Volume 3, Number 2
June 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 2, June 2011

Editorial

Dear Member of the SIGMM Community,

Welcome to the second issue of the SIGMM Records in 2011!

Learn about five newly concluded PhD theses in the multimedia area. Summaries of the work of five young researchers are included in this issue of the Records to raise your interest, and to point you to the right to look for further information.

You can stay up to date with the MediaEval, the benchmarking initiative by PetaMedia, which has published new tasks that will be answered in September 2011. MMSys 2011, on the other hand, has already been held in February. You can read about the experience in this issue of the records, and follow the direct links to papers in the ACM digital library. You can of course also find links for the latest issues of TOMCCAP and MMSJ.

The Educational Column presents you a new and exciting book that is meant as an introductory textbook into multimedia computing.

You find a summary of the last MPEG meeting, which is meant as an introduction to our new column on the developments in MPEG, which will feature in future issues of the Records.

If you intend to host ACM Multimedia in North America in 2014, you should take note of the call for bids. The requirements and links to the relevant documents are presented in this issue.

We can also report that one of our colleagues has been granted an IEEE I&M Technical Award, and you can learn about several other opportunities for your research or career.

The Editors
Stephan Kopf
Viktor Wendel
Lei Zhang
Wei Tsang Ooi
Carsten Griwodz

Awards for SIGMM members

IEEE I&M Technical Award

Professor El Saddik was awarded the IEEE Instrumentation and Measurement Society technical Award for his outstanding contributions to multimedia computing

Dr. El Saddik's research focus is establishing a framework of software tools to define an ambient-based collaborative haptic audio visual environment (C-HAVE). Thereby making the exchange of data in that environment fully interactive and personalized, allowing users on the Internet to experience an enhanced telepresence, made up of images, sounds, and touch.

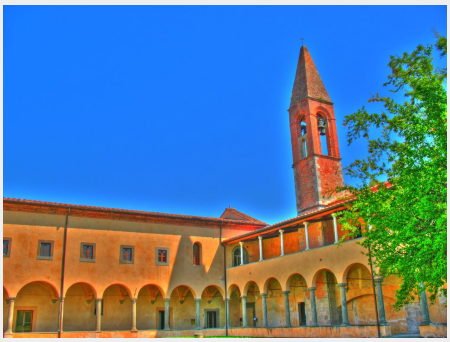
The I&M Society presented its Society Awards at I2MTC 2011 in Binjiang, Hangzhou, China.

MediaEval 2011

Authors: Martha Larson
URL: http://www.multimediaeval.org

by Martha Larson

MediaEval (http://www.multimediaeval.org/) is a benchmarking initiative that offers tasks promoting research and innovation on multimodal approaches to multimedia annotation and retrieval. Its focus is on speech, language, context and social aspects of multimedia, in addition to visual content. The MediaEval 2011 benchmarking season culminates with the MediaEval 2011 workshop, held on 1&2 September MediaEval 2011 Workshop at Santa Croce in Fossabanda, Pisa, Italy. The workshop is an official satellite event of Interspeech 2011 (http://www.interspeech2011.org), the 12th Annual Conference of the International Speech Communication Association (ISCA).



Santa Croce (photo: Flickr User Marius B)

Currently, the 2011 season of MediaEval is under way. For each task, participants receive a task definition, task data and accompanying resources (dependent on task) such as shot boundaries, keyframes, visual features, speech transcripts and social metadata. Participants are tackling the following tasks in the 2011 MediaEval season:

Genre Tagging Given a set of genre tags (how-to, interview, review etc.) and a video collection, participants are required to automatically assign genre tags to each video based on a combination of modalities, i.e., speech, metadata, audio and visual (Data: Creative Commons internet video, multiple languages mostly English)

Rich Speech Retrieval Given a set of queries and a video collection, participants are required to automatically identify relevant jump-in points into the video based on the combination of modalities, i.e., speech, metadata, audio and visual. The task can be approached as a multimodal task, but also as strictly a searching speech task. (Data: Creative Commons internet video, multiple languages mostly English)

Spoken Web Search This task involves searching FOR audio content WITHIN audio content USING an audio content query. It is particularly interesting for speech researchers in the area of spoken term detection. (Data: Audio from four different Indian languages -- English, Hindi, Gujarati and Telugu. Each of the ca. 400 data item is an 8 KHz audio file 4-30 secs in length)

Affect Task: Violent Scenes Detection This task requires participants to deploy multimodal features to automatically detect portions of movies containing violent material. Any features automatically extracted from the video, including the subtitles, can be used by participants. (Data: A set of ca. 15 Hollywood movies that must be purchased by the participants)

Social Event Detection This task requires participants to discover events and detect media items that are related to either a specific social event or an event-class of interest. By social events we mean that the events are planned by people, attended by people and that the social media are captured by people. (Data: A large set of URLs of videos and images together with their associated metadata)

Placing Task This task involves automatically assigning geo-coordinates to Flickr videos using one or more of: Flickr metadata, visual content, audio content, social information (Data: Creative Commons Flickr data, predominantly English language)

MediaEval is coordinated by the EU FP7 PetaMedia Network of Excellence http://www.petamedia.eu and also by the ICT Labs of EIT http://eit.ictlabs.eu/ and is made possible by the many projects, institutions and researchers that contribute to the organization of the individual tasks.

For more information on MediaEval, please contact Martha Larson m.a.larson@tudelft.nl.

SIGMM Education Column

Authors: Wei Tsang Ooi
URL: http://www.sigmm.org/Education
by Wei Tsang Ooi

In this issue's SIGMM Education Column, we highlight a new textbook on multimedia computing, written by two active members of the SIGMM community, Ramesh Jain and Gerald Friedland. The textbook, titled

``Introduction to Multimedia Computing,'' is targeted at senior undergraduate and beginning graduate students. It is scheduled to appear in 2011 and is published by Cambridge University Press.

``Introduction to Multimedia Computing'' adopts a comprehensive approach in presenting the field of multimedia computing. Instead of introducing multimedia computing based on the media types (image, audio, video), this book considers multimedia as a fundamental and unique discipline that must use all media necessary to solve problems and presents a unified introduction to the field by focusing on the fundamental techniques and mathematical foundations of multimedia computing. For instance, the book introduces lossy compression by presenting the principles of quantization and differential coding, before showing how these principles are applied to compression of image, audio, and video.

Even though the theoretical foundations are very similar, multimedia researchers and developers are usually divided along specific medium boundaries, such as speech processing, natural language processing or computer vision. The ways humans process information as well as the processing power of modern computers, however, suggests that by handling different media synergistically, one may understand even an individual media better and may design more practical engineering systems. Integrated processing promises improved robustness in many situations and is closer to what humans do. This textbook provides such a unified perspective to multimedia computing for computer scientists. The authors hope that, people interested in multimedia will rise above the media boundaries and will worry more about the content than the medium.

Of course, a field as broad as multimedia cannot be covered comprehensively in one book as it would require covering large parts of mathematics, physics, physiology, psychology, electrical engineering, in addition to computer science topics. The authors therefore adopt a different approach: They introduce the main concepts in a ``capsule'' form and provide pseudo-code for algorithms such as ADPCM encoding or the Paeth Predictor. Very often, chapters point to unexplored possibilities and the historical reasons for them (e.g., patent quarrels), encouraging the student to experiment on their own. For example, given the Weber-Fechner law of logarithmic perception, why are the TV formats encoding colors and brightness linearly? Further literature and web references are provided, allowing the reader to explore the topics beyond the coverage of the book.

Topics covered in this textbook include fundamentals of sound and light, compression, authoring, content analysis, retrieval, HCI, as well as privacy and security. The authors of the textbook have made draft chapters of the book available online for comments. The chapters can be downloaded from http://www.mm-creole.org.

In the future, the website will provide additions to the book chapters, which will dynamically evolve as the field progresses, support for the exercises, as well as further and updated information on the community.

Call for Bids: ACM Multimedia 2014

Authors: Mohan Kankanhalli
by Mohan Kankanhallo (National University of Singapore)

Required Bid Documents:

Two documents are required:

1. Bid Proposal: This document outlines all of the details except the budget. The proposal should contain:
   a. The organizing team: Names and brief bios of General Chairs, Program Chairs and Local Arrangements Chairs. Names and brief bios of at least one chair (out of the two) each for Workshops, Panels, Video, Brave New Ideas, Interactive Arts, Open Source Software Competition, Multimedia Grand Challenge, Tutorials, Doctoral Symposium, Preservation and Technical Demos.

      It is the responsibility of the General Chairs to obtain consent of all of the proposed team members.

      Please note that the SIGMM Executive Committee may suggest changes in the team composition for the winning bids.
   b. The Venue: the details of the proposed conference venue including the location, layout and facilities. The layout should facilitate maximum interaction between the participants. It should provide for the normal required facilities for multimedia presentations including internet access.

      Please note that the 2014 ACM Multimedia Conference will be held in North America.
   c. Accommodation: the bids should indicate a range of accommodations catering for student, academic and industry attendees with easy as well as quick access to the conference venue. Indicative costs should be provided. Indicative figures for lunches/ dinners and local transport costs for the location must be provided.

   d. Accessibility: the venue should be easily accessible to participants from Americas, Europe and Asia (the primary sources of attendees). Indicative cost of travel from these major destinations should be provided.
   e. Other aspects:
      i. commitments from the local government and organizations
      ii. committed financial and in-kind sponsorships
      iii. institutional support for local arrangement chairs
      iv. conference date in October/November which does not clash with any major holidays or other major related conferences
      v. social events to be held with the conference
      vi. possible venue(s) for the TPC Meeting
      vii. any innovations to be brought into the conference
      viii. cultural/scenic/industrial attractions
2. Tentative Budget: The entire cost of holding the conference with realistic estimated figures should be provided. This template budget sheet should be used for this purpose:

   http://www.comp.nus.edu.sg/~mohan/ACMMM_Budget_Template.xls

   Please note that the sheet is quite detailed and you may not have all of the information. Please try to fill it as much as possible. All committed sponsorships for conference organization, meals, student subsidy and awards must be highlighted.

   Please note that estimated registration costs for ACM members, non-members and students will be required for preparing the budget.

   Estimates of the number of attendees will also be required.

Bid Evaluation Procedure:

Bids will be evaluated on the basis of:

1. Quality of the Organizing Team (both technical strengths and conference organization experience)
2. Quality of the Venue (facilities and accessibility)
3. Affordability of the Venue (travel, stay and registration) to the participants
4. Viability of the Budget: Since SIGMM fully sponsors this conference and it does not have reserves, the aim is to minimize the probability of making a loss and maximize the chances of making a small surplus.

The winning bid will be decided by the SIGMM Executive Committee by vote.

Bid Submission Procedure:

Please up-load the two required documents and any other supplementary material to a web-site. The general chairs then should email the formal intent to host along with the bid documents web-site URL to the SIGMM Chair (klara@cs.uiuc.edu) and the Director of Conferences (mohan@comp.nus.edu.sg) by Oct 1, 2011.

Time-line:

Oct 01, 2011	Bid URL to be submitted to SIGMM Chair and Director of Conferences
Oct 2011	Bids open for viewing by SIGMM Executive Committee and Feedback about any missing information to Bidders
Nov 15, 2011	Bids open for viewing by all SIGMM Members
Nov 29, 2011	10-min Presentation of each Bid at ACM Multimedia 2011
Nov 30, 2011	Decision by the SIGMM Executive Committee

Please note that there is a separate conference organization procedure which kicks in for the winning bids whose details can be seen at: http://www.acm.org/sigs/volunteer_resources/conference_manual

MPEG Internet Video Coding

Authors: Christian Timmerer
URL: http://multimediacommunication.blogspot.com/search/label/MPEG
by Christian Timmerer

At its 96th meeting, MPEG issued two document in the area of Internet video coding which are publicly available:

- Draft Call for Proposals (CfP) for Internet Video Coding Technologies
- Requirements for Internet Video Coding Technologies

Draft Call for Proposals (CfP) for Internet Video Coding Technologies

The draft CfP comprises requirements for a proposal (i.e., on the actual submission), information on the

evaluation, source code & IPR details, and the timeline. In particular, the aim of this work item is to address the diversified needs of the Internet:

- To satisfy the requirements of this application domain MPEG will evaluate the submissions and will develop a specification (the Standard) that MPEG expects shall include a profile qualified as a "Option-1 licensing" and may include other profiles.

The timeline for the Call for Proposal is as follows:

- Draft Call-for-Proposals ready: 2011/03
- Final Call-for-Proposals issued: 2011/07
- Proposals received and evaluation starts: 2011/10

Option-1 Codec specification development plan:

- Committee Draft: 2012/07
- Draft International Standard: 2013/01
- Final Draft International Standard: 2013/07

Requirements for Internet Video Coding Technologies

Requirements fall into the following major categories:

- IPR requirements
- Technical requirements
- Implementation complexity requirements

Interestingly, the standard shall provide better compression performance than MPEG-2 and possibly comparable to AVC baseline profile. The resolution shall be from QVGA to HD and various color spaces, color sampling, and bit-depth coding shall be supported. Other technical requirements include (the usual ones) high perceptual quality, random access, support for trick modes, network friendliness, error resilience, video buffer management, bitstream scalability, transcoding, and overlay channel. Finally, the implementation complexity shall allow for real-time encoding and decoding on both stationary and mobile devices.

PhD thesis abstracts

Dinesh Babu Jayagopi

Computational Modeling of Face-to-Face Social Interaction

The computational modeling of face-to-face interactions using nonverbal behavioral cues is an emerging and relevant problem in social computing. Studying face-to-face interactions in small groups helps in understanding the basic processes of individual and group behavior; and improving team productivity and satisfaction in the modern workplace. Apart from the verbal channel, nonverbal behavioral cues form a rich communication channel through which people infer - often automatically and unconsciously - emotions, relationships, and traits of fellow members.

There exists a solid body of knowledge about small groups and the multimodal nature of the nonverbal phenomenon in social psychology and nonverbal communication. However, the problem has only recently begun to be studied in the multimodal processing community. A recent trend is to analyze these interactions in the context of face-to-face group conversations, using multiple sensors and make inferences automatically without the need of a human expert. These problems can be formulated in a machine learning framework involving the extraction of relevant audio, video features and the design of supervised or unsupervised learning models. While attempting to bridge social psychology, perception, and machine learning, certain factors have to be considered. Firstly, various group conversation patterns emerge at different time-scales. For example, turn-taking patterns evolve over shorter time scales, whereas dominance or group-interest trends get established over larger time scales. Secondly, a set of audio and visual cues that are not only relevant but also robustly computable need to be chosen. Thirdly, unlike typical machine learning problems where ground truth is well defined, interaction modeling involves data annotation that needs to factor in interannotator variability. Finally, principled ways of integrating the multimodal cues have to be investigated. In the thesis, we have investigated individual social

constructs in small groups like dominance and status (two facets of the so-called vertical dimension of social relations). In the first part of this work, we have investigated how dominance perceived by external observers can be estimated by different nonverbal audio and video cues, and affected by annotator variability, the estimation method, and the exact task involved. In the second part,we jointly study perceived dominance and role-based status to understand whether dominant people are the ones with high status and whether dominance and status in small group conversations be automatically explained by the same nonverbal cues. We employ speaking activity, visual activity, and visual attention cues for both the works. In the second part of the thesis, we have investigated group social constructs using both supervised and unsupervised approaches. We first propose a novel framework to characterize groups. The two-layer framework consists of a individual layer and the group layer. At the individual layer, the floor-occupation patterns of the individuals are captured. At the group layer, the identity information of the individuals is not used. We define group cues by aggregating individual cues over time and person, and use them to classify group conversational contexts - cooperative vs competitive and brainstorming vs decision-making. We then propose a framework to discover group interaction patterns using probabilistic topic models. An objective evaluation of our methodology involving human judgment and multiple annotators, showed that the learned topics indeed are meaningful, and also that the discovered patterns resemble prototypical leadership styles - autocratic, participative, and free-rein - proposed in social psychology.

Advisor(s): Thesis supervisor: Daniel Gatica-Perez
SIG MM member(s): Daniel Gatica-Perez
http://library.epfl.ch/theses/?nr=4986

Social Computing group

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

Social computing is an emerging research domain focused on the automatic sensing, analysis, and interpretation of human and social behavior from sensor data. Through microphones and cameras in multi-sensor spaces, mobile phones, and the web, sensor data depicting human behavior can increasingly be obtained at large-scale - longitudinally and population-wise. The research group integrates models and methods from multimedia signal processing and information systems, statistical machine learning, ubiquitous computing, and applying knowledge from social sciences to address questions related to the discovery, recognition, and prediction of short-term and long-term behavior of individuals, groups, and communities in real life. This can range from people at work having meetings, users of social media sites, or people with mobile phones in urban environments. The group's research methods are aimed at creating ethical, personally and socially meaningful applications that support social interaction and communication, in the contexts of work, leisure, healthcare, and creative expression.

Katayoun Farrahi

A Probabilistic Approach to Socio-Geographic Reality Mining

As we live our daily lives, our surroundings know about it. Our surroundings consist of people, but also our electronic devices. Our mobile phones, for example, continuously sense our movements and interactions. This socio-geographic data could be continuously captured by hundreds of millions of people around the world and promises to reveal important behavioral clues about humans in a manner never before possible. Mining patterns of human behavior from large-scale mobile phone data has deep potential impact on society. For example, by understanding a community's movements and interactions, appropriate measures may be put in place to prevent the threat of an epidemic. The study of such human-centric massive datasets requires advanced mathematical models and tools. In this thesis, we investigate probabilistic topic models as unsupervised machine learning tools for large-scale socio-geographic activity mining.
We first investigate two types of probabilistic topic models for large-scale location-driven phone data

mining. We propose a methodology based on Latent Dirichlet Allocation, followed by the Author Topic Model, for the discovery of dominant location routines mined from the MIT Reality Mining data set containing the activities of 97 individuals over the course of a 16 month period. We investigate the many possibilities of our proposed approach in terms of activity modeling, including differentiating users with high and low varying lifestyles and determining when a user's activities fluctuate from the norm over time. We then consider both location and interaction features from cell tower connections and Bluetooth, in single and multimodal forms for routine discovery, where the daily routines discovered contain information about the interactions of the day in addition to the locations visited. We also propose a method for the prediction of missing multimodal data based on Latent Dirichlet Allocation. We further consider a supervised approach for day type and student type classification using similar socio-geographic features. We then propose two new probabilistic approaches to alleviate some of the limitations of Latent Dirichlet Allocation for activity modeling. Large duration activities and varying time duration activities can not be modeled with the initially proposed methods due to problems with input and model parameter size explosion. We first propose a Multi-Level Topic Model as a method to incorporate multiple time duration sequences into a probabilistic generative topic model. We then propose the Pairwise-Distance Topic Model as an approach to address the problem of modeling long duration activities with topics. Finally, we consider an application of our work to the study of influencing factors in human opinion change with mobile sensor data. We consider the Social Evolution Project Reality Mining dataset, and investigate other mobile phone sensor features including communication logs. We consider the difference in behaviors of individuals who change political opinion and those who do not. We combine several types of data to form multimodal exposure features, which express the exposure of individuals to others' political opinions. We use the previously defined methodology based on Latent Dirichlet Allocation to define each group's behaviors in terms of their exposure to opinions, and determine statistically significant features which differentiate those who change opinions and those who do not. We also consider the difference in exposure features of individuals that increases their interest in politics versus those who do not. Overall, this thesis addresses several important issues in the recent body of work called Computational Social Science. Investigations principled on mathematical models and multiple types of mobile phone sensor data are performed to mine real life human activities in largescale scenarios.

Advisor(s): Daniel Gatica-Perez, thesis supervisor

SIG MM member(s): Daniel Gatica-Perez
ISBN number: DOI: 10.5075/epfl-thesis-5018
http://library.epfl.ch/theses/?nr=5018 http://www.idiap.ch/~kfarrahi/finalthesis.pdf

Kimiaki Shirahama

Intelligent Video Processing Using Data Mining Techniques

Due to the rapidly increasing video data on the Web, much research effort has been devoted to develop video retrieval methods which can efficiently retrieve videos of interest. Considering the limited man-power, it is much expected to develop retrieval methods which use features automatically extracted from videos. However, since features only represent physical contents (e.g. color, edge, motion, etc.), retrieval methods require knowledge of how to use/integrate features for retrieving videos relevant to queries. To obtain such knowledge, this thesis concentrates on `video data mining' where videos are analyzed using data mining techniques which extract previously unknown, interesting patterns in underlying data. Thereby, patterns for retrieving relevant shots to queries are extracted as explicit knowledge.
Queries can be classified into three types. For the first type of queries, a user can find keywords suitable for retrieving relevant videos. For the second type of queries, the user cannot find such keywords due to the lexical ambiguity, but can provide some example videos. For the final type of queries, the user has neither keywords nor example videos. Thus, this thesis develops a video retrieval system with `multi-modal' interfaces by implementing three video data mining methods to support each of the above three query types. For the first query type, the system provides a `Query-By-Keyword' (QBK) interface where patterns which characterize videos relevant to certain keywords are extracted. For the second query type, a `Query-By-Example' (QBE) interface is provided where relevant videos are retrieved based on their similarities to example videos provided by the user. So, patterns for defining meaningful shot similarities are extracted using example videos. For the final qu ery type, a `Query-By-Browsing' (QBB) interface is devised where abnormal video editing patterns are detected to characterize impressive segments in videos, so that the user can browse these videos to find keywords or example videos. Finally, to improve retrieve performance, the integration of QBK and QBE is explored where informations from text and image/video modalities are interchanged using knowledge base which represents relations among semantic contents. The developed video data mining methods and the integration method are summarized as follows. The method for the QBK interface focuses that a certain semantic content is presented by concatenating several

shots taken by different cameras. Thus, this method extracts `sequential patterns' which relate adjacent shots relevant to certain keyword queries. Such patterns are extracted by connecting characteristic features in adjacent shots. However, the extraction of sequential patterns requires an expensive computation cost because a huge number of sequences of features have to be examined as candidates of patterns. Hence, time constraints are adopted to eliminate semantically irrelevant sequences of features. The method for the QBE interface focuses on a large variation of relevant shots. This means that even for the same query, relevant shots contain significantly different features due to varied camera techniques and settings. Thus, `rough set theory' is used to extract multiple patterns which characterize different subsets of example shots. Although this pattern extraction requires counter-example shots which are compared to example shots, they are not provided. Hence, `partially supervised learning' is used to collect counter-example shots from a large set of shots left behind in the database. Particularly, to characterize the boundary between relevant and irrelevant shots, the method collects counter-example shots which are as similar to example shots as possible. The method for the QBB interface assumes that impressive actions of a character are presented by abnormal video editing patterns. For example, thrilling actions of the character are presented by shots with very short durations while his/her romantic actions are presented by shots with very long durations. Based on this, the method detects `bursts' as patterns consisting of abnormally short or long durations of the character's appearance. The method firstly performs a probabilistic time-series segmentation to divide a video into segments characterized by distinct patterns of the character's appearance. It then examines whether each segment contains a burst or not. The integration of QBK and QBE is achieved by constructing a `video ontology' where concepts such as Person, Car and Building are organized into a hierarchical structure. Specifically, this is constructed by considering the generalization/ specialization relation among concepts and their co-occurrences in the same shots. Based on the video ontology, concepts related to a keyword query are selected by tracing its hierarchical structure. Shots where few of selected concepts are detected are filtered, and then QBE is performed on the remaining shots. Experimental results validate the effectiveness of all the developed methods. In the future, the multi-modal video retrieval system will be extended by adding a `Query-By-Gesture' (QBG) interface based on virtual reality techniques. This enables a user to create example shots for any arbitrary queries by synthesizing his/her gesture, 3DCG and background images.

Advisor(s): Prof. Dr. Kuniaki Uehara (supervisor)
SIG MM member(s): Kimiaki Shirahama

http://www.ai.cs.kobe-u.ac.jp/~kimi/papers/shirahama_thesis.pdf

CS 24 Uehara Laboratory at Graduate School of System Informatics, Kobe University

http://www.ai.cs.scitec.kobe-u.ac.jp/

Our research group aims at developing fundamental and practical technologies to utilize knowledge extracted from multimedia data. To this end, we are conducting research in broad areas of artificial intelligence, more specifically, machine learning, video data mining, time-series data analysis, information retrieval, trend analysis, knowledge discovery, etc. with typically a large amount of data.

As a part of the research efforts, we are developing a multi-modal video retrieval system where different media, such as text, image, video, and audio, are analyzed using machine learning and data mining techniques. We formulate video retrieval as a classification problem to discriminate between relevant and irrelevant shots to a query. Various techniques, such as rough set theory, partially supervised learning, multi-task learning, and Hidden Markov Model (HMM), are applied to the classification. Recently, we began to develop a gesture-based video retrieval system where information from various sensors, including RGB cameras, depth sensors, and magnetic sensors, are fused using virtual reality and computer vision techniques. In addition, transfer learning and collaborative filtering are utilized to refine the video annotation. Another pillar of our research group is concerned with more deeper analysis of natural language text. Our primary focus is to distill both explicit and implicit information contained therein. The former is generally seen as the problems of information extraction, question answering, passage retrieval, and annotation, and the latter as hypothesis discovery and text mining. Explicit information is directly described in text but not readily accessible by computers as it is embedded in complex human language. On the other hand, implicit information cannot be found in a single document and is only understood by synthesizing knowledge fragments scattered across a large number of documents. We take statistical natural language processing (NLP)-

and machine learning-based approaches, such as kernel-based online learning and transductive transfer learning, to tackling these problems. As described above, the common foundation underlying our research methodologies is machine learning, which requires more and more computing power reflecting increasingly available large-scale data and more complex algorithms. To deal with it, we are also engaged in developing parallel machine learning frameworks using MapReduce, MPI, Cell, and GPGPU. These works are ongoing and will be shared with the research community soon. More details of our research group can be found on our web site at http://www.ai.cs.scitec.kobe-u.ac.jp.

Pinaki Sinha

Automatic Summarization of Personal Photo Collections

Photo taking and sharing devices (e.g., smart phones, digital cameras, etc) have become extremely popular in recent times. Photo enthusiasts today capture moments of their personal lives using these devices. This has resulted in huge collections of photos stored in various personal archives. The exponential growth of online social networks and web based photo sharing platforms have added fuel to this fire. According to recent estimates [46], three billion photos are uploaded on the social network Facebook per month. This photo data overload has created some major challenges. One of the them is automatic generation of representative overviews from large photo collections. Manual browsing of photo corpora is not only tedious, but also time inefficient. Hence, development of an automatic photo summarization system is not only a research but also a practical challenge. In this dissertation, we present a principled approach for generation of size constrained overview summaries from large personal photo collections.

We define a photo summary as an extractive subset, which is a good representative of the larger photo set. We propose three properties that an effective summary should satisfy: Quality, Diversity and Coverage. Modern digital photos come with heterogeneous content and context data. We propose models which can combine this multimodal data to compute the summary properties. Our summarization objective is modeled as an optimization of these properties. Further, the summarization framework can integrate user preferences in form of inputs. Thus, different summaries may be generated from the same corpus to accommodate preference variations among the users. A traditional way of intrinsic evaluation in information retrieval is comparing the retrieved result set with a manually generated ground truth. However, given the variability of human behavior in selection of appealing photos, it may be difficult and non-intuitive to generate a unique ground truth summary of a larger data corpus. Due to the personal nature of the dataset, only the contributor of a particular photo corpus can possibly summarize it (since personal photos typically come with lots of background personal knowledge). While considerable efforts have been directed towards evaluation of annotation and ranking in multimedia, relatively few experiments have been done to evaluate photo summaries. We conducted extensive user studies on summarization of photos from single life events. The experiments showed certain uniformity and some diversity of user preferences in generating and evaluating photo summaries. We also posit that photo summaries should serve the twin objectives of information discovery and reuse. Based on this assumption, we propose novel objective metrics which enables us to evaluate summaries from large personal photo corpora without user generated ground truths. We also create dataset of personal photos along with hosts of contextual data which can be helpful in future research. Our experiments show that the summarization properties and framework proposed can indeed be used to generate effective summaries. This framework can be extended to include other types information (e.g., social ties among multiple users present in a dataset) and to create personalized photo summaries.

Advisor(s): Professor Ramesh Jain (supervisor),
Professor Sharad Mehrotra (committee member),
Professor Padhraic Smyth (committee member),
Professor Deva Ramanan (committee member)
SIG MM member(s): Ramesh Jain

http://www.ics.uci.edu/~psinha/research/thesis/pinaki_thesis.pdf

Radu Andrei Negoescu

Modeling and understanding communities in online social

The amount of multimedia content is on a constant increase, and people interact with each other and with content on a daily basis through social media systems. The goal of this thesis was to model and understand emerging online communities that revolve around multimedia content, more specifically photos, by using large-scale data and probabilistic models in a quantitative approach. The disertation has four contributions. First, using data from two online photo management systems, this thesis examined different aspects of the behavior of users of these systems pertaining to the uploading and sharing of photos with other users and online groups. Second, probabilistic topic models were used to model online entities, such as users and groups of users, and the new proposed representations were shown to be useful for further understanding such entities, as well as to have practical applications in search and recommendation scenarios. Third, by jointly modeling users from two different social photo systems, it was shown that differences at the level of vocabulary exist, and different sharing behaviors can be observed. Finally, by modeling online user groups as entities in a topic-based model, hyper-communities were discovered in an automatic fashion based on various topic-based representations. These hyper-communities were shown, both through an objective and a subjective evaluation with a number of users, to be generally homogeneous, and therefore likely to constitute a viable exploration technique for online communities.

Advisor(s): Daniel Gatica-Perez, supervisor
SIG MM member(s): Daniel Gatica-Perez

ISBN number: DOI: 10.5075/epfl-thesis-5059

Social computing group

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

Our recent work has investigated methods to analyze small groups at work in multisensor spaces, populations of mobile phones users in urban environments, and on-line communities in social media.

Event and publication reports

MMSJ, Volume 17, Issue 1, February 2010

NewsBox=Editor-in-Chief: Thomas Plagemann
URL: http://www.springer.de/
Published: February 2011

Papers

- Cristian Hesselman, Pablo Cesar and David Geerts: *Guest editorial: Networked television*
- Z. Avramova, D. De Vleeschauwer, S. Wittevrongel and H. Bruneel: *Performance analysis of a caching algorithm for a catch-up television service*
- Marcelo G. Manzato, Danilo B. Coimbra and Rudinei Goularte: *An enhanced content selection mechanism for personalization of video news programmes*
- Nairon S. Viana and Vicente F. de Lucena: *iDTV Home Gateway convergence: an open software model integrating the Ginga middleware and the OSGi framework*
- Morten Lindeberg, Stein Kristiansen, Thomas Plagemann and Vera Goebel: *Challenges and techniques for video streaming over mobile ad hoc networks*

MMSJ, Volume 17, Issue 2, March 2010

NewsBox=Editor-in-Chief: Thomas Plagemann
URL: http://www.springer.de/
Published: March 2011

Papers

- Arnar Ólafsson, Björn Þór Jónsson, Laurent Amsaleg and Herwig Lejsek: *Dynamic behavior of balanced NV-trees*
- Lei Xie, Zhong-Hua Fu, Wei Feng and Yong Luo: *Pitch-density-based features and an SVM binary tree approach for multi-class audio classification in broadcast news*
- Chung-Hua Chu, De-Nian Yang, Ya-Lan Pan and Ming-Syan Chen: *Stabilization and extraction of 2D barcodes for camera phones*
- Mohammed Belkhatir: *A three-level architecture for bridging the image semantic gap*
- Gorka Marcos, Arantza Illarramendi, Igor G. Olaizola and Julián Flórez: *A middleware to enhance current multimedia retrieval systems with content-based functionalities*

ACM Multimedia Systems Conference 2011

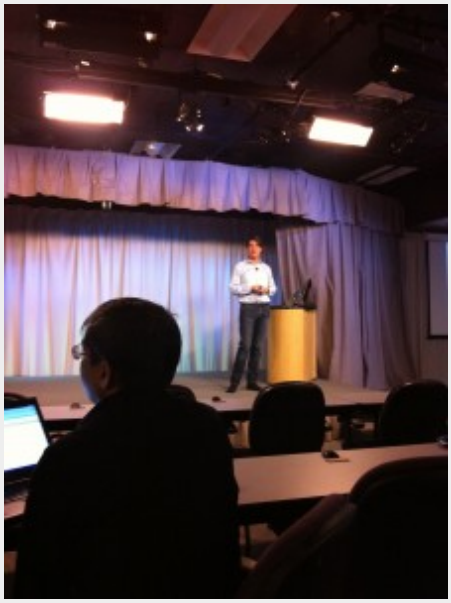
Conference Chairs: Mark Claypool, Christian Timmerer, Ketan Mayer-Patel, Ali C. Begen
Event location: San Jose, CA, USA
Event date: February 22-23, 2010
URL: http://www.mmsys.org
Sponsored by ACM SIG Multimedia, SIGOPS, SIGCOMM and Cisco
Reported by Mark Claypool, Kristian Evensen, Carsten Griwodz

The second MMSys conference was held in San Jose, CA, in February on the Cisco campus. The second MMSys conference saw a multiplication of the number of participants compared to the inaugural conference and was considered a great success by organizers and participants. We attribute the growth of the conference to several factors. MMSys becomes better know in the community after replacing MMCN; co-chair Christian Timmerer connected MMSys to the MPEG community by organizing a special session called "Multimedia Transport: DASH"; MMSys offered a dataset track that allows researchers to share the often-ignored work of collecting data and get credit for it; and conference fees were kept extremely low thanks to the sponsorship of our host, Cisco. An great addition was the provision of live feeds from MMSys to remote participants through Cisco's WebEx and Cisco TV.

With 15 papers accepted for the main conference, 5 long and 4 short for the special session, 5 for the dataset track, 2 keynotes and demos by Cisco, MMSys was arranged as a 3-day event. The keynotes were held by Alain Fiocco of Cisco and Mark Watson of Netflix. The themes of the conference covered multimedia systems, real-time support for multimedia, modeling of multimedia systems, mobile multimedia systems, multimedia databases, networked games, virtual and augmented worlds, and several more. A special issue journal containing papers relevant for MMT: DASH is currently under preparation. Readers should note that slides from all presentations and the keynotes are available from the MMSys web page, http://www.mmsys.org.

For the audience, MMSys started with a keynote from Cisco, where Alain Fiocco presented what Cisco sees as the future of the Internet. Unsurprisingly, this includes coping with the increased amount of traffic generated by video. As more and more multimedia devices get access to the internet, they believe the demand on the core network will increase significantly and smarter solutions than what we have today will be needed. One suggestion was that more logic and advanced devices will be placed in the network.

The theme of the first session was wireless and mobile, where the papers combined wireless networks with the capabilities of modern mobile devices (like GPS, cameras and so on). For example, the paper "GPS-aided Recognition-based User Tracking System with Augmented Reality in Extreme Large-scale Areas" was about using GPS to improve augmented reality. The second session was titled "Networking", which contained three papers. Then, three papers on QoS and data transmission were presented, before several datasets were introduced in the last session. These will be hosted at the MMSys website and include traces from virtual worlds and a visual search dataset.



Day two began with a keynote by Mark Watson from Netflix. He spoke about their HTTP-streaming solution and gave details on and insight into how it

works, and what challenges they are currently trying to solve. Several of these challenges concern the problem of understanding video quality, but also multipath streaming.

MMSys continued with the first set of DASH-papers. It was a bit difficult to follow the presentations without prior knowledge. However, if you have an interest in the future of dynamic adaptive streaming, the papers are worth checking out. We were then shown several impressive Cisco demos, before the final session of the day took place. This session was about system performance, and included a very interesting presentation about using flash memory as video-on-demand storage, documented in the paper "Impact of Flash Memory on Video-on-Demand Storage: Analysis of Tradeoffs". The authors had performed several experiments and evaluated the performance of different combinations of storage, and concluded that while SSD was unsuitable as main storage, it was very efficient when combined with a traditional HDD. When using the SSD as a cache, the performance was better than HDD+DRAM, as well as cheaper. The first session of the third day was about encoding, while the conference finished with a second session on DASH.

During the three days the conference lasted, we had several interesting discussions and got many ideas. After the final presentation, everyone agreed that it had been a good conference with interesting presentations and lots of cool people. If you are interested in multimedia systems, MMSys is highly recommended.

Papers

Wireless and mobile

- Wei Guan, Suya You, Ulrich Neumann: *GPS-aided recognition-based user tracking system with augmented reality in extreme large-scale areas*
- Jia Hao, Seon Ho Kim, Sakire Arslan Ay, Roger Zimmermann: *Energy-efficient mobile video management using smartphones*
- Navin K. Sharma, David E. Irwin, Prashant J. Shenoy, Michael Zink: *MultiSense: fine-grained multiplexing for steerable camera sensor networks*

Networking

- Lawrence Stewart, David A. Hayes, Grenville Armitage, Michael Welzl, Andreas Petlund: *Multimedia-unfriendly TCP congestion control and home gateway queue management*
- Mukundan Venkataraman, Mainak Chatterjee: *Effects of internet path selection on video-QoE*
- Kristian Evensen, Dominik Kaspar, Carsten Griwodz, Pål Halvorsen, Audun Hansen, Paal Engelstad: *Improving the performance of quality-adaptive video streaming over multiple heterogeneous access networks*

Data transmission and QoS

- Zixia Huang, Wanmin Wu, Klara Nahrstedt, Raoul Rivas, Ahsan Arefin: *SyncCast: synchronized dissemination in multi-site interactive 3D tele-immersion*
- Kuan-Ta Chen, Chen-Chi Wu, Yu-Chun Chang, Chin-Laung Lei: *Quantifying QoS requirements of network services: a cheat-proof framework*
- Dominik Seiler, Ernst Juhnke, Ralph Ewerth, Manfred Grauer, Bernd Freisleben: *Efficient data transmission between multimedia web services via aspect-oriented programming*

Dataset track

- Yichuan Wang, Cheng-Hsin Hsu, Jatinder Pal Singh, Xin Liu: *Network traces of virtual worlds: measurements and applications*
- Martin Ellis, Colin Perkins, Dimitrios P. Pezaros: *End-to-end and network-internal measurements of real-time traffic to residential users*
- Vijay R. Chandrasekhar, David M. Chen, Sam S. Tsai, Ngai-Man Cheung, Huizhong Chen, Gabriel Takacs, Yuriy Reznik, Ramakrishna Vedantham, Radek Grzeszczuk, Jeff Bach, Bernd Girod: *The stanford mobile visual search data set*
- Yeng-Ting Lee, Kuan-Ta Chen, Yun-Maw Cheng, Chin-Laung Lei: *World of warcraft avatar history dataset*
- Ricardo A. Calix, Gerald M. Knapp: *Affect corpus 2.0: an extension of a corpus for actor level emotion magnitude detection*

Modern media transport 1

- Thomas Stockhammer: *Dynamic adaptive streaming over HTTP --: standards and design principles*
- Luca De Cicco, Saverio Mascolo, Vittorio Palmisano: *Feedback control for adaptive live video streaming*
- Saamer Akhshabi, Ali C. Begen, Constantine Dovrolis: *An experimental evaluation of rate-adaptation algorithms in adaptive streaming over HTTP*
- Chenghao Liu, Imed Bouazizi, Moncef Gabbouj: *Rate adaptation for adaptive HTTP streaming*

System performance

- Moonkyung Ryu, Hyojun Kim, Umakishore Ramachandran: *Impact of flash memory on video-on-demand storage: analysis of tradeoffs*
- Samamon Khemmarat, Renjie Zhou, Lixin Gao, Michael Zink: *Watching user generated videos with prefetching*

- Kewin O. Stoeckigt, Hai L. Vu, Philip Branch: *Dynamic codec with priority for voice over IP in WLAN*

Encoding and repair

- Khiem Quang Minh Ngo, Ravindra Guntur, Wei Tsang Ooi: *Adaptive encoding of zoomable video streams based on user access pattern*
- Osama Abboud, Thomas Zinner, Konstantin Pussep, Sabah Al-Sabea, Ralf Steinmetz: *On the impact of quality adaptation in SVC-based P2P video-on-demand systems*
- David Varodayan, Wai-tian Tan: *Error-resilient live video multicast using low-rate visual quality feedback*

Modern media transport 2

- Robert Kuschnig, Ingo Kofler, Hermann Hellwagner: *Evaluation of HTTP-based request-response streams for internet video streaming*
- Yago Sánchez de la Fuente, Thomas Schierl, Cornelius Hellge, Thomas Wiegand, Dohy Hong, Danny De Vleeschauwer, Werner Van Leekwijck, Yannick Le Louédec: *iDASH: improved dynamic adaptive streaming over HTTP using scalable video coding*
- Cyril Concolato, Jean Le Feuvre, Romain Bouqueau: *Usages of DASH for rich media services*
- Christopher Müller, Christian Timmerer: *A test-bed for the dynamic adaptive streaming over HTTP featuring session mobility*
- Frank Hartung, Sinan Kesici, Daniel Catrein: *DRM protected dynamic adaptive HTTP streaming*

TOMCCAP, Volume 7, Issue 1, January 2011

NewsBox=Editor-in-Chief: Ralf Steinmetz
URL: http://tomccap.acm.org/
Sponsored by ACM SIG Multimedia
Published: January 2011

Papers

- Marek Meyer, Christoph Rensing, Ralf Steinmetz: *Multigranularity reuse of learning resources*
- Samia Bouyakoub, Abdelkader Belkhir: *SMIL builder: An incremental authoring tool for SMIL Documents*
- M. Anwar Hossain, Pradeep K. Atrey, Abdulmotaleb El Saddik: *Modeling and assessing quality of information in multisensor multimedia monitoring systems*
- Jianke Zhu, Steven C. H. Hoi, Michael R. Lyu, Shuicheng Yan: *Near-duplicate keyframe retrieval by semi-supervised learning and nonrigid image matching*
- Cheng-Hsin Hsu, Mohamed Hefeeda: *A framework for cross-layer optimization of video streaming in wireless networks*
- Surendar Chandra, Xuwen Yu: *An empirical analysis of serendipitous media sharing among campus-wide wireless users*

TOMCCAP, Volume 7, Issue 2, February 2011

Editor-in-Chief: Ralf Steinmetz
URL: http://tomccap.acm.org/
Sponsored by ACM SIG Multimedia
Published: February 2011

Papers

- Ajay Gopinathan, Zongpeng Li: *Optimal layered multicast*
- Cheng-Hsin Hsu, Mohamed Hefeeda: *Using simulcast and scalable video coding to efficiently control channel switching delay in mobile tv broadcast networks*
- Yohan Jin, Balakrishnan Prabhakaran: *Knowledge discovery from 3D human motion streams through semantic dimensional reduction*
- Wei Cheng, Wei Tsang Ooi, Sebastien Mondet, Romulus Grigoras, Géraldine Morin: *Modeling progressive mesh streaming: Does data dependency matter?*
- Susmit Bagchi: *A fuzzy algorithm for dynamically adaptive multimedia streaming*
- Cheng-Hsin Hsu, Mohamed Hefeeda: *Statistical multiplexing of variable-bit-rate videos streamed to mobile devices*

Calls for contributions

Calls for SIGMM Sponsored and Co-sponsored Events

International ACM Workshop on Multimedia in Forensics and Intelligence

Full paper Deadline: July 11, 2011
Event location: Scottsdale, AZ, USA
Event date: Nov 28 - Dec 1, 2011
URL: http://madm.dfki.de/mifor2011/

MiFor 2011 offers a forum for bringing solutions from multimedia research into forensics and intelligence.

International Workshop on Automated Media Analysis and Production for Novel TV Services

Full paper Deadline: July 11, 2011
Event location: Scottsdale, AZ, USA
Event date: Nov 28 - Dec 1, 2011
URL: http://aiempro2011.inria.fr/

The workshop aims at exploring the application and evaluation of automated information extraction techniques and audiovisual content analysis tools to support future media production for novel TV services.

The Third ACM SIGMM Workshop on Social Media

Full paper Deadline: July 11, 2011
Event location: Scottsdale, AZ, USA
Event date: Nov 28 - Dec 1, 2011
URL: http://www.cais.ntu.edu.sg/~wsm2011/

The workshop seeks contributions on various aspects of social media including papers on related theory, methodology, algorithms and issues associated to social media content creation, modeling, manipulation, content analysis, information extraction, storage, search, learning and mining.

ACM Workshop on Social and Behavioral Networked Media Access

Full paper Deadline: July 11, 2011
Event location: Scottsdale, AZ, USA
Event date: Nov 28 - Dec 1, 2011
URL: http://www.elec.qmul.ac.uk/mmv/sbnma/

The first aim of this workshop is to address the question on how multimedia content analysis can be combined with information derived from behavioral modeling and social interactions in order to improve personalized content distribution. The second relates to the way in which personalized services can be offered to users, in a real time, ubiquitous seamless and comprehensive way.

Workshop on Sparse Representation for Event Detection in Multimedia

Full paper Deadline: July 11, 2011
Event location: Scottsdale, AZ, USA
Event date: Nov 28 - Dec 1, 2011
URL: http://research.microsoft.com/en-us/um/people/zhang/SRED11/

The goal of this workshop is to model, detect and recognise events using sparsity analysis and the applications that make use of sparse learning for event analysis in the context of multimedia data.

Third International Workshop on Social Signal Processing

Full paper Deadline: July 11, 2011
Event location: Scottsdale, AZ, USA
Event date: Nov 28 - Dec 1, 2011
URL: http://sspnet.eu/2011/05/wssp/

We seek to attract contributions representing the state-of-the-art efforts to develop algorithms that can process naturally occurring human social communication, decode communicative intent, and generate the appropriate socially-adept responses.

International ACM Workshop on Multimedia access to 3D Human Objects

Full paper Deadline: July 11, 2011
Event location: Scottsdale, AZ, USA
Event date: Nov 28 - Dec 1, 2011
URL: http://www-rech.telecom-lille1.eu/ma3ho/

This workshop aims at taking a leap forward in emerging research of multimedia access of 3D human objects, aggregating together basic research in 3D graphics, 3D recognition and retrieval.

The ACM International Workshop on Medical Multimedia Analysis and Retrieval

Full paper Deadline: July 11, 2011
Event location: Scottsdale, AZ, USA
Event date: Nov 28 - Dec 1, 2011
URL: http://web2.utc.edu/~swf134/Service/MMAR2011.htm

The goal of this workshop is to bring together researchers in the area of medical multimedia analysis and retrieval.

Story Representation Mechanism and Context

Full paper Deadline: July 11, 2011
Event location: Scottsdale, AZ, USA
Event date: Nov 28 - Dec 1, 2011
URL: http://www.srmc2011.org/

This half-day workshop will foster discussion on issues related to story generation, representation, discovery and understanding.

Internation ACM Workshop on Ubiquitous Meta User Interfaces

Full paper Deadline: July 11, 2011
Event location: Scottsdale, AZ, USA
Event date: Nov 28 - Dec 1, 2011
URL: http://www.dai-labor.de/Ubi-MUI2011/

The aim of this workshop is to bring together different research groups to foster the developments of highly intuitive, multimedia supported meta user interfaces that bring transparency, predictability, and control into intelligent environments.

Third ACM International Workshop on Multimedia Technologies for Distance Leaning

Full paper Deadline: July 11, 2011
Event location: Scottsdale, AZ, USA
Event date: Nov 28 - Dec 1, 2011
URL: http://mtdl2011.mine.tku.edu.tw/

The workshop will provide a public forum for researchers and practitioners to exchange new ideas and information regarding advancements in the state of the art of emerging multimedia and e-learning related issues from multidisciplinary perspectives.

International Workshop on Interactive Multimedia on Mobile and Portable Devices

Full paper Deadline: July 11, 2011
Event location: Scottsdale, AZ, USA
Event date: Nov 28 - Dec 1, 2011
URL: http://www.eecs.qmul.ac.uk/~cfshan/IMMPD.html

This workshop will bring together researchers from both academia and industry in domains including computer vision, audio and speech processing, machine learning, pattern recognition, communications, human-computer interaction, and media technology to share and discuss recent advances in interactive multimedia.

First International ACM Workshop on Music Information Retrieval with User-Centered and Multimodal Strategies

Full paper Deadline: July 11, 2011
Event location: Scottsdale, AZ, USA
Event date: Nov 28 - Dec 1, 2011
URL: http://mirum11.tudelft.nl/

The workshop aims to gather experts from the Music Information Retrieval community and neighboring fields at a premier multimedia venue, to initiate a cross-disciplinary dialogue on open challenges in the field of Music Information Retrieval with user-centered and/or multimodal strategies.

3DLife/Huawei ACM MM Grand Challenge 2011

Full paper Deadline: August 6th, 2011
Event location: Scottsdale, Arizona
Event date: Nov 28th - Dec 1st, 2011
URL: http://www.3dlife-noe.eu/3DLife/emc2/grand-challenge/

EMC2 have teamed up with Huawei to present a Grand Challenge at ACM Multimedia 2011 on "Realistic Interaction in Online Virtual Environments". The challenge provides a captured dataset that consists of recordings of a number of Salsa dancers from a variety of modalities, together with ground-truth annotations of the choreographies.

ACM Multimedia Systems (MMSys)

Full paper Deadline: September 19, 2010
Event location: Chapel Hill, NC, USA
Event date: February 22-24, 2012
URL: http://www.mmsys.org/

The ACM Multimedia Systems conference provides a forum for researchers, engineers, and scientists to present and share their latest research findings in multimedia systems. While research about specific aspects of multimedia systems is regularly published in the various proceedings and transactions of the networking, operating system, real-time system, and database communities, MMSys aims to cut across these domains in the context of multimedia data types.

Calls for Events held in cooperation with SIGMM

International Workshop on Network and Systems Support for Games (NetGames)

Full paper Deadline: July 10, 2011
Event location: Ottawa, Canada

Event date: October 6-7 2011
URL: http://www.discover.uottawa.ca/netgames2011/

NetGames brings together researchers and practitioners from both academia and industry to present the latest research results and challenges of todays networked games.

International Conference on Advances in Mobile Computing & Multimedia (MoMM)

Full paper Deadline: July 15, 2011
Event location: Hue City, Vietnam
Event date: December 5-7, 2011
URL: http://www.iiwas.org/conferences/momm2011/

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.

Asia Information Retrieval Societies Conference

Full paper Deadline: July 17, 2011
Event location: Dubai, United Arab Emirates
Event date: December 18-20, 2011
URL: http://www.uowdubai.ac.ae/airs2011/

AIRS aims to bring together researchers and developers to exchange new ideas and latest achievements in the field of information retrieval. The scope of the conference covers applications, systems, technologies and theory aspects of information retrieval in text, audio, image, video, and multimedia data.

Other multimedia-related Events

The 18th International Conference on MultiMedia Modeling (MMM)

Full paper Deadline: July 22, 2011
Event location: Klagenfurt, Austria
Event date: January 4-6, 2012
URL: http://mmm2012.org/

The conference calls for research papers reporting original investigation results and demonstrations in areas related to multimedia modeling technologies and applications.

IEEE International Symposium on Multimedia

Full paper Deadline: July 6, 2011
Event location: Dana Point, CA, USA
Event date: December 5-7, 2011
URL: http://ism.eecs.uci.edu/

ISM is an international forum for researchers to exchange information regarding advances in the state of the art and practice of multimedia computing, as well as to identify the emerging research topics and define the future of multimedia computing.

Free Material

Huawei/3DLife Dataset for ACM Multimedia 2011 Grand Challenge

URL: http://www.acmmm11.org/content-huawei3dlife-challenge.html

Huawai/3DLife have just released an exciting data set to the community as part of our coordination the ACM MM Grand Challenge entitled "Realistic Interaction in Online Virtual Environments". The dataset consists of groundtruths/recordings of multiple Salsa dancers from a variety of modalities, which include microphones, cameras, inertial and depth sensors.

Job Opportunities

PhD Open Position at Telecom ParisTech: New representation and compression framework for 3D video

Employer: Institut Telecom - Telecom ParisTech
Valid until: Autumn 011
More info: http://www.tsi.telecom-paristech.fr/mm/en

3D video is widely perceived as the next major advancement in video technology. However, current solutions based on stereoscopic vision only exploit limited depth cues. In this PhD thesis, we will explore a

new representation and compression framework for 3D video, with the potential to support all depth cues and to greatly enhanced user experience.

The candidate must hold a master degree in electrical engineering, computer science or equivalent. The candidate will have a strong knowledge in image and video processing, solid basis in mathematics, and good programming skills in Matlab or C/C++. The candidate will be fluent in English.

The doctoral research work will be carried out within the Multimedia Group (http://www.tsi.telecom-paristech.fr/mm/en/), in the Signal and Image Processing Department at Telecom ParisTech. Telecom ParisTech (http://www.telecom-paristech.fr/eng/telecom-paristech.html), one of France's top five graduate engineering schools, is considered the leading French school in Information and Communication Technology (ICT).

The duration of the PhD program is 3 years. The candidate will be supported by a fellowship for the whole duration with a competitive salary. Starting date is expected in Fall 2011.

To apply:

Applicants should send a motivation letter and a complete CV to Dr. Frederic Dufaux (frederic.dufaux@telecom-paristech.fr). Applications will be considered until the position is filled.

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
	Wei Tsang Ooi, National University of Singapore