guarantees for distributed multimedia applications, leveraging the best effort services provided by the underlying operating system and networks. Toward this goal, we are doing research in a broad area including (but not limited to):
- Multimedia operating systems - Multimedia communication protocols - QoS middleware and large scale distributed systems - Multimedia security and trustworthy computing systems - Advanced tele-immersive and multimedia applications - High speed QoS routing and ad hoc networks

Calls for contributions

Calls for SIGMM Sponsored and Co-sponsored Events

ACM TOMCCAP Special Issue on 3D Mobile Multimedia

Full paper Deadline: January 15 2012
URL: http://www.site.uottawa.ca/~shervin/3dmmsi/

Despite the promising nature of 3D media, significant challenges remain to be solved in order to bring them to mobile devices: limited battery, limited bandwidth, higher network latency, limited processing and display, and heterogeneity. In this special issue, we solicit papers about supporting interactive 3D multimedia in a mobile setting.

Calls for Events held in cooperation with SIGMM

ACM International Conference on Multimedia Retrieval (ICMR)

Full paper Deadline: Jan. 15, 2012
Event location: Hong Kong
Event date: June 5 - 8, 2012
URL: http://www.icmr2012.org/

ICMR 2012 is seeking original high quality submissions addressing innovative research in the broad field of multimedia retrieval.

ACM Workshop on Mobile Video (MoVid) (in conjunction with ACM Multimedia Systems)

Full paper Deadline: October 20, 2011
Event location: Chapel Hill, NC, USA
Event date: Feb. 24, 2012
URL: http://www.eecs.ucf.edu/movid/

MoVid solicits original and unpublished research achievements in various aspects of mobile video delivery.

Other multimedia-related Events

International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM)

Full paper Deadline: December 2, 2011
Event location: San Francisco, CA, USA
Event date: June 25-28, 2012
URL: http://wowmom2012.it.uc3m.es/

WoWMoM is addressing research challenges and advances towards a world of wireless, mobile, and multimedia pervasive communications.

International World Wide Web Conference (WWW)

Full paper Deadline: November 7th, 2011
Event location: Lyon, France
Event date: April 16-20, 2012
URL: http://www2012.org/

WWW is the premier venue for academics and industry to present, demonstrate, and discuss the latest ideas about the Web, its infrastructure, relevant algorithms and new innovative applications.

IEEE International Conference on Multimedia and Expo (ICME)

Full paper Deadline: November 28, 2011
Event location: Melbourne, Australia
Event date: July 9-13, 2012
URL: http://www.icme2012.org/

ICME serves as a forum to promote the exchange of the latest advances in multimedia technologies, systems, and applications from both the research and development perspectives of the circuits and systems, communications, computer, and signal processing communities.

International Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS)

Full paper Deadline: Feb. 3, 2012
Event location: Dublin, Ireland
Event date: May 23-25, 2012
URL: http://wiamis.dcu.ie/

The objective of the workshop is to bring together researchers and developers from academia and industry working in the areas of image, video and audio applications, with a special focus on analysis.

EURASIP Journal on Advances in Signal Processing Special Issue On Social Media Processing and Semantic Modeling

Full paper Deadline: October 15, 2011
Event date: April 15, 2012
URL: http://asp.eurasipjournals.com/

Collaborative image/video tagging has become very popular for people to share, tag and search images/ videos. The topics for this special issue will include,

but are not limited to: Advanced semantic models for social media processing, especially for multimedia social media data; Social media computation and applications on advanced semantic models, such as clustering, reasoning and retrieval; Novel semantic models for non-traditional signals, such as touch models for haptic devices, gesture models for touch screens, and 3D object models; Automatic extraction algorithms for semantic models, either model driven or data driven; Computation algorithms and infrastructures for the problems of model extraction and applications

Springer MMSJ Special Issue on Network and Systems Support for Games

Full paper Deadline: November 15 2011
URL: http://www.site.uottawa.ca/~shervin/netgamessi/

As online games become more popular, research needs to be devoted to manage and support their emerging massiveness, their traffic, and the users' quality of experience. In this special issue, we focus on scientific, engineering, and research topics in systems and networking support to enable multiplayer networked games.

Call For Special Session: IEEE International Conference on Multimedia & Expo (ICME) 2012

Full paper Deadline: October 15, 2011
Event location: Melbourne, Australia
Event date: 9th - 13th July, 2012
URL: http://www.icme2012.org/CallForPapers_SpecialSession.php

Special sessions supplement the regular program for ICME 2012. They are intended to cover new and emerging topics in the fields of multimedia. Each special session should provide an overview of the state-of-the-art and highlight important research directions in a field of special interest to ICME participants.

Job Opportunities

Speech technology: one postdoc and one PhD position

Employer: University of Eastern Finland, Joensuu, Finland

Valid until: Friday November 4, 2011
More info: http://cs.joensuu.fi/sipu/UEF_SIPU_speechtech_job_announcement.pdf

One postdoctoral researcher and one PhD position available in our research group. The topics include speaker recognition and voice conversion. The applicants should have background in signal processing, machine learning or spoken language technologies.

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
Technical assistance	Preben Olsen, University of Oslo