

ACM multimedia



October 15 – 19, 2016
Amsterdam | The Netherlands



Chairs' Welcome

ACM Multimedia 2016

We warmly welcome you to the 24th ACM Multimedia conference, which is hosted for the first time in the Netherlands, in the wonderful city of Amsterdam. ACM Multimedia 2016 brings an extensive program consisting of technical sessions covering all aspects of the multimedia field in the form of oral and poster presentations, tutorials, panels, exhibits, demonstrations and workshops, bringing into focus the principal subjects of investigation, competitions of research teams on challenging problems, and an interactive art program stimulating artists and computer scientists to meet and discover together the frontiers of artistic communication. The call for contributions attracted submissions from all over the world, which were all thoroughly reviewed for their merit in terms of scientific quality, innovation, and match to the conference. In summary, the various dedicated program committees reviewed and accepted the following:

Track	Reviewed	Accepted	
Long papers	237	52	22%
Short papers	413	127	31%
Brave new idea papers	14	4	29%
Demos	54	31	57%
Doctoral symposium papers	14	9	64%
Multimedia grand challenge submissions	20	12	60%
Video demonstrations	3	3	100%
Open source competition submissions	22	16	73%
Interactive art papers	23	10	43%

In addition to these, the main program has two exciting keynote presentations, by Dirk Helbing from ETH Zurich, Switzerland, titled "A digital world to thrive in – How the Internet of Things can make the 'invisible hand' work" and by Jack van Wijk from Eindhoven University of Technology, the Netherlands, titled "Visual Analytics for Multimedia: Challenges and Opportunities". Furthermore, a visionary presentation will be given by the winner of the SIGMM Award for Outstanding Technical Contributions to Multimedia Computing, Communications and Applications 2016, Alberto Del Bimbo, from the University of Florence, Italy. The main program will conclude with the SIGMM Rising Stars Symposium, highlighting the scientific results and vision of the invited young researchers, who demonstrated

great potential in multimedia research and who are considered to become future leaders in the multimedia field.

The main program is accompanied by eight workshops to discuss challenging topics and six tutorials to bring you up to speed on important foundations of our multimedia field. The unique co-location of ACM Multimedia 2016 with the European Conference on Computer Vision (ECCV 2016) has brought the opportunity to have additional twelve invited tutorials from world leaders covering both multimedia and computer vision.

Putting together a conference like ACM Multimedia is a team effort with many people volunteering their precious time. We would like to thank all who contributed to making this year's conference a success: the authors for submitting their latest work, the chairs and reviewers of the various tracks and activities for realizing a high quality program, and all others who assured smooth organization and realization of the conference. Finally, we would like to thank our sponsors and supporters, SIGMM, University of Amsterdam, Delft University of Technology, Qualcomm Research, FxPal, COMMIT, Facebook, Microsoft, Google, Unified Streaming, IBM and ViSenze for allowing us to bring you more than just the scientific program.

ACM Multimedia 2016 General Chairs

Alan Hanjalic

Delft University of
Technology

Cees Snoek

Qualcomm Research &
University of Amsterdam

Marcel Worring

University of Amsterdam

Message from the Program Coordinator

ACM Multimedia 2016

I am very happy to welcome you to Amsterdam for Multimedia 2016. This is the 24th edition of what can be considered as the premier academic conference on multimedia systems, content and applications.

The conference general and technical program chairs have given a wealth of information on the nature, structure and process used to assemble the program that will unfold this week. I, too, would like to add a few comments to help participants better understand the nature of the sessions and the combination of program elements that we hope will serve as the basis of an inspiring visit.

There are always two broad audiences for the collections of papers presented as part of the ACM Multimedia series. First are researchers who access individual papers via the ACM Digital Library. Over the coming years, this is where the greatest impact of the contributions presented here is likely to manifest itself. The Technical Program Committee has done an excellent job in providing a corpus of research that reflects the current state of the various sub-disciplines covered at the conference. As a Program Coordinator, I have little influence on the selection of these papers: this is the work of the reviewers, various area chairs and the technical program committee chairs.

The second audience for this year's papers, posters and presentations are the attendees of the conference in Amsterdam. These people come to the conference not only to hear the latest individual research results – this is something that they can do more quickly and cheaply via the ACM DL! – but also to be inspired by work that is perhaps (just) outside their own areas of expertise. Here, the Program Coordinator has the responsibility to 'package' the papers at the conference in a way that makes new work available to as many people as possible while also maintaining the community bonds within our various historical and emerging interests groups.

This year, the conference program has been architected so that each session block contains a Topics in Multimedia session, along with three more specialized special-interest sessions. The Topics session contains a collection of papers from different disciplines but often on (somewhat) related topics. The specialized

sessions allow members of particular groups to interact more directly. In this way, there should always be a session of interest (either in breadth or depth) to capture the imagination of the participant. We hope that you find this approach useful and inspiring.

Another responsibility of the Program Coordinator is to coordinate the selection of the best paper at the conference. This year, the program committee nominated eight candidate papers. A second committee was formed (using recommendations from the technical program chairs) that selected four of these papers as candidates for the award at the conference. This was a particularly difficult task, as the quality of each of the candidates was very high. Since ACM Multimedia is a generalist conference, we used this as a guide in the ultimate selection. I encourage you to attend the best paper session to see this broad sampling of first-class research in our community.

Coming to a conference is an expense and time-consuming venture, but I hope that the work of all of the members of our community who volunteered to put together ACM Multimedia 2016 will inspire your own research, allow you to make new contacts and help us all advance the quality of our fields.

Dick C.A. Bulterman

CWI and Vrije Universiteit Amsterdam

Message from the Technical Program Chairs ACM Multimedia 2016

We are pleased to introduce to you to the Technical Program for the ACM International Conference on Multimedia (ACM Multimedia) 2016, which will take place in Amsterdam in October 2016. While adopting the best practices from previous ACM Multimedia conferences, this year we also introduced new approaches and guidelines in the review process to ensure an efficient and fair evaluation of all submissions in providing a high quality conference.

Following last year's structure, fifteen areas of interest were grouped into four research themes, including a System Theme (covering Multimedia Transport and Delivery, Multimedia Systems and Middleware, Multimedia Telepresence and Virtual/Augmented Reality, Multimedia Scalability and Management, Mobile Multimedia), an Experience Theme (covering Multimedia HCI and Quality of Experience, Music, Speech and Audio Processing in Multimedia, Multimedia Art, Entertainment and Culture, Multimedia for Collaboration in Education & Distributed Environments), an Understanding Theme (covering Deep Learning for Multimedia, Multimodal Analysis and Description, Multimedia and Vision), and an Engagement Theme (covering Emotional and Social Signals In Multimedia, Social Multimedia, Multimedia Search and Recommendation). Minor changes compared to last year's structure included moving the Multimedia Art area from the Engagement to the Experience theme and the addition of the Multimedia Scalability and Management area within the System's theme.

Among the noticeable changes for ACM Multimedia 2016 is the addition of a fifth page (reserved for citations) for the short papers, demos, grand challenge and open source software competitions, doctoral symposium, and interactive art submissions. Additionally, long paper submissions can range from 8 to 10 pages, with the minimum page length introduced this year.

Each research theme was overseen by one of this year's Technical Program Chairs. For each area of interest, a team of Area Chairs (ACs), ranging from one to eight members, was recruited to handle the paper review process. In total, 46 ACs were appointed. A new contribution this year was the preparation of detailed instructions for ACs and reviewers, describing each step of the process and an accompanying time schedule. Organizationally, we believe it helped ACs and

reviewers with overall planning logistics and time management. It is suggested that this process should be updated and used as a standard practice in future conferences. The second innovation this year was the use of the Toronto Paper Matching System (TPMS) for paper-reviewer assignment. This system has gained adoption in both machine learning and computer vision conferences and is being used, for example by CVPR, ICCV and ECCV. The TPMS worked in conjunction with the conference management system (CMT), whereby invited reviewers uploaded their own papers for analysis as part of the automatic assignment process. Overall, the matching system provided reasonably reliable results in most cases, but required close supervision and tweaking by the technical program chairs, to avoid issues such as reviewers from one area being allocated papers from another. The conference ACs played an important role here also in as well

The paper review and submission is described in detail below:

- Submission of full (abstract pre-submission was allowed for better review planning) and short papers were separated from each other, with the short paper submission deadline one month later than the deadline for full paper submissions.
- Following the previous year's approach, a two-stage review process for full paper submissions was applied. Authors were invited to submit a rebuttal based on the first round of reviews. The reviewers were then asked to consider the rebuttal, to revise their reviews accordingly, and to enter discussion with other reviewers with the aim of improving review consensus on each paper. The results of the second round reviews and discussions were then summarized by the ACs into meta-reviews, which included tentative recommendations, to be discussed and finalized during the TPC meeting.
- In the second review stage, authors had the opportunity to communicate with the Author's Advocate was consulted...“ if they believed there were issues or anomalies with the reviews. During this year's review cycle, the Author's Advocate was utilized in several cases to ensure a more equitable and satisfying review procedure.
- Submissions from any authors who also acted as ACs/TPCs were handled personally by non-conflicting Technical Program Chairs.
- Deadlines were predominantly adhered to, with minor delays on occasion.

In total, 248 full and 425 short paper submissions were received by their respective deadlines and entered into the regular review process.

The table below gives the exact submission numbers across the various research areas.

Area	Full	Short
Emotional and Social Signals In Multimedia	7	26
Multimedia Search and Recommendation	41	52
Social Multimedia	16	19
Multimedia Art, Entertainment and Culture	9	7
Multimedia for Collaboration in Education & Distributed Environments	2	3
Multimedia HCI and Quality of Experience	12	21
Music, Speech and Audio Processing in Multimedia	8	17
Mobile Multimedia	6	12
Multimedia Scalability and Management	5	5
Multimedia Systems and Middleware	13	12
Multimedia Telepresence and Virtual/Augmented Reality	10	12
Multimedia Transport and Delivery	17	10
Deep Learning for Multimedia	20	57
Multimedia and Vision	45	125
Multimodal Analysis and Description	26	35

Decisions on full and short paper acceptances were made at the one-day TPC meeting held at Columbia University on June 10, 2016. All four Technical Program Chairs attended the meeting and handled their assigned themes in different allocated rooms. The emphasis of the TPC meeting was on high-score, borderline, and controversial paper submissions. The discussions were detailed and intensive, taking into account all the available data, including reviews, the authors' rebuttals, the post-rebuttal discussions, the initial recommendations by the ACs (in the form of a meta-review) and any subsequent reviews performed by other ACs during the meeting itself. For some papers that span across multiple themes, the TPC/AC Chairs collaborate across theme to reach a consensus on the decision. Technical Program Chairs and ACs were requested to leave the room when a conflicting submission was discussed.

After the separate meetings, a plenary session coordinated by the TPCs took place to make final decisions, where the overall quality and balance of different tracks was taken into consideration. By the end of the day, the participants accepted 52 full and 127 short papers, leading to acceptance rates of 22.2% and 30.3% for full and short papers, respectively. Five high quality full paper submissions were also recommended for acceptance as short papers.

The TPC meeting was organized in conjunction with a supporting workshop as per previous years. This years' workshop (<http://acmmmtpcws16.eurecom.fr/>) was

integrated with the ICMR conference which also took place in New York in June. Workshop participants had the chance to listen to 13 enlightening talks about recent research developments from the entire spectrum of topics covered by ACM Multimedia.

The Technical Program Committee has worked diligently over the last year to ensure that the technical program of this year's conference is of the highest possible quality. We would like to recognize the great effort and hard work from the volunteer team of ACs and reviewers who helped shape an exciting technical program.

We look forward to seeing you this autumn, in Amsterdam, Netherlands for a compelling ACM Multimedia technical program.

Benoit Huet

EURECOM (FR)

Yiannis Kompatsiaris

CERTH-ITI (GR)

Aisling Kelliher

Virginia Tech (USA)

Jin Li

Microsoft (USA)

Practical Information

ACM Multimedia 2016

About Amsterdam

Amsterdam is one of the greatest small cities in the world. From Amsterdam canals to world-famous Amsterdam museums and historical Amsterdam sights, it is one of the most romantic and beautiful cities in Europe. Canal cruises are a popular way to see the city from the perspective of its canals.

Amsterdam is also a city of tolerance and diversity. It has all the advantages of a big city: rich culture, lively Amsterdam nightlife, international restaurants, good transport - but is quiet, and largely thanks to its extensive canals, has a little road traffic. In this city your destination is never far away, but get a bike for an authentic local experience.

Important.

Your badge serves as your entrance ticket.

No badge is no entrance (If you have a full registration you also need your ticket for the conference dinner).



--- Walking route

Practical Information

Venue Information

A **Roeterseiland**
(Faculty of Economics and Business / Faculty of Social and Behavioural Sciences)

For **Workshops and invited tutorials**
Conference Registration (Saturday 15th and Sunday 16th October, from 8am)

Address Roeterstraat 11
1018 WB Amsterdam

Directions To get to the Roeterseiland from the Amsterdam Central Station, you can use all the metro's.

Get off at the third stop (*Weesperplein*).

Head north on *Rhijnspoorplein* toward *Weesperplein*, Turn right at *Weesperplein*,

Turn left toward *Sarphatistraat*, Turn right onto *Sarphatistraat* and Turn left onto *Roetersstraat*
Destination will be on the left.

B **Theatre Tuschinski**
(Pathé Tuschinski is a movie theater in the Netherlands in Amsterdam commissioned by Abraham Ick Tuschinski in 1921 at a cost of 4 million guilders. The interior and exterior are a spectacular mix of styles, as designed by Hijman Louis de Jong; Amsterdam School, Jugendstil, Art Nouveau and Art Deco. The main auditorium hosts many premieres of Dutch films. It is considered to be one of the most beautiful cinemas in the world.)

For **Workshops**
Conference Registration (Monday 17 – Wednesday 19 October, from 8am)

Address Reguliersbreestraat 26-34
1017 CN Amsterdam

Directions To get to the Theatre Tuschinski from the Amsterdam Central Station, you can use tramlines 4, 9, 16, 24, or 25.

Get off at the second stop (*Spui*). The street you are on now is called the *Rokin*. Take a right and cross the '*Muntplein*' to enter the alley *Reguliersbreestraat*.

C**De Bazel****For Address** **Poster and Demo sessions (Monday 17th October)**

Vijzelstraat 32
1017 HL Amsterdam

Directions To get to the Bazel from the Amsterdam Central Station, you can use tramline 16.
Get off at the fourth stop (*Keizersgracht*). Cross the street and enter the Bazel.

D**I-Dock (Conference Dinner)****For Directions** **Conference Dinner (Tuesday 18th October)**

From the conference we will take the boat through the canals together to the restaurant.

E**Canal cruise to Conference Dinner****For Address** **Conference Dinner (Tuesday 18th October - 6pm)**

Rokin t.o. 125
1012 KK Amsterdam

F**Openbare Bibliotheek****For Address** **Art Show**

Oosterdokskade 143
1011 DL Amsterdam

Directions A 5 minute walk from Amsterdam Central Station.

Practical Information**Practical Matters**

More information about Amsterdam can be found at: www.iamsterdam.nl.

Safety & Crime

Amsterdam is considered to be a relatively safe city. In general, violent crimes are as rare as in other European Cities. However, as in most crowded cities, beware of pick-pockets; they are mainly active in crowded areas, and on the train connecting Amsterdam Central Station and Schiphol Airport. Watch your belongings carefully.

National Emergency numbers / Police

Police, fire brigade, ambulance: National: **+31 112**

Police, no emergency: **+31 900 8844**

Electricity

Within the Netherlands 230 Volts/50 Hertz electrical supply is used. Visitors may need converters and adapters.

Cloak room/luggage deposit

Is available at all venues.

Lost and found

There is a lost and found service at the registration desk in the hall of all the locations.

Lunch

At some days we are offering lunch bags. Please check the program.

Questions and requests

Volunteer assistants – wearing a red ACM-shirt – will be present at the registration desk conference rooms. They will provide assistance to speakers and other participants with practical answers.

Wifi

A limited capacity is available at all conference venues.

Program

ACM Multimedia 2016

Saturday, October 15 / Roeterseiland

	08:00	08:30	09:00	09:30	10:00	10:30	11:00	11:30	12:00		12:30	13:00	13:30	14:00	14:30	15:00	15:30	16:00	16:30	17:00	17:30	18:00	18:30	
Hall REC E	Registration																							
Hall	Coffee and drinks are served all day																							
Invited tutorials C1.03 11, 21, 31, 41, 51			Segmentation and Tracking				Developments in deep learning					Output embedding for large-scale visual recognition			Human body shape modelling			Deep visual understanding from deep learning						
Invited tutorials C0.01 12, 22, 32, 42, 52, 61			Variational wrapping deformations for image processing applications				Multimodal human-robot interaction					Video event detection, localization and cross-media linking			A guide to creating and managing lifelog collections			Bridging video and language with deep learning			Color and spectral information in modern computer vision			

Sunday, October 16 / Roeterseiland / Openbare Bibliotheek

	08:00	08:30	09:00	09:30	10:00	10:30	11:00	11:30	12:00		12:30	13:00	13:30	14:00	14:30	15:00	15:30	16:00	16:30	17:00	17:30	18:00	18:30	
Hall REC E	Registration																							
Hall	Coffee and drinks are served all day																							
Workshop 23 C0.02			Computer Vision for Audio-visual Media																					
Workshop 24 M1.03			Computer VISION for ART Analysis																					
Workshop 25 C1.03			Virtual/Augmented Reality for Visual Artificial Intelligence																					
Workshop 26 C1.04			Joint Workshop on Storytelling with Images and Videos and Large										Scale Movie Description and Understanding Challenge											
Workshop M1.02			Workshop on Vision and Language Integration Meets Multimedia																					
Workshop C2.06			Workshop on Multimedia Alternate Realities																					
Workshop C0.01			Workshop on Human Behavior Understanding																					
Workshop C2.05			Workshop on Multimedia Assisted Dietary Management																					
Workshop C2.03			Workshop on Multimedia for personal health and health care (p.98)																					
Workshop C2.04			Workshop on Lifelogging Tools and Applications																					
Workshop C2.02			Audio/Visual Emotion Challenge and Workshop																					
Workshop C2.03			Multimedia COMMONS Workshop 2016																					
Tutorial C3.01			Tutorial on Social and Affective Robotics																					
Tutorial C2.07			Tutorial on Technology and Art in Stimulating Creative Place-										making in Public-Use Spaces											
Tutorial C0.02																								
Tutorial C3.02																								
Tutorial C3.03																								
Openbare Bibliotheek Amsterdam																								Art show reception Theme: Data Esthetics

Monday, October 17 / Tuschinski / De Bazel

		08:00	08:30	09:00	09:30	10:00	10:30	11:00	11:30	12:00		12:30	13:00	13:30	14:00	14:30	15:00	15:30	16:00	16:30	17:00	17:30	18:00	18:30
Tuschinski	Entree hal						Break					Lunch Break												
Tuschinski	Zaal 1		Opening	Keynote I: Dirk Helbing			Break	Best paper session				Lunch Break												
Tuschinski	Zaal 2						Break					Lunch Break												
Tuschinski	Zaal 3						Break					ToMM meeting												
Tuschinski	Zaal 4						Break					Lunch Break												
Tuschinski	VIP						Break					Women in SIGMM lunch												
De Bazel	Lobby														Poster session I		Poster session II				Reception			
De Bazel	Demo floor														Demo session						Reception			
De Bazel	Lobby														Video program									

Tuesday, October 18 / Tuschinski

		08:00	08:30	09:00	09:30	10:00	10:30	11:00	11:30	12:00		12:30	13:00	13:30	14:00	14:30	15:00	15:30	16:00	16:30	17:00	17:30	18:00	18:30
Tuschinski	Entree hal						Break					Lunch Break						Break						Social event
Tuschinski	Zaal 1			SIGMM Award presentations / talks			Break	Panel				Lunch Break			Deep Learning			Break						
Tuschinski	Zaal 2						Break	Analysis & Search				Lunch Break			Brave New Topics			Break						
Tuschinski	Zaal 3						Break	Topics in Multimedia I				MMSJ meeting			Topics in Multimedia II			Break						
Tuschinski	Zaal 4						Break	Video Analysis & Streaming				MM-Mag Meeting			Events and Context			Break						
Tuschinski	VIP						Break					Doctoral Symposium						Break						

Wednesday, October 19 / Tuschinski

		08:00	08:30	09:00	09:30	10:00	10:30	11:00	11:30	12:00		12:30	13:00	13:30	14:00	14:30	15:00	15:30	16:00	16:30	17:00	17:30	18:00	18:30
Tuschinski	Entree hal						Break					Lunch Break						Break						
Tuschinski	Zaal 1			Keynote II: Jack van Wijk			Break	Open source software competition				SIGMM Business Meeting & ACMM M16 Awards			Topics in Multimedia IV			Break	SIGMM Rising Star Symposium					
Tuschinski	Zaal 2						Break	Topics in Multimedia III										Break						
Tuschinski	Zaal 3						Break	Learning & Hashing							Analysis & Middleware			Break						
Tuschinski	Zaal 4						Break	Transport & Experience							Emotions, People and Faces			Break						
Tuschinski	VIP						Break								Exchange meeting			Break						

Award presentations

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

The 2016 winner of the prestigious ACM Special Interest Group on Multimedia (SIGMM) award for Outstanding Technical Contributions to Multimedia Computing, Communications and Applications is Prof. Dr. Alberto del Bimbo. The award is given in recognition of his outstanding, pioneering and continued research contributions in the areas of multimedia processing, multimedia content analysis, and multimedia applications, his leadership in multimedia education, and his outstanding and continued service to the community.

Prof. del Bimbo was among the very few who pioneered the research in image and video content-based retrieval in the late 80's. Since that time, for over 25 years, he has been among the most visionary and influential researchers in Europe and world-wide in this field. His research has influenced several generations of researchers that are now active in some of the most important research centers world-wide. Over the years, he has made significant innovative research contributions.

In the early times of the discipline he explored all the modalities for retrieval by visual similarity of images and video. In his early paper Visual Image Retrieval by Elastic Matching of User Sketches published in IEEE Trans. on Pattern Analysis and Machine Intelligence in 1997, he presented one of the first and top performing methods for image retrieval by shape similarity from user's sketches. He also published in IEEE Trans. on Pattern Analysis and Machine Intelligence and IEEE Trans. on Multimedia his original research on representations for spatial relationships between image regions based on spatial logic. This ground-breaking research was accompanied by the definition of efficient index structures to permit retrieval from large datasets. He was one of the first to address this large datasets aspect that has now become very important for the research community.

Since the early 2000s, with the advancement of 3D imaging technologies and the availability of a new generation of acquisition devices capable of capturing the geometry of 3D objects in the three-dimensional physical space, Prof. del Bimbo and his team initiated research in 3D content based retrieval that has now become increasingly popular in mainstream research. Again, he was among the very first researchers to initiate this research. Particularly, he focused on 3D face recognition

extending the weighted walkthrough representation of spatial relationships between image regions to model the 3D relationships between facial stripes. His solution of 3D Face Recognition Using Iso-geodesic Stripes scored the best performance at SHREC Shape Retrieval Contest in 2008, and was published in IEEE Trans. on Pattern Analysis and Machine Intelligence, in 2010. At CVPR'15 he presented a novel idea for representing 3D textured mesh manifolds using Local Binary Patterns, that is highly effective for 3D face retrieval. This was the first attempt to combine 3D geometry and photometric texture into a single unified representation. In 2016 he has co-authored a forward looking survey on content-based image retrieval in the context of social image platforms, that has appeared on ACM Computing Surveys. It includes an extensive treatise of image tag assignment, refinement and tag-based retrieval and explores the differences between traditional image retrieval and retrieval with socially generated images.

One very important aspect of his contribution to the community is Professor del Bimbo's educational impact during his career. He was the author of the monograph, Visual Information Retrieval, published by Morgan Kaufmann in 1999 which became one of the most cited and influential books from the early years of image and video content-based retrieval. Many young researchers have used this book as the main reference in their studies, and their career has been shaped by the ideas discussed in this book. Being the first and sole book on that subject in the early times of the discipline, it played a key role to develop content-based retrieval from a research niche to a largely populated field of research and to make it central to Multimedia research.

Professor del Bimbo has an extraordinary and long-lasting track record of services to the scientific community through the last 20 years. As the General Chair he organized two of the most successful conferences in Multimedia, namely IEEE ICMCS'99, the Int'l Conf. on Multimedia Computing and Systems (now renamed IEEE ICME) and ACM MULTIMEDIA'10. The quality and success of these conferences were highly influential to attract new young researchers in the field and form the present research community. Since 2016, he is the Editor-in-Chief for ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM).

Award presentations

ACM SIGMM Rising Star Award 2016

Dr. Bart Thomee received his Ph.D. from Leiden University in 2010. In his thesis, he focused on multimedia search and exploration, specifically targeting artificial imagination and duplicate detection. On the topic of artificial imagination, he aimed to more rapidly understand the user's search intent by generating imagery that resemble the ideal image the user is looking for. Using the synthesized images as queries instead of existing images from the database boosted the relevance of the image results by up to 23%. On the topic of duplicate detection, he designed descriptors to compactly represent web-scale image collections and to accurately detect transformed versions of the same image. This work led to an Outstanding Paper Citation at ACM Conference on Multimedia Information Retrieval 2008.

In 2011, he joined Yahoo Labs, where Dr. Thomee's interests grew into geographic computing in Multimedia. He began characterizing spatiotemporal regions from labeled (e.g. tagged) georeferenced media, for which he devised a technique based on scale-space theory that could process billions of georeferenced labels in a matter of hours. This work was published at WWW 2013 and became a reference example at Yahoo for how to disambiguate multi-language and multi-meaning labels from media with noisy annotations.

He also started to use an overlooked piece of information that is found in most camera phone images: compass information. He developed a technique to accurately pinpoint the locations and surface area of landmarks, solely based on the positions and orientations of photos taken of them which may have been taken hundreds of yards to miles away.

Dr. Thomee's recent work on the YFCC100M dataset has had important impacts on the multimedia and SIGMM research community. This new dataset was real in size and structure to fuel and change the landscape of research in Multimedia. What started as an initiative to release a geo-Flickr dataset, Dr. Thomee quickly saw the broader impact and worked rapidly to scale the size. He had to push the limits of openness without violating licensing terms, copyright, or privacy. He worked closely with many lawyers to overturn the default, restrictive terms of use by making it also available to non-academics all over the world. He

coordinated and led the efforts to share the data and effort horizontally with ICSI, LLNL, and Amazon Open Data. It was highlighted in the 2016 February issue of the Communications of ACM (CACM). The dataset has been requested over 1200 times in just a few months and cited many times since launch. Dr. Thomee has continued by releasing expansion packs to the YFCC100M. This dataset is expected to impact Multimedia research significantly over the future years.

Dr. Thomee has also been an exemplary community member of the Multimedia community. For example, he organized the ImageCLEF photo annotation task (2012-2013) and MediaEval placing task (2013-2016) as well as designed the ACM Grand Challenge on Event Summarization (2015) and on Tag & Caption Prediction (2016).

In summary, Dr. Bart Thomee receives the 2016 ACM SIGMM Rising Star Award Thomee for significant contributions in the areas of geo-multimedia computing, media evaluation, and open datasets for research.

Award presentations

SIGMM Award for Outstanding Ph.D. Thesis in Multimedia Computing, Communications and Applications 2016

ACM Special Interest Group on Multimedia (SIGMM) is pleased to present the 2016 SIGMM Outstanding Ph.D. Thesis Award to Dr. Christoph Kofler. The award committee considers Dr. Kofler's dissertation entitled "User Intent in Online Video Search" worthy of the recognition as the thesis is the first to innovatively consider a user's intent in multimedia search yielding significantly improved results in satisfying the information need of the user. The work has high originality and is expected to have significant impact, especially in boosting the search performance for multimedia data.

Dr. Kofler's thesis systematically explores a user's video search intent that is behind a user's information need in three steps: (1) analyzing a real-world transaction log produced by a large video search engine to understand why searches fail, (2) understanding the possible intents of users behind video search and uploads, and (3) designing an intent-aware video search result optimization approach that re-ranks initial video search results so as to yield the highest potential to satisfy the users' search intent.

The effectiveness of the framework developed in the thesis has been successfully justified by a thorough range of experiments. The thesis topic itself is highly topical and the framework makes groundbreaking contributions to our understanding and knowledge in the area of users' information seeking, user intent, user satisfaction, and multimedia search engine usability. The publications related to the thesis clearly demonstrate the impact of this work across several research disciplines including multimedia, web, and information retrieval. Overall, the committee recognizes that the thesis has significant impact and makes considerable contributions to the multimedia community.

Bio of Awardee: Dr. Christoph Kofler is a software engineer and data scientist at Bloomberg L.P., NY, USA. He holds a Ph.D. degree from Delft University of Technology, The Netherlands, and an M.Sc. and B.Sc. degree from Klagenfurt University, Austria – all in Computer Science. His research interests include the broad fields of multimedia and text-based information retrieval with focus on search intent inference and its applications for search results optimization

throughout the entire search engine pipeline (indexing, ranking, query formulation). In addition to "what" a user is looking for using search, Dr. Kofler is particularly interested in the "why" component behind the search and in the related opportunities for improving the efficiency and effectiveness of information retrieval systems. Dr. Kofler has co-authored more than 20 scientific publications with predominant focus on venues such as ACM Multimedia, IEEE Transactions on Multimedia, and ACM Computing Surveys. He has been a task co-organizer of the MediaEval Benchmark initiative. He received the Grand Challenge Best Presentation Award at ACM Multimedia and the Best Paper nomination at the European Conference on Information Retrieval. Dr. Kofler is a recipient of the Google Doctoral Fellowship in Information Retrieval (Video Search). He has held positions at Microsoft Research, Beijing, China; Columbia University, NY, USA; and Google, NY, USA.

Keynotes

A digital world to thrive in – How the Internet of Things can make the ‚invisible hand‘ work

Speaker

Dirk Helbing, ETH Zurich

Abstract

Managing data-rich societies wisely and reaching sustainable development are among the greatest challenges of the 21st century. We are faced with existential threats and huge opportunities. If we don't act now, large parts of our society will not be able to economically benefit from the digital revolution. This could lead to mass unemployment and social unrest. It is time to create the right framework for the digital society to come.

Bio

Dirk Helbing is Professor of Computational Social Science at the Department of Humanities, Social and Political Sciences and affiliate of the Computer Science Department at ETH Zurich. He earned a PhD in physics and was Managing Director of the Institute of Transport & Economics at Dresden University of Technology in Germany. He is internationally known for his work on pedestrian crowds, vehicle traffic, and agent-based models of social systems. Furthermore, he coordinates the FuturICT Initiative (<http://www.futurict.eu>), which focuses on the understanding of techno-socio-economic systems, using smart data. His work is documented in hundreds of scientific articles, keynote lectures and media reports worldwide. Helbing is an elected member of the prestigious German Academy of Sciences "Leopoldina" and worked for the World Economic Forum's Global Agenda Council on Complex Systems. He is also co-founder of the Physics of Socio-Economic Systems Division of the German Physical Society and of ETH Zurich's Risk Center. In January 2014 Prof. Helbing received a honorary PhD from Delft University of Technology (TU Delft). Since June 2016 he is affiliate professor at the faculty of Technology, Policy and Management at TU Delft, where he leads the PhD school in „Engineering Social Technologies for a Responsible Digital Future“.

Keynotes

Visual Analytics for Multimedia: Challenges and Opportunities

Speaker

Jack van Wijk, Eindhoven University of Technology

Abstract

Understanding huge multimedia collections is a big challenge. Given a set of hundreds of thousands or millions of images, how to understand its contents and how to find images that are relevant for the task at hand? The use of a tight integration of automated methods, visualization and interaction, known as visual analytics, is probably the only way to go, combining the strengths of man and machine. An overview of trends in data visualization and visual analytics is given, and examples of recent work in multimedia analytics are presented. Exploiting meta-data, using interaction with relatively simple visual representations, and alignment with the work flow of users are promising routes, but scalability and evaluation are still challenging issues.

Bio

Jack (Jarke J.) van Wijk is full professor in visualization at the Department of Mathematics and Computer Science of Eindhoven University of Technology (TU/e). He received a MSc degree in industrial design engineering in 1982 and a PhD degree in computer science in 1986, both from Delft University of Technology. He has worked for ten years at the Netherlands Energy Research Foundation ECN. He joined Eindhoven University of Technology in 1998, where he became a full professor of visualization in 2001. His main research interests are information visualization and visual analytics, with a focus on the development of new methods for the interactive exploration of large data-sets. The work of his group has led to two start-up companies: MagnaView BV and SynerScope BV. He has (co-)authored more than 150 papers in visualization and computer graphics and received six best paper awards. He received the IEEE Visualization Technical Achievement Award in 2007 and the Eurographics 2013 Outstanding Technical Contributions Award.

Organising Committee

ACM Multimedia 2016

General Chairs

general.chairs@acmmm.org

Alan Hanjalic

Delft University of Technology

Cees Snoek

Qualcomm Research Netherlands / University of Amsterdam

Marcel Worrying

University of Amsterdam

Honorary Chair

honorary.chair@acmmm.org

Arnold Smeulders

University of Amsterdam

Program Coordinator

program.coordinator@acmmm.org

Dick Bulterman

CWI / VU

Program Chairs

technical.program@acmmm.org

Benoit Huet

EURECOM

Aisling Kelliher

Virginia Tech

Yiannis Kompatsiaris

CERTH-ITI

Jin Li

Microsoft

Local Arrangement Chairs

local.organisation@acmmm.org

Hayley Hung

Delft University of Technology

Cynthia Liem

Delft University of Technology

Author's Advocate

advocate@acmmm.org

Shervin Shirmohammadi

University of Ottawa

Workshop Chairs

workshop.chairs@acmmm.org

Lexing Xie

Australian National University

Alan Smeaton

Dublin City University

Brave New Idea Chairs

bni.chairs@acmmm.org

Martha Larson

Delft University of Technology / Radboud University Nijmegen

Hari Sundaram

University of Illinois

Tutorial Chairs

tutorial.chairs@acmmm.org

Rita Cucchiara

University of Modena

Susanne Boll

University of Oldenburg

Doctoral Symposium Chairs

doctoral.symposium@acmmm.org

Winston Hsu

National Taiwan University & IBM TJ Watson

Eckehard Steinbach

Technische Universität München

Panel Chairs

panel.chairs@acmmm.org

Alberto Del Bimbo

University of Florence

David Shamma

Yahoo Labs

Multimedia Grand Challenge Chairs

grand.challenge@acmmm.org

Xavier Anguera

ELSA Corp.

Shuicheng Yan

Qihoo 360 & NUS

Demo Chairs

technical.demo@acmmm.org

Pablo Cesar

CWI

Max Mühlhäuser

Technische Universität Darmstadt

Open Source Competition Chairs

open.source@acmmm.org

Tao Mei

Microsoft Research Asia

Christian Timmerer

Alpen-Adria-Universität Klagenfurt

Travel Grant Chairs

travel.grants@acmmm.org

Wessel Kraaij

Radboud University Nijmegen / TNO

Remco Veltkamp

Utrecht University

Video Chairs

video.program@acmmm.org

Marco Bertini

University of Florence

Shin'ichi Satoh

National Institute of Informatics

Interactive Arts Chairs

interactive.art@acmmm.org

Lucas Evers

Waag Society

Frank Nack

University of Amsterdam

Sponsor Chairs

sponsor.chairs@acmmm.org

Touradj Ebrahimi

EPFL

Gerald Friedland

ICSI

Publicity Chairs

publicity.chairs@acmmm.org

Carsten Griwodz

University of Oslo

Meng Wang

Hefei University of Technology

Publication Chairs

proceedings.chairs@acmmm.org

Chong-Wah Ngo

City University of Hong Kong

Balakrishnan Prabhakaran

University of Texas at Dallas

Web Chairs

web.chairs@acmmm.org

Stevan Rudinac

University of Amsterdam

Pascal Mettes

University of Amsterdam

History Preservation Chairs

history.chairs@acmmm.org

Yong Rui

Microsoft Research Asia

Heng Tao Shen

University of Queensland

Technical Program Committee

Area chairs

Theme

Engagement

Emotional and Social Signals in Multimedia

Julien Epps

University of New South Wales

Mohammad Soleymani

University of Geneva

Jean-Marc Odobez

IDIAP

Multimedia Search and Recommendation

Joemon Jose

Glasgow

Stéphane Marchand-Maillet

UNIGE

Michele Merler

IBM TJ Watson Research Center

Muriel Visani

University of La Rochelle

Zheng-Jun Zha

University of Science and Technology of China

Social Multimedia

Luca Maria Aiello

Yahoo Labs

Symeon Papadopoulos

CERTH-ITI

Jonathon Hare

University of Southampton

Theme Experience

Multimedia Art, Entertainment and Culture

Jinsil Hwaryoung Seo

Simon Fraser

David Ayman Shamma

Yahoo

Multimedia for Collaboration in Education & Distributed Environments

Matt Cooper

FXPAL

Multimedia HCI and Quality of Experience

Teresa Chambel

Lisbon University

Judith Redi

Delft

Bart Thomee

Yahoo

Music, Speech and Audio Processing in Multimedia

Brian Pardo

Northwestern

Theme System

Mobile Multimedia

Winston Hsu

National Taiwan University

Jochen Huber

MIT Media Lab

Multimedia Scalability and Management System

Xiansheng Hua

Alibaba

Lei Zhang

Microsoft

Multimedia Systems and Middleware

Jiangchuan Liu

Simon Fraser University

Wenwu Zhu

Tsinghua University

Roger Zimmerman

NUS

Multimedia Telepresence and Virtual/Augmented Reality

Pablo Cesar

CWI

Zicheng Liu

Microsoft

Multimedia Transport and Delivery

Ali Begen

CISCO

Christian Timmerer

Alpen-Adria-Universität Klagenfurt

Multimedia Transport and Delivery

Ali Begen

CISCO

Christian Timmerer

Alpen-Adria-Universität Klagenfurt

Theme

Understanding

Deep Learning for Multimedia

Ed Chang

HTC

Guo-Jun Qi

University of Central Florida

Elisa Ricci

FBK Trento

Efstratios Gavves

University of Amsterdam

Fiona Yan Liu

Hong Kong Polytechnic University

Multimedia and Vision

Francois Bremond

INRIA

Wen-Huang Cheng

Academia Sinica

Xavier Giró-i-Nieto

Universitat Politècnica de Catalunya

Yu-Gang Jiang

Fudan

Noel O'Connor

DCU

Maja Pantic

Imperial College

Qi Tian

University of Texas at San Antonio

Concetto Spampinato

Uni Catania

Multimodal Analysis and Description

Guillaume Gravier

IRISA

Alex Hauptmann

CMU

Ichiro Ide

Nagoya University

Sofia Tsekeridou

Intrasoft International, Research and Innovation Dept.

Invited Tutorials

Saturday, October 15

Location: Roeterseiland

To celebrate the special occasion that both the European Conference on Computer Vision (8 – 16 October) and the ACM Multimedia Conference (15 – 19 October) meet in Amsterdam, the University of Amsterdam offers to registered attendees of either conference the opportunity to attend a maximum of two Invited tutorials by a renowned scientist in vision and/or multimedia on Saturday October 15th, free of charge.

The Invited tutorials are:

- 9:30 – 10:30** **IT11** / Room C1.03
Segmentation and Tracking, Mubarak Shah
- 9:30 – 10:30** **IT12** / Room C0.01
Variational wrapping deformations for image processing applications, Olga Sorkine-Hornung
- 10:45 – 11:45** **IT21** / Room C1.03
Developments in deep learning, Yoshua Bengio
- 10:45 – 11:45** **IT22** / Room C0.01
Multimodal human-robot interaction, Elisabeth Andre
- 11:45 – 13:00** **Break**
- 13:00 – 14:00** **IT31** / Room C1.03
Output embedding for large-scale visual recognition, Florent Perronnin
- 13:00 – 14:00** **IT32** / Room C0.01
Video event detection, localization and cross-media linking, Shih-Fu Chang
- 14:15 – 15:15** **IT41** / Room C1.03
Human body shape modelling, Michael Black
- 14:15 – 15:15** **IT42** / Room C0.01
A guide to creating and managing lifelog collections, Cathal Gurrin
- 15:30 – 16:30** **IT51** / Room C1.03
Deep visual understanding from deep learning, Jitendra Malik
- 15:30 – 16:30** **IT52** / Room C0.01
Bridging video and language with deep learning, Tao Mei
- 16:45 – 17:45** **IT61** / Room C0.01
Color and spectral information in modern computer vision, Sabine Süsstrunk

Workshops and Tutorials

Sunday, October 16

Location: Roeterseiland

09:00 – 12:30 Workshops and Tutorials

ECCV Workshop on Computer VISION for ART Analysis

Room M1.03

ECCV Workshop on Computer Vision for Audio-visual Media

Room C0.02

ECCV Workshop on Storytelling with Images and Video and Large Scale Movie Description and Understanding Challenge

Room C1.04

Workshop on Vision and Language Integration Meets Multimedia Fusion

Room M1.02

ECCV Workshop on Virtual/Augmented Reality for Visual Artificial Intelligence

Room C1.03

Workshop on Multimedia Alternate Realities

Room C2.06

Workshop on Human Behavior Understanding

Room C0.01

Workshop on Multimedia Assisted Dietary Management

Room C2.05

Workshop on Multimedia for personal health and health care

Room C2.03

Workshop on Lifelogging Tools and Applications

Room C2.04

Audio/Visual Emotion Challenge and Workshop

Room C2.02

Multimedia COMMONS Workshop 2016

Room C2.01

Tutorial on Social and Affective Robotics

Room C3.01

Tutorial on Technology and Art in Stimulating Creative Placemaking in Public-Use Spaces

Room C2.07

12:30 – 14:00 Break

14:00 – 17:00 Workshops and Tutorials

ECCV Workshop on Computer VISION for ART Analysis

Room M1.03

Tutorial on Emerging topics in learning from noisy and missing data

Room C0.02

ECCV Workshop on Storytelling with Images and Video and Large Scale Movie Description and Understanding Challenge

Room C1.04

Workshop on Vision and Language Integration Meets Multimedia Fusion

Room M1.02

Workshop on Multimedia Alternate Realities

Room C2.06

Workshop on Human Behavior Understanding

Room C0.01

Workshop on Multimedia Assisted Dietary Management

Room C2.05

Workshop on Multimedia for personal health and health care

Room C2.03

Workshop on Lifelogging Tools and Applications

Room C2.04

Audio/Visual Emotion Challenge and Workshop

Room C2.02

Multimedia COMMONS Workshop 2016

Room C2.01

Tutorial on Social and Affective Robotics

Room C3.01

Tutorial on Multimedia Privacy

Room C2.07

Tutorial on Situation Recognition from Multimodal Data

Room C3.02

Tutorial on The Lifecycle of Geotagged Multimedia Data

Room C3.03

Workshops, Tutorials and Full Papers**Monday, October 17**

Location: Tuschinski / De Bazel

08:45 – 09:00 Keynote I: Dirk Helbing (Zaal 1)**Session chair:** Alan Hanjalic (Delft University of Technology)**A Digital World to Thrive In — How the Internet of Things Can Make the “Invisible Hand” Work**

Dirk Helbing (ETH Zurich)

10:30 – 11:00 Break**11:00 – 12:30 Best paper session (Zaal 1)****Session chair:** Benoit Huet (Eurecom)**1 Multi-modal Multi-view Topic-opinion Mining for Social Event Analysis**

Shengsheng Qian (National Lab of Pattern Recognition, Institute of Automation, CAS), Tianzhu Zhang (National Lab of Pattern Recognition, Institute of Automation, CAS), Changsheng Xu (National Lab of Pattern Recognition, Institute of Automation, CAS)

2 Patterns of Free-form Web Curation: Visual Thinking and Expression of Relationships

Nic Lupfer (Texas A&M University), Andruid Kerne (Texas A&M University), Andrew M. Webb (Texas A&M University), Rhema Linder (Texas A&M University)

3 DASH2: Exploring HTTP/2 for Internet Streaming to Mobile Device

Mengbai Xiao (George Mason University), Viswanathan Swaminathan (Adobe Systems Inc.), Sheng Wei (University of Nebraska-Lincoln), Songqing Chen (George Mason University)

4 Deep-based Ingredient Recognition for Cooking Recipe Retrieval

Jingjing Chen (City university of HongKong), Chong-wah NGO (City university of HongKong)

12:30 – 14:00 Lunch break**12:30 – 14:00 ToMM meeting (Zaal 3)****On invitation only****12:30 – 14:00 Women in SIGMM lunch (VIP)****On invitation only****De Bazel****Multimedia Technolog Expo****14:00 – 17:00 Poster session I and II (Lobby)****Session chair:** Ichiro Ide (Nagoya University), Yiannis Kompatsiaris (CERTH), Judith Redi (Delft University of Technology)**5 GeoTracks: Adaptive Music for Everyday Journeys**

Chris Greenhalgh (University of Nottingham), Adrian Hazzard (University of Nottingham), Sean McGrath (University of Nottingham), Steve Benford (University of Nottingham)

6 Abnormal Event Discovery in User Generated Photos

Xiaoshan Yang (Institute of Automation, Chinese Academy of Sciences), Tianzhu Zhang (Institute of Automation, Chinese Academy of Sciences), Changsheng Xu (Institute of Automation, Chinese Academy of Sciences)

- 7 **Deep Bi-directional Cross-triplet Embedding for Cross-Domain Clothing Retrieval**
Shuhui Jiang (Northeastern University), Yue Wu (Northeastern University), Yun Fu (Northeastern University)
- 8 **A Discriminative and Compact Audio Representation for Event Detection**
Liping Jing (Beijing Jiaotong University), Bo Liu (Beijing Jiaotong University), Jaeyoung Choi (International Computer Science Institute & Delft University of Technology, Delft, Netherlands), Adam Janin (International Computer Science Institute), Julia Bernd (International Computer Science Institute), Michael W. Mahoney (International Computer Science Institute & University of California), Gerald Friedland (International Computer Science Institute & University of California)
- 9 **Jockey Time: Making Video Playback to Enhance Emotional Effect**
Kyeong Ah Jeong (Korea Advanced Institute of Science and Technology), Hyeon-Jeong Suk (Korea Advanced Institute of Science and Technology)
- 10 **Discriminative Paired Dictionary Learning for Visual Recognition**
Hui-Hung Wang (National Chung Cheng University), Yi-Ling Chen (National Taiwan University), Chen-Kuo Chiang (National Chung Cheng University)
- 11 **From Seed Discovery to Deep Reconstruction: Predicting Saliency in Crowd via Deep Networks**
Yanhao Zhang (Harbin Institute of Technology), Lei Qin (CAS), Qingming Huang (CAS & Harbin Institute of Technology), Kuiyuan Yang (Microsoft Research Aisa), Jun Zhang (Hefei University of Technology), Hongxun Yao (Harbin Institute of Technology)
- 12 **Facial Age Estimation Using Robust Label Distribution**
Ke Chen (Tampere University of Technology), Joni-Kristian Kämäräinen (Tampere University of Technology), Zhaoxiang Zhang (Chinese Academy of Sciences)
- 13 **What Makes a Good Movie Trailer? Interpretation from Simultaneous EEG and Eyetracker Recording**
Sidi Liu (The University of Georgia), Jinglei Lv (The University of Georgia), Yimin Hou (Northeast Dianli University), Ting Shoemaker (The University of Georgia), Qinglin Dong (The University of Georgia), Kaiming Li (West China Hospital of Sichuan Univerity), Tianming Liu (The University of Georgia)
- 14 **LIME: A Method for Low-light IMage Enhancement**
Xiaojie Guo (Institute of Information Engineering, Chinese Academy of Sciences)
- 15 **Multi-Protocol Video Delivery with Late Trans-Muxing**
Rufael Mekuria (Unified Streaming B.V), Jelte Fennema (University of Amsterdam), Dirk Griffioen (Unified Streaming B.V)
- 16 **Analyzing Structural Characteristics of Object Category Representations From Their Semantic-part Distributions**
Ravi Kiran Sarvadevabhatla (Indian Institute of Science), Venkatesh Babu R (Indian Institute of Science)
- 17 **Action Recognition Based on Joint Trajectory Maps Using Convolutional Neural Networks**
Pichao Wang (University of Wollongong), Zhaoyang Li (Tianjin University), Yonghong Hou (Tianjin University), Wanqing Li (University of Wollongong)
- 18 **Efficient Digital Holographic Image Reconstruction on Mobile Devices**
Chung-Hua Chu (National Taichung University Of Science And Technology)
- 19 **Scene Image Synthesis from Natural Sentences Using Hierarchical Syntactic Analysis**
Tetsuaki Mano Mano (The University of Tokyo), Hiroaki Yamane Yamane (The University of Tokyo), Tatsuya Harada Harada (The University of Tokyo)
- 20 **A Fast 3D Retrieval Algorithm via Class-Statistic and Pair-Constraint Model**
Zan Gao (Tianjin University of Technology), Deyu Wang (Tianjin University of Technology), Hua Zhang (Tianjin University of Technology), Yanbing Xue (Tianjin University of Technology), Guangping Xu (Tianjin University of Technology)
- 21 **Analyzing and Predicting GIF Interestingness**
Michael Gygli (ETH Zurich), Mohammad Soleymani (University of Geneva)
- 22 **Emotion in Context: Deep Semantic Feature Fusion for Video Emotion Recognition**
Chen Chen (Fudan University), Zuxuan Wu (Fudan University), Yu-Gang Jiang (Fudan University)
- 23 **Exploiting Hierarchical Activations of Neural Network for Image Retrieval**
Ying Li (Dalian University of Technology), Xiangwei Kong (Dalian University of Technology), Liang Zheng (University of Texas at San Antonio), Qi Tian (University of Texas at San Antonio)
- 24 **A Deeply-Supervised Deconvolutional Network for Horizon Line Detection**
Lorenzo Porzi (Fondazione Bruno Kessler), Samuel Rota Bulò (Fondazione Bruno Kessler), Elisa Ricci (Fondazione Bruno Kessler)
- 25 **Exploiting Objects with LSTMs for Video Categorization**
Yongqing Sun (NTT Media Intelligence Laboratories), Zuxuan Wu (Fudan University), Xi Wang (Fudan University), Hiroyuki Arai (NTT Media Intelligence Laboratories), Tetsuya Kinebuchi (NTT Media Intelligence Laboratories), Yu-Gang Jiang (Fudan University)
- 26 **Assessing 3D Scan Quality Through Paired-comparisons Psychophysics**
Jacob Thorn (University College London), Rodrigo Pizarro (Universitat de Barcelona), Bernhard Spanlang (Universitat de Barcelona), Pablo Bermell-Garcia (Airbus Group), Mar Gonzalez-Franco (Airbus Group)
- 27 **Partial Multi-Modal Sparse Coding via Adaptive Similarity Structure Regularization**
Zhou Zhao (Zhejiang University), Hanqing Lu (Zhejiang University), Cai Deng (Zhejiang University), Xiaofei He (Zhejiang University), Yueting Zhuang (Zhejiang University)
- 28 **Improving Speaker Diarization of TV Series using Talking-Face Detection and Clustering**
Hervé Bredin (Université Paris-Saclay), Grégory Gelly (Paris-Sud, Université Paris-Saclay)
- 29 **Location-Independent WiFi Action Recognition via Vision-based Methods**
Jen-Yin Chang (National Taiwan University), Kuan-Ying Lee (National Taiwan University), Yu-Lin Wei (National Taiwan University), Kate Ching-Ju Lin (National Chiao Tung University), Winston Hsu (National Taiwan University)
- 30 **INRS Audiovisual Quality Dataset**
Edip Demirbilek (Institut National de la Recherche Scientifique), Jean-Charles Grégoire (Institut National de la Recherche Scientifique)
- 31 **Learning to Make Better Mistakes: Semantics-aware Visual Food Recognition**
Hui Wu (IBM Thomas J. Watson Research Center), Michele Merler (IBM Thomas J. Watson Research Center), Rosario Uceda-Sosa (IBM Thomas J. Watson Research Center), John R Smith (IBM Thomas J. Watson Research Center)
- 32 **Dictionary Learning Based Hashing for Cross-Modal Retrieval**
Xin-Shun Xu (Shandong University)
- 33 **SocialFX: Studying a Crowdsourced Folksonomy of Audio Effects Terms**
Taylor Zheng (Northwestern University), Prem Seetharaman (Northwestern University), Bryan Pardo (Northwestern University)
- 34 **SwiDeN : Convolutional Neural Networks For Depiction Invariant Object Recognition**
Ravi Kiran Sarvadevabhatla (Indian Institute of Science), Shiv Surya (Indian Institute of Science), Srinivas S S Kruthiventi (Indian Institute of Science), Venkatesh Babu R. (Indian Institute of Science)
- 35 **Multi-Scale Triplet CNN for Person Re-Identification**
Jiawei Liu (University of Science and Technology of China), Zheng-Jun Zha (University of Science and Technology of China), Qi Tian (University of Texas at San Antonio), Dong Liu (University of Science and Technology of China), Ting Yao (Microsoft Research Asia), Qiang Ling (University of Science and Technology of China), Tao Mei (Microsoft Research Asia)

- 36 **Multimodal Popularity Prediction of Brand-related Social Media Posts**
Masoud Mazloom (University of Amsterdam), Robert Rietveld (University of Amsterdam), Stevan Rudinac (University of Amsterdam), Marcel Worrying (University of Amsterdam), Willemijn van Dolen (University of Amsterdam)
- 37 **Learning Multimodal Temporal Representation for Dubbing Detection in Broadcast Media**
Nam Le (Idiap Research Institute & Ecole Polytechnique Federal de Lausanne), Jean-Marc Odobez (Idiap Research Institute & Ecole Polytechnique Federal de Lausanne)
- 38 **Joint Image-Text Representation by Gaussian Visual-Semantic Embedding**
Zhou Ren (University of California, Los Angeles), Hailin Jin (Adobe Research), Zhe Lin (Adobe Research), Chen Fang (Adobe Research), Alan Yuille (Adobe Research)
- 39 **A Domain Robust Approach For Image Dataset Construction**
Yazhou Yao (Nanjing University of Science and Technology), Xian-sheng Hua (Alibaba Group), Fumin Shen (University of Electronic Science and Technology of China), Jian Zhang (University of Technology Sydney), Zhenmin Tang (Nanjing University of Science and Technology)
- 40 **A Supervised Approach for Text Illustration**
Harsh Jhamtani (Adobe Systems), Shubham Varma (IIT Varanasi), Midhun Gundapuneni (IIT Kharagpur), Siddhartha Kumar Dutta (IIT Bombay)
- 41 **Learning Music Emotion Primitives via Supervised Dynamic Clustering**
Yang Liu (Hong Kong Baptist University), Yan Liu (The Hong Kong Polytechnic University), Xiang Zhang (The Hong Kong Polytechnic University), Gong Chen (The Hong Kong Polytechnic University), Kejun Zhang (Zhejiang University)
- 42 **Cross-modal Retrieval by Real Label Partial Least Squares**
Jianfeng He (University of Chinese Academy of Sciences), Bingpeng Ma (University of Chinese Academy of Sciences), Shuhui Wang (Institute of Computing Technology, Chinese Academy of Sciences), Yugui Liu (University of Chinese Academy of Sciences), Qingming Huang (University of Chinese Academy of Sciences)
- 43 **LSOD: Local Sparse Orthogonal Descriptor for Image Matching**
Yiru Zhao (Shanghai Jiao Tong University), Yaoyi Li (Shanghai Jiao Tong University), Zhiwen Shao (Shanghai Jiao Tong University), Hongtao Lu (Shanghai Jiao Tong University)
- 44 **Frustratingly Easy Cross-Modal Hashing**
Dekui Ma (Dalian University of Technology), Jian Liang (National Laboratory of Pattern Recognition, CASIA), Xiangwei Kong (Dalian University of Technology), Ran He (National Laboratory of Pattern Recognition, CASIA)
- 45 **Families in the Wild (FIW): Large-Scale Kinship Image Database and Benchmarks**
Joseph P. Robinson (Northeastern University), Ming Shao (Northeastern University), Yue Wu (Northeastern University), Yun Fu (Northeastern University)
- 46 **Enabling My Robot To Play Pictionary : Recurrent Neural Networks For Sketch Recognition**
Ravi Kiran Sarvadevabhatla (Indian Institute of Science), Jogendra Kundu (Indian Institute of Science), Venkatesh Babu R (Indian Institute of Science)
- 47 **Experience Individualization on Online TV Platforms through Persona-based Account Decomposition**
Payal Bajaj (Adobe Systems), Sumit Shekhar (Adobe Systems)
- 48 **Improved Dense Trajectory with Cross Streams**
Katsunori Ohnishi (The University of Tokyo), Masatoshi Hidaka (The University of Tokyo), Tatsuya Harada (The University of Tokyo)
- 49 **Joint Image and Text Representation for Aesthetics Analysis**
Ye Zhou (Fudan University), Xin Lu (Adobe Systems Inc.), Junping Zhang (Fudan University), James Z. Wang (The Pennsylvania State University)
- 50 **Who is where? Matching People in Video to Wearable Acceleration During Crowded Mingling Events**
Laura Cabrera-Quiros (TU Delft), Hayley Hung (Delft University of Technology)
- 51 **Supervised Recurrent Hashing for Large Scale Video Retrieval**
Yun Gu (Shanghai Jiao Tong University), Chao Ma (Shanghai Jiao Tong University), Jie Yang (Shanghai Jiao Tong University)
- 52 **Adaptation of Word Vectors using Tree Structure for Visual Semantics**
Nakamasa Inoue (Tokyo Institute of Technology), Koichi Shinoda (Tokyo Institute of Technology)
- 53 **Adaptive Bitrate Selection for Video Encoding with Reduced Block Artifacts**
Min-Kook Choi (Inha University), Hyun-Gyu Lee (Inha University), Minseok Song (Inha University), Sang-Chul Lee (Inha University)
- 54 **What Makes Photo Cultures Different?**
Miriam Redi (Yahoo), Damon Crockett (UCSD), Lev Manovich (CUNY), Simon Osindero (Flickr)
- 55 **Synchronization among Groups of Spectators for Highlight Detection in Movies**
Michal Muszynski (University of Geneva), Theodoros Kostoulas (University of Geneva), Patrizia Lombardo (University of Geneva), Thierry Pun (University of Geneva), Guillaume Chanel (University of Geneva)
- 56 **On Estimating Air Pollution from Photos Using Convolutional Neural Network**
Chao Zhang (Beijing Normal University & IBM Research – China), Junchi Yan (East China Normal University & IBM Research – China), Changsheng Li (IBM Research – China), Xiaoguang Rui (IBM Research – China), Liang Liu (IBM Research – China), Rongfang Bie (Beijing Normal University)
- 57 **Cross-modal Retrieval with Label Completion**
Xing Xu (University of Electronic Science and Technology of China), Fumin Shen (University of Electronic Science and Technology of China), Yang Yang (University of Electronic Science and Technology of China), Heng Tao Shen (The University of Queensland), Li He (Qualcomm R&D Center), Jingkuan Song (University of Trento)
- 58 **Objectness-aware Semantic Segmentation**
Yuhang Wang (National Laboratory of Pattern Recognition, CASIA & University of Chinese Academy of Sciences), Jing Liu (National Laboratory of Pattern Recognition, CASIA), Yong Li (National Laboratory of Pattern Recognition, CASIA & University of Chinese Academy of Sciences), Junjie Yan (SenseTime Group Limited), Hanqing Lu (National Laboratory of Pattern Recognition, CASIA)
- 59 **ReadMe: A Real-Time Recommendation System for Mobile Augmented Reality Ecosystems**
Dimitris Chatzopoulos (The Hong Kong University of Science and Technology), Pan Hui (The Hong Kong University of Science and Technology)
- 60 **Action Recognition Using Local Consistent Group Sparse Coding with Spatio-Temporal Structure**
Yi Tian (Beijing Jiaotong University), Qiuqi Ruan (Beijing Jiaotong University), Gaoyun An (Beijing Jiaotong University), Yun Fu (Northeastern University)
- 61 **Super Resolution of the Partial Pixelated Images With Deep Convolutional Neural Network**
Haiyi Mao (Northeastern University), Yue Wu (Northeastern University), Jun Li (Northeastern University), Yun Fu (Northeastern University)
- 62 **Adaptive Visual Feedback Generation for Facial Expression Improvement with Multi-task Deep Neural Networks**
Takuhiro Kaneko (NTT Corporation), Kaoru Hiramatsu (NTT Corporation), Kunio Kashino (NTT Corporation)

- 63 **Fast Supervised LDA for Discovering Micro-Events in Large-Scale Video Datasets**
Angelos Katharopoulos (Aristotle University of Thessaloniki), Despoina Paschalidou (Aristotle University of Thessaloniki), Christos Diou (Aristotle University of Thessaloniki), Anastasios Delopoulos (Aristotle University of Thessaloniki)
- 64 **Semantic Description of Timbral Transformations in Music Production**
Ryan Stables (Birmingham City University), Brecht De Man (Queen Mary University of London), Sean Enderby (Birmingham City University), Joshua D Reiss (Queen Mary University of London), Gyo?rgy Fazekas (Queen Mary University of London), Thomas Wilmering (Queen Mary University of London)
- 65 **Multimodal Learning via Exploring Deep Semantic Similarity**
Di Hu (Northwestern Polytechnical University), Xiaoqiang Lu (Chinese Academy of Sciences), Xuelong Li (Chinese Academy of Sciences)
- 66 **Multi-pose Facial Expression Recognition Using Transformed Dirichlet Process**
Feifei Zhang (Jiangsu University), Qirong Mao (Jiangsu University), Ming Dong (Wayne State University), Yongzhao Zhan (Jiangsu University)
- 67 **Neighborhood-Preserving Hashing for Large-Scale Cross-Modal Search**
Botong Wu (Peking University), Yizhou Wang (Peking University)
- 68 **Attention-based LSTM with Semantic Consistency for Videos Captioning**
Zhao Guo (University of Electronic Science and Technology of China), Lianli gao (University of Electronic Science and Technology of China), jingquan song (Columbia University), Xing Xu (University of Electronic Science and Technology of China), Jie Shao (University of Electronic Science and Technology of China), Heng Tao Shen (The University of Queensland)
- 69 **Efficient Mobile Implementation of A CNN-based Object Recognition System**
Keiji Yanai (The University of Electro-Communications, Tokyo), Ryosuke Tanno (The University of Electro-Communications, Tokyo), Koichi Okamoto (The University of Electro-Communications, Tokyo)
- 70 **Context-aware Geometric Object Reconstruction for Mobile Education**
Jinxin Zheng (Peking University), Yongtao Wang (Peking University), Zhi Tang (Peking University)
- 71 **Automatic Music Video Generation Based on Emotion-Oriented Pseudo Song Prediction and Matching**
Jen-Chun Lin (Academia Sinica), Wen-Li Wei (Academia Sinica), Hsin-Min Wang (Academia Sinica)
- 72 **Novel Word Embedding and Translation-based Language Modeling for Extractive Speech Summarization**
Kuan-Yu Chen (Academia Sinica), Shih-Hung Liu (Academia Sinica), Berlin Chen (National Taiwan Normal University), Hsin-Min Wang (Academia Sinica), Hsin-Hsi Chen (National Taiwan University)
- 73 **Micro-Expression Recognition with Expression-State Constrained Spatio-Temporal Feature Representations**
Dae Hoe Kim (korea advanced institute of science and technology), Wissam J. Baddar (korea advanced institute of science and technology), Yong Man Ro (korea advanced institute of science and technology)
- 74 **Multimodal Interest Level Estimation via Variational Bayesian Mixture of Robust CCA**
Yuma Sasaka (Hokkaido University), Takahiro Ogawa (Hokkaido University), Miki Haseyama (Hokkaido University)
- 75 **Transportation Mode Detection on Mobile Devices Using Recurrent Nets**
Toan H Vu (National Central University), Le Dung (Hanoi University of Science and Technology), Jia-Ching Wang (National Central University)
- 76 **Deeply-Supervised Recurrent Convolutional Neural Network for Saliency Detection**
Youbao Tang (Harbin Institute of Technology), Xiangqian Wu (Harbin Institute of Technology), Wei Bu (Harbin Institute of Technology)
- 77 **Deep Correlation Features for Image Style Classification**
Wei-Ta Chu (National Chung Cheng University), Yi-Ling Wu (National Chung Cheng University)
- 78 **CNN vs. SIFT for Image Retrieval: Alternative or Complementary?**
Ke Yan (Peking University), Yaowei Wang (Beijing Institute of Technology), Dawei Liang (Peking University), Tiejun Huang (Peking University), Yonghong Tian (Peking University)
- 79 **Looking Good With Flickr Favos: Gaussian Processes for Finding Difference Makers in Personality Impressions**
Xiaoyu Xiong (University of Glasgow), Maurizio Filippone (EURECOM), Alessandro Vinciarelli (University of Glasgow)
- 80 **Ad Recommendation for Sponsored Search Engine via Composite Long-Short Term Memory**
Dejiang Kong (Zhejiang Univerisity), Fei Wu (Zhejiang University), Siliang Tang (Zhejiang University), Yueting Zhuang (Zhejiang University)
- 81 **Learning a Multi-class Discriminative Dictionary with Nonredundancy Constraints for Visual Classification**
Zhao Liu (Beijing Institute of Technology), Yuwei Wu (Beijing Institute of Technology), Junsong Yuan (Nanyang Technological University), Yap-peng Tan (Nanyang Technological University)
- 82 **A Compact Binary Aggregated Descriptor via Dual Selection for Visual Search**
Yuwei Wu (Nanyang Technological University), Zhe Wang (Nanyang Technological University), Junsong Yuan (Nanyang Technological University), Lingyu Duan (Peking University)
- 83 **Capped Lp-Norm Graph Embedding for Photo Clustering**
Mengfan Tang (University of California, Irvine), Feiping Nie (Northwestern Polytechnical University), Ramesh Jain (University of California, Irvine)
- 84 **Bidirectional Long-Short Term Memory for Video Description**
Yi Bin (University of Electronic Science and Technology of China), Yang Yang (University of Electronic Science and Technology of China), Fumin Shen (University of Electronic Science and Technology of China), Xing Xu (University of Electronic Science and Technology of China), Heng Tao Shen (The University of Queensland)
- 85 **A Robust Distance with Correlated Metric Learning for Multi-Instance Multi-Label Data**
Yashaswi Verma (IIIT Hyderabad (India)), C. V. Jawahar (IIIT Hyderabad (India))
- 86 **Multiview Video Super-Resolution via Information Extraction and Merging**
Yawei Li (University of Electronic Science and Technology of China), Xiaofeng Li (University of Electronic Science and Technology of China), Zhizhong Fu (University of Electronic Science and Technology of China), Wenli Zhong (University of Electronic Science and Technology of China)
- 87 **InnerView: Learning Place Ambiance from Social Media Images**
Darshan Santani (Idiap Research Institute and EPFL), Rui Hu (Idiap Research Institute), Daniel Gatica-Perez (Idiap Research Institute and EPFL)
- 88 **Quartet-net Learning for Visual Instance Retrieval**
Jiewei Cao (The University of Queensland), Zi Huang (The University of Queensland), Peng Wang (The University of Queensland), Chao Li (The University of Queensland), Xiaoshuai Sun (The University of Queensland), Heng Tao Shen (The University of Queensland)
- 89 **AKSDA-MSVM: A GPU-accelerated Multiclass Learning Framework for Multimedia**
Stavros Arestis-Chartampilas (CERTH-ITI), Nikolaos Gkalelis (CERTH-ITI), Vasileios Mezaris (CERTH-ITI)
- 90 **Automatic Reflection Removal using Gradient Intensity and Motion Cues**
Chao Sun (University of Electronic Science and Technology of China), Shuaicheng Liu (University of Electronic Science and Technology of China), Taotao Yang (University of Electronic Science and Technology of China), Bing Zeng (University of Electronic Science and Technology of China)

- of China), Zhengning Wang (University of Electronic Science and Technology of China), Guanghui Liu (University of Electronic Science and Technology of China)
- 91 **Personal Multi-view Viewpoint Recommendation based on Trajectory Distribution of the Viewing Target**
Xueting Wang (Nagoya University), Kensho Hara (Nagoya University), Yu Enokibori (Nagoya University), Takatsugu Hirayama (Nagoya University), Kenji Mase (Nagoya University)
- 92 **Motion Segmentation using Visual and Bio-mechanical Features**
Stefano Alletto (University of Modena and Reggio Emilia), Giuseppe Serra (University of Modena and Reggio Emilia), Rita Cucchiara (University of Modena and Reggio Emilia)
- 93 **Locality-preserving K-SVD Based Joint Dictionary and Classifier Learning for Object Recognition**
Yuan-Shan Lee (National Central University), Chien-Yao Wang (National Central University), Seksan Mathulapransan (National Central University), Jia-Hao Zhao (National Central University), Jia-Ching Wang (National Central University)
- 94 **Label Tree Embeddings for Acoustic Scene Classification**
Huy Phan (University of Lübeck), Lars Hertel (University of Lübeck), Marco Maass (University of Lübeck), Philipp Koch (University of Lübeck), Alfred Mertins (University of Lübeck)
- 95 **Deep Learning for Image Memorability Prediction: the Emotional Bias**
Yoann Baveye (Université de Nantes, France), Romain Cohendet (Université de Nantes, France), Matthieu Perreira Da Silva (Université de Nantes, France), Patrick Le Callet (Université de Nantes, France)
- 96 **Demand-adaptive Clothing Image Retrieval Using Hybrid Topic Model**
Zhengzhong Zhou (MOE-Microsoft Key Laboratory for Intelligent Computing and Intelligent Systems), Jingjin Zhou (MOE-Microsoft Key Laboratory for Intelligent Computing and Intelligent Systems), Liqing Zhang (MOE-Microsoft Key Laboratory for Intelligent Computing and Intelligent Systems)
- 97 **Deep Multi-task Learning with Label Correlation Constraint for Video Concept Detection**
Foteini Markatopoulou (Centre for Research and Technology Hellas (CERTH), Information Technologies Institute (ITI)), Vasileios Mezaris (Centre for Research and Technology Hellas (CERTH)), Information Technologies Institute (ITI)), Ioannis Patras (Queen Mary University of London)
- 98 **Application-Layer Rate-Adaptive Multicast Video Streaming over 802.11 for Mobile Devices**
Raheeb Muzaffar (Alpen-Adria-Universität Klagenfurt and Queen Mary University of London), Evsen Yanmaz (Lakeside Labs GmbH), Christian Bettstetter (Alpen-Adria-Universität Klagenfurt), Andrea Cavallaro (Queen Mary University of London)
- 99 **Scalable Compression of Deep Neural Networks**
Xing Wang (Simon Fraser University & AltumView Systems Inc.), Jie Liang (Simon Fraser University & AltumView Systems Inc.,)
- 100 **UnitBox: An Advanced Object Detection Network**
Jiahui Yu (University of Illinois at Urbana-Champaign), Yuning Jiang (Megvii Inc.), Zhangyang Wang (University of Illinois at Urbana-Champaign), Zhimin Cao (Megvii Inc.), Thomas Huang (University of Illinois at Urbana-Champaign)
- 101 **Alone versus In-a-group: A Comparative Analysis of Facial Affect Recognition**
Wenxuan Mou (Queen Mary University of London), Hatice Gunes (University of Cambridge), Ioannis Patras (Queen Mary University of London)
- 102 **Local Diffusion Map Signature for Symmetry-aware Non-rigid Shape Correspondence**
Meng Wang (New York University Abu Dhabi & New York University), Yi Fang (New York University Abu Dhabi & New York University)

- 103 **How Cosmopolitan Are Emojis?**
Francesco Barbieri (Universitat Pompeu Fabra, Spain), German Kruszewski (University of Trento), Francesco Ronzano (Universitat Pompeu Fabra, Spain), Horacio Saggion (Universitat Pompeu Fabra, Spain)
- 104 **Online Weighted Clustering for Real-time Abnormal Event Detection in Video Surveillance**
Hanhe Lin (University of Otago), Jeremiah D Deng (University of Otago), Brendon J Woodford (University of Otago), Ahmad Shahi (University of Otago)
- 105 **Accelerating Convolutional Neural Networks for Mobile Applications**
Peisong Wang (National Laboratory of Pattern Recognition, CASIA & University of Chinese Academy of Sciences), Jian Cheng (National Laboratory of Pattern Recognition, CASIA & University of Chinese Academy of Sciences)
- 106 **News Program Detection in TV Broadcast Videos**
Raghendra Kannao (Indian Institute of Technology Guwahati), Durgaprasad Dandi (Indian Institute of Technology Guwahati), Swamy Yellapu (Indian Institute of Technology Guwahati), Prithwiji Guha (Indian Institute of Technology Guwahati)
- 107 **Detecting Arbitrary Oriented Text in the Wild with a Visual Attention Model**
Wenyi Huang (The Pennsylvania State University), Dafang He (The Pennsylvania State University), Xiao Yang (The Pennsylvania State University), Zihan Zhou (The Pennsylvania State University), Daniel Kifer (The Pennsylvania State University), C. Lee Giles (The Pennsylvania State University)
- 108 **Global Consistent Shape Correspondence for Efficient and Effective Active Shape Models**
Meng Wang (New York University Abu Dhabi & New York University), Yi Fang (New York University Abu Dhabi & New York University)
- 109 **Towards Ultra-Low-Bitrate Video Conferencing Using Facial Landmarks**
Pin-Chun Wang (National Tsing Hua University), Ching-Ling Fan (National Tsing Hua University), Chun-Ying Huang (National Chiao Tung University), Kuan-Ta Chen (Academia Sinica), Cheng-Hsin Hsu (National Tsing Hua University)
- 110 **Generating Diverse Image Datasets with Limited Labeling**
Niluthpol Chowdhury Mithun (University of California, Riverside), Rameswar Panda (University of California, Riverside), Amit K. Roy-Chowdhury (University of California, Riverside)
- 111 **Multi-modal Conditional Attention Fusion for Dimensional Emotion Prediction**
Shizhe Chen (Renmin University of China), Qin Jin (Renmin University of China)
- 112 **Video Generation Using 3D Convolutional Neural Network**
Shohei Yamamoto (The University of Tokyo), Tatsuya Harada (The University of Tokyo)
- 113 **Processing-Aware Privacy-Preserving Photo Sharing over Online Social Networks**
Weiwei Sun (University of Macau), Jiantao Zhou (University of Macau), Ran Lyu (University of Macau), Shuyuan Zhu (UESTC)
- 114 **Detecting Violence in Video using Subclasses**
Xirong Li (Renmin University of China), Yujia Huo (Renmin University of China), Qin Jin (Renmin University of China), Jieping Xu (Renmin University of China)
- 115 **Deep Representation for Abnormal Event Detection in Crowded Scenes**
Yachuang Feng (Xi'an Institute of Optics and Precision Mechanics, Chinese Academy of Sciences & University of Chinese Academy of Sciences), Yuan Yuan (Xi'an Institute of Optics and Precision Mechanics, Chinese Academy of Sciences), Xiaoqiang Lu (Xi'an Institute of Optics and Precision Mechanics, Chinese Academy of Sciences)
- 116 **Exploration of Large Image Corpuses in Virtual Reality**
Sanket Khanwalkar (University of California, Irvine), Shonali Balakrishna (University of California, Irvine), Ramesh Jain (University of California, Irvine)

- 117 **HEVC-compliant Tile-based Streaming of Panoramic Video for Virtual Reality Applications**
Alireza Zare (Tampere University of Technology), Alireza Aminlou (Nokia Technologies), Miska M. Hannuksela (Nokia Technologies), Moncef Gabbouj (Tampere University of Technology)
- 118 **MatchDR: Image Correspondence by Leveraging Distance Ratio Constraint**
Rui Wang (State Key Laboratory of Information Security, Institute of Information Engineering, Chinese Academy of Sciences), Dong Liang (State Key Laboratory of Information Security, Institute of Information Engineering, Chinese Academy of Sciences), Wei Zhang (State Key Laboratory of Information Security, Institute of Information Engineering, Chinese Academy of Sciences), Xiaochun Cao (State Key Laboratory of Information Security, Institute of Information Engineering, Chinese Academy of Sciences)
- 119 **A Novel Shadow-Free Feature Extractor for Real-Time Road Detection**
Zhenqiang Ying (Shenzhen Graduate School, Peking University), Ge Li (Shenzhen Graduate School, Peking University), Xianghao Zang (Shenzhen Graduate School, Peking University), Ronggang Wang (Shenzhen Graduate School, Peking University), Wenmin Wang (Shenzhen Graduate School, Peking University)
- 120 **Facial Expression Recognition with Deep two-view Support Vector Machine**
Chongliang Wu (University of Science and Technology of China), Shangfei Wang (University of Science and Technology of China), Bowen Pan (University of Science and Technology of China), Huaping Chen (University of Science and Technology of China)
- 121 **Mental Visual Indexing: Towards Fast Video Browsing**
Richang Hong (School of Computer and Information, Hefei University of Technology), Jun He (School of Computer and Information, Hefei University of Technology), Hanwang Zhang (School of Computing, National University of Singapore), Tat-Seng Chua (School of Computing, National University of Singapore)
- 122 **One Sensor is not Enough: Adapting and Fusing Sensors for the Quality Assessment of User Generated Video**
Stefan Wilk (TU Darmstadt), Manisha Luthra (TU Darmstadt), Wolfgang Effelsberg (TU Darmstadt)
- 123 **Boosting Video Description Generation by Explicitly Translating from Frame-Level Captions**
Yuan Liu (Ricoh Software Research Center (Beijing) Co., Ltd.), Zhongchao Shi (Ricoh Software Research Center (Beijing) Co., Ltd.)
- 124 **Artist-based Classification via Deep Learning with Multi-scale Weighted Pooling**
Kevin Alfianto Jangtjik (National Taiwan University of Science and Technology), Mei-Chen Yeh (National Taiwan Normal University), Kai-Lung Hua (National Taiwan University of Science and Technology)
- 125 **CrowdNet: A Deep Convolutional Network for Dense Crowd Counting**
Lokesh Boominathan (Indian Institute of Science), Srinivas S S Kruthiventi (Indian Institute of Science), R. Venkatesh Babu (Indian Institute of Science)
- 126 **Do Textual Descriptions Help Action Recognition?**
Matteo Bruni (University of Florence – Media Integration and Communication Center), Tiberio Uricchio (University of Florence – Media Integration and Communication Center), Lorenzo Seidenari (University of Florence – Media Integration and Communication Center), Alberto Del Bimbo (University of Florence – Media Integration and Communication Center)
- 127 **Frame Untangling for Unobtrusive Display-Camera Visible Light Communication**
Xiao Shu (McMaster University), Xiaolin Wu (Shanghai Jiao Tong University)
- 128 **Performance Measurements of Virtual Reality Systems: Quantifying the Timing and Positioning Accuracy**
Chun-Ming Chang (Academia Sinica), Cheng-Hsin Hsu (National Tsing Hua University), Chih-Fan Hsu (Academia Sinica), Kuan-Ta Chen (Academia Sinica)

- 129 **Synthesizing Emerging Images from Photographs**
Cheng-Han Yang (National Tsing Hua University), Ying-Miao Kuo (National Tsing Hua University), Hung-Kuo Chu (National Tsing Hua University)
- 130 **Predicting and Optimizing Image Compression**
Oleksandr Murashko (University of St Andrews), John Thomson (University of St Andrews), Hugh Leather (University of Edinburgh)
- 131 **Spectral and Cepstral Audio Noise Reduction Techniques in Speech Emotion Recognition**
Jouni Pohjalainen (University of Passau), Fabien Fabien Ringeval (Universite Grenoble Alpes), Zixing Zhang (University of Passau), Björn Schuller (University of Passau)

14:00 – 17:00 Video program (Lobby)

- Session chair:** Shin'ichi Satoh (National Institute of Informatics)
- 132 **AntiLoiter: A Loitering Discovery System for Longtime Videos across Multiple Surveillance Cameras**
Jianquan Liu (NEC Corporation), Shoji Nishimura (NEC Corporation), Takuya Araki (NEC Corporation)
- 133 **Magic Mirror: A Virtual Fashion Consultant**
Yejun Liu (Tsinghua University), Jia Jia (Tsinghua University), Jingtian Fu (Tsinghua University), Yihui Ma (Tsinghua University), Jie Huang (Tsinghua University), Zijian Tong (Sogou Corporation)
- 134 **Placing Broadcast News Videos in their Social Media Context Using Hashtags**
Joseph G. Ellis (Columbia University), Svebor Karaman (Columbia University), Hongzhi Li (Columbia University), Hong Bin Shim (Columbia University), Shih-Fu Chang (Columbia University)

14:00 – 17:00 Demo session (Demo floor)

- Session chair:** Pablo Cesar (CWI), Max Mühlhäuser (Technische Universität Darmstadt)
- 135 **MARIM: Mobile Augmented Reality for Interactive Manuals**
Tam V Nguyen (University of Dayton), Dorothy Tan (Singapore Polytechnic), Bilal Mirza (Singapore Polytechnic), Jose Sepulveda (Singapore Polytechnic)
- 136 **A Live Face Swapper**
Shengtao Xiao (National University of Singapore), Luoqi Liu (360), Xuecheng Nie (National University of Singapore), Jiashi Feng (National University of Singapore), Ashraf A Kassim (National University of Singapore), Shuicheng Yan (360)
- 137 **WorkCache: Salvaging siloed knowledge**
Scott Carter (FX Palo Alto Laboratory, Inc), Laurent Denoue (FX Palo Alto Laboratory, Inc), Matthew Cooper (FX Palo Alto Laboratory, Inc)
- 138 **Hypervideo Production Using Crowdsourced Youtube Videos**
Stefan John (Philipps-Universität Marburg), Christian Handschigl (University of Passau), Britta Meixner (FX Palo Alto Laboratory, Inc.), Michael Granitzer (University of Passau)
- 139 **SceneTextReg: A Real-Time Video OCR System**
Haojin Yang (Hasso Plattner Institute for Software Systems Engineering GmbH), Cheng Wang (Hasso Plattner Institute for Software Systems Engineering GmbH), Christian Bartz (Hasso Plattner Institute for Software Systems Engineering GmbH), Christoph Meinel (Hasso Plattner Institute for Software Systems Engineering GmbH)
- 140 **Beauty eMakeup: A Deep Makeup Transfer System**
Xinyu Ou (Chinese Academy of Sciences, Huazhong University of Science and Technology, Yunnan Open University), Si Liu (Chinese Academy of Sciences), Xiaochun Cao (Chinese Academy of Sciences), Hefei Ling (Huazhong University of Science and Technology)

- 141 **Real-time Wearable Computer Vision System for Improved Museum Experience**
Giovanni Taverriti (Università di Firenze), Stefano Lombini (Università di Firenze), Lorenzo Seidenari (Università di Firenze), Marco Bertini (Università di Firenze), Alberto Del Bimbo (Università di Firenze)
- 142 **An Intention-Aware Interactive System for Mobile Video Browsing**
Jun He (School of Computer and Information, Hefei University of Technology), Hanwang Zhang (School of Computing, National University of Singapore), Ling Shen (School of Computer and Information, Hefei University of Technology), Richang Hong (School of Computer and Information, Hefei University of Technology), Tat-Seng Chua (School of Computing, National University of Singapore)
- 143 **A Multimodal Gamified Platform for Real-Time User Feedback in Sports Performance**
David Monaghan (Dublin City University), Freddie Honohan (Dublin City University), Amin Ahmadi (Dublin City University), Troy McDaniel (Arizona State University), Ramin Tadayon (Arizona State University), Ajay Karpur (Arizona State University), kieran morran (Dublin City University), noel e o'connor (Dublin City University), Sethuraman Panchanathan (Arizona State University)
- 144 **PlaylistCreator: An Assisted Approach for Playlist Creation**
Ricardo Dias (INESC-ID, Instituto Superior Técnico, Universidade de Lisboa), Daniel Gonçalves (INESC-ID, Instituto Superior Técnico, Universidade de Lisboa), Manuel J. Fonseca (LaSIGE, Faculdade de Ciências, Universidade de Lisboa)
- 145 **WIMBY: What's in My Backyard?**
Michael Dorkhom (Queensland University of Technology), Alan Woodley (Queensland University of Technology), Shlomo Geva (Queensland University of Technology), Richi Nayak (Queensland University of Technology)
- 146 **SuperSelect – An Interactive Superpixel-Based Segmentation Method for Touch Displays**
Christoph Korinke (OFFIS), Tim Claudius Stratmann (University of Oldenburg), Tim Laue (University of Oldenburg), Susanne Boll (University of Oldenburg)
- 147 **ThePlantGame: Actively Training Human Annotators for Domain-specific Crowdsourcing**
Maximilien Servajean (INRIA/LIRMM), Alexis Joly (INRIA/LIRMM), Dennis Shasha (NYU), Julien Champ (INRIA/LIRMM), Esther Pacitti (INRIA/LIRMM)
- 148 **A Multi-Video Browser for Endoscopic Videos on Tablets**
Marco A. Hudelist (Klagenfurt University), Sabrina Kletz (Klagenfurt University), Klaus Schoeffmann (Klagenfurt University)
- 149 **A Tablet Annotation Tool for Endoscopic Videos**
Marco A. Hudelist (Klagenfurt University), Sabrina Kletz (Klagenfurt University), Klaus Schoeffmann (Klagenfurt University)
- 150 **News Archive Exploration Combining Face Detection and Tracking with Network Visual Analytics**
Benjamin Renoust (National Institute of Informatics & JFLI UMI 3527), Thanh Duc Ngo (University of Information Technology), Duy-Dinh Le (National Institute of Informatics), Shin'Ichi Satoh (National Institute of Informatics)
- 151 **A New Tool for Collaborative Video Search via Content-based Retrieval and Visual Inspection**
Wolfgang Hürst (Utrecht University), Algernon Ip Vai Ching (Utrecht University), Marco A. Hudelist (Klagenfurt University), Manfred J. Primus (Klagenfurt University), Klaus Schoeffmann (Klagenfurt University), Christian Beecks (RWTH Aachen University)
- 152 **A Browsing and Retrieval System for Broadcast Videos using Scene Detection and Automatic Annotation**
Lorenzo Baraldi (University of Modena and Reggio Emilia), Costantino Grana (University of Modena and Reggio Emilia), Alberto Messina (RAI – Radiotelevisione Italiana), Rita Cucchiara (University of Modena and Reggio Emilia)
- 153 **First-Person Shooter Game for Virtual Reality Headset with Advanced Multi-Agent Intelligent System**
Ilya Makarov (National Research University Higher School of Economics), Mikhail Tokmakov (National Research University Higher School of Economics), Pavel Polyakov (National Research University Higher School of Economics), Peter Zyuzin (National Research University Higher School of Economics), Maxim Martynov (National Research University Higher School of Economics), Oleg Konoplya (National Research University Higher School of Economics), George Kuznetsov (National Research University Higher School of Economics), Ivan Guschenko-Cheverda (National Research University Higher School of Economics), Maxim Uriev (National Research University Higher School of Economics), Ivan Mokeev (National Research University Higher School of Economics), Olga Gerasimova (National Research University Higher School of Economics), Lada Tokmakova (National Research University Higher School of Economics), Alexey Kosmachev (National Research University Higher School of Economics)
- 154 **SuperStreamer: Enabling Progressive Content Streaming in a Game Engi**
Yong Xue Eu (National University of Singapore), Jermyn Tanu (National University of Singapore), Justin Jieting Law (National University of Singapore), Muhammad Hanif B Ghazali (National University of Singapore), Shuan Siang Tay (National University of Singapore), Wei Tsang Ooi (National University of Singapore), Anand Bhojan (National University of Singapore)
- 155 **DeepSketch2Image: Deep Convolutional Neural Networks for Partial Sketch Recognition and Image Retrieval**
Omar Seddati (UMONS), Stéphane Dupont (UMONS), Said Mahmoudi (UMONS)
- 156 **A Fast Cattle Recognition System using Smart devices**
Santosh Kumar (Indian Institute of Technology (B.H.U), Varanasi-221005), Sanjay Kumar Singh (Indian Institute of Technology (B.H.U), Varanasi-221005), Tanima Datta (Indian Institute of Technology (B.H.U), Varanasi-221005), Hari Prabhat Gupta (Indian Institute of Technology (B.H.U), Varanasi-221005)
- 157 **Vibrotactile Experiences for Augmented Reality**
Wolfgang Hürst (Utrecht University), Nina Rosa (Utrecht University), Jean-Paul van Bommel (Utrecht University)
- 158 **Image2Text: A Multimodal Image Captioner**
Chang Liu (Tsinghua University), Changhu Wang (Multimedia Search and Mining/Microsoft Research Asia), Fuchun Sun (Tsinghua University), Yong Rui (Multimedia Search and Mining/Microsoft Research Asia)
- 159 **History Rhyme: Searching Historic Events by Multimedia Knowledge**
Yifan Xiong (Renmin University of China), Jia Chen (Carnegie Mellon University), Qin Jin (Renmin University of China), Chao Zhang (Renmin University of China)
- 160 **Intelli-Wrench: Smart Navigation Tool for Mechanical Assembly and Maintenance**
Toru Takahashi (NEC Corporation), Yuta Kudo (NEC Corporation), Rui Ishiyama (NEC Corporation)
- 161 **Interactive Image Search for Clothing Recommendation**
Zhengzhong Zhou (MOE-Microsoft Key Laboratory for Intelligent Computing and Intelligent Systems), Yifei Xu (MOE-Microsoft Key Laboratory for Intelligent Computing and Intelligent Systems), Jingjin Zhou (MOE-Microsoft Key Laboratory for Intelligent Computing and Intelligent Systems), Liqing Zhang (MOE-Microsoft Key Laboratory for Intelligent Computing and Intelligent Systems)

- 162 **Video ChatBot: Triggering Live Social Interactions by Automatic Video Commenting**
Yehao Li (Sun Yat-Sen University), Ting Yao (Microsoft Research Asia), Rui Hu (Microsoft Research Asia), Tao Mei (Microsoft Research Asia), Yong Rui (Microsoft Research Asia)
- 163 **bBridge: A Big Data Platform for Social Multimedia Analytics**
Aleksandr Farseev (National University of Singapore), Ivan Samborskii (ITMO University), Tat-Seng Chua (National University of Singapore)
- 164 **Scalable Multimedia Streaming in Wireless Networks with Device-to-Device Cooperation**
Karim Jahed (Lebanese American University (LAU)), Sanaa Sharafeddine (Lebanese American University (LAU)), Abdallah Moussawi (American University of Beirut (AUB)), Abbas Abou Daya (American University of Beirut (AUB)), Hassan Dbouk (American University of Beirut (AUB)), Saadallah Kassir (American University of Beirut (AUB)), Zaher Dawy (American University of Beirut), Preethi Valsalan (Qatar Mobility Innovation Center (QMIC)), Wael Cherif (Qatar Mobility Innovation Center (QMIC)), Fethi Filali (Qatar Mobility Innovation Center (QMIC))
- 165 **Leveraging ICN for Secure Content Distribution in IP Networks**
Syed Obaid Amin (Huawei Research Center), Qingji Zheng (Huawei Research Center), Ravishankar Ravindran (Huawei Research Center), GQ Wang (Huawei Research Center)

17:00 – 19:00 Reception

Topics and Meetings

Tuesday, October 18

Location: Tuschinski

09:00 – 10:30 SIGMM Award presentations / talks (Zaal 1)

Session chair: Rainer Lienhart (University of Augsburg)

10:30 – 11:00 Break

11:00 – 12:30 Panel: Deep Learning: Finding Generalization and Applications (Zaal 1)

Session chair: Alberto Del Bimbo (University of Florence), David Shamma (Yahoo Labs)

11:00 – 12:30 Analysis & Search (Zaal 2)

Session chair: Michele Merler (IBM TJ Watson Research Center)

166 **Event Specific Multimodal Pattern Mining for Knowledge Base Construction**
Hongzhi Li (Columbia University), Joseph G. Ellis (Columbia University), Heng Ji (Rensselaer Polytechnic Institute), Shih-Fu Chang (Columbia University)

167 **Joint Graph Learning and Video Segmentation Via Multiple Cues and Topology Calibration**

Jingkuan Song (University of Trento), Lianli Gao (University of Electronic Science and Technology of China), Mihai Marian Puscas (University of Trento), Feiping Nie (Northwestern Polytechnical University), Fumin Shen (University of Electronic Science and Technology of China), Nicu Sebe (University of Trento)

168 **Parsimonious Mixed-Effects HodgeRank for Crowdsourced Preference Aggregation**
Qianqian Xu (State Key Laboratory of Information Security, Institute of Information Engineering, CAS & BICMR, Peking University), Jiechao Xiong (School of Mathematical Sciences, BICMR-LMAM-LMEQF-LMP, Peking University), Xiaochun Cao (State Key Laboratory of Information Security, Institute of Information Engineering, CAS), Yuan Yao (School of Mathematical Sciences, BICMR-LMAM-LMEQF-LMP, Peking University)

169 **Weighted Linear Fusion of Multimodal Data – A Reasonable Baseline?**
Ognjen Arandjelovic (University of St Andrews)

11:00 – 12:30 Topics in Multimedia I (Zaal 3)

Session chair: Noel O'Connor (Dublin City University)

170 **Play and Rewind: Optimizing Video Binary Representations by Self-Supervised Temporal Hashing**

Hanwang Zhang (NUS), Meng Wang (Hefei University of Technology), Richang Hong (Hefei University of Technology), Tat-Seng Chua (NUS)

171 **Multi-Stream Multi-Class Fusion of Deep Networks for Video Classification**
Zuxuan Wu (Fudan University), Yu-Gang Jiang (Fudan University), Xi Wang (Fudan University), Hao Ye (Fudan University), Xiangyang Xue (Fudan University)

172 **QoE Prediction for Enriched Assessment of Individual Video Viewing Experience**
Yi Zhu (Delft University of Technology), Alan Hanjalic (Delft University of Technology), Judith A. Redi (Delft University of Technology)

- 173 **Deep CTR Prediction in Display Advertising**
Junxuan Chen (Shanghai Jiao Tong University), Baigui Sun (Alibaba Group), Hao Li (Alibaba Group), Hongtao Lu (Shanghai Jiao Tong University), Xian-Sheng Hua (Alibaba Group)

11:00 – 12:30 Video Analysis & Streaming (Zaal 4)

Session chair: Yu-Gang Jiang (Fudan University)

- 174 **DRIVING: Distributed Scheduling for Video Streaming in Vehicular Wi-Fi Systems**
Xi Chen (McGill University), Lei Rao (General Motors Company), Qiao Xiang (Yale University), Xue Liu (McGill University), Fan Bai (General Motors Company)
- 175 **Dynamic Resource Provisioning with QoS Guarantee for Video Transcoding in Online Video Sharing Service**
Guanyu Gao (Nanyang Technological University), Yonggang Wen (Nanyang Technological University), Cedric Westphal (Huawei Innovation Center & University of California, Santa Cruz)
- 176 **High-speed Depth Stream Generation from A Hybrid Camera**
Xinxin Zuo (Northwestern Polytechnical University & University of Kentucky), Sen Wang (Northwestern Polytechnical University & University of Kentucky), Jiangbin Zheng (Northwestern Polytechnical University), ruigang Yang (University of Kentucky)
- 177 **Spatio-Temporal Analysis of Bandwidth Maps for Geo-Predictive Video Streaming in Mobile Environments**
Bayan Taani (National University of Singapore), Roger Zimmermann (National University of Singapore)

12:30 – 14:00 Lunch break

12:30 – 14:00 MMSJ meeting (Zaal 3)

On invitation only

12:30 – 14:00 MM-Mag Meeting (Zaal 4)

On invitation only

12:30 – 14:00 Doctoral Symposium (VIP)

On invitation only

Session chair: Winston Hsu (National Taiwan University), Eckehard Steinbach (Technische Universität München)

- 178 **Multimodal-based Multimedia Analysis, Retrieval, and Services in Support of Social Media Applications**
Rajiv Ratn Shah (National University of Singapore)
- 179 **Geospatial Multimedia Data for Situation Recognition**
Mengfan Tang (University of California, Irvine)
- 180 **Image Emotion Computing**
Sicheng Zhao (Harbin Institute of Technology)
- 181 **First Person View Video Summarization Subject to the User Needs**
Ana Garcia del Molino (Institute for Infocomm Research, A*STAR & School of Computer Science and Engineering, NTU)
- 182 **Sentiment and Emotion Analysis for Social Multimedia: Methodologies and Applications**
Quanzeng You (University of Rochester)

183 **n-Dimensional Display Interface**

Charles D Estes (University of North Carolina)

184 **Multi-Modal Learning: Study on A Large-Scale Micro-Video Data Collection**

Jingyuan Chen (National University of Singapore)

185 **Weakly-Supervised Recognition, Localization, and Explanation of Visual Entities**

Pascal Mettes (University of Amsterdam)

186 **Zero-Example Multimedia Event Detection and Recounting with Unsupervised Evidence Localization**

Yi-Jie Lu (City University of Hong Kong)

14:00-15:30 Deep Learning (Zaal 1)

Session chair: Elisa Ricci (FBK Trento)

- 187 **Multilayer and Multimodal Fusion of Deep Neural Networks for Video Classification**
Xiaodong Yang (NVIDIA), Pavlo Molchanov (NVIDIA), Jan Kautz (NVIDIA)
- 188 **Image Captioning with Deep Bidirectional LSTMs**
Cheng Wang (Hasso Plattner Institute, University of Potsdam), Haojin Yang (Hasso Plattner Institute, University of Potsdam), Christian Bartz (Hasso Plattner Institute, University of Potsdam), Christoph Meinel (Hasso Plattner Institute, University of Potsdam)
- 189 **Deep Cross Residual Learning for Multitask Visual Recognition**
Brendan Jou (Columbia University), Shih-Fu Chang (Columbia University)
- 190 **Robust Visual-Textual Sentiment Analysis: When Attention meets Tree-structured Recursive Neural Networks**
Quanzeng You (University of Rochester), Liangliang Cao (Yahoo Labs), Hailin Jin (Adobe), Jiebo Luo (University of Rochester)

14:00-15:30 Brave new topics (Zaal 2)

Session chair: Martha Larson (Radboud University)

- 191 **Research Challenges in Developing Multimedia Systems for Managing Emergency Situations**
Mengfan Tang (University of California, Irvine), Siripen Pongpaichet (University of California, Irvine), Ramesh Jain (University of California, Irvine)
- 192 **Multimedia on the Mountaintop: Using Public Snow Images to Improve Water Systems Operation**
Andrea Castelletti (Politecnico di Milano), Roman Fedorov (Politecnico di Milano), Piero Fraternali (Politecnico di Milano), Matteo Giuliani (Politecnico di Milano)
- 193 **Crowdsourcing Biodiversity Monitoring: How Sharing your Photo Stream can Sustain our Planet**
Alexis Joly (INRIA), Hervé Goëau (INRIA), Julien Champ (INRIA), Samuel Dufour-Kowalski (INRIA), Henning Müller (HS-SO), Pierre Bonnet (CIRAD)
- 194 **Multimedia and Medicine: Teammates for Better Disease Detection and Survival**
Michael Riegler (Simula Research Laboratory and University of Oslo), Mathias Lux (Klagenfurt University), Carsten Gridwodz (Simula Research Laboratory and University of Oslo), Concetto Spampinato (University of Catania), Thomas de Lange (Cancer Registry of Norway and Vestre Viken Hospital Trust), Sigrun L. Eskeland (Vestre Viken Hospital Trust), Konstantin Pogorelov (Simula Research Laboratory and University of Oslo), Wallapak Tavanapong (Iowa State University), Peter T. Schmidt (Karolinska Institutet, Sweden and Center for Digestive Diseases, Solna and Karolinska University Hospital), Cathal Gurrin (Dublin City University), Dag Johansen (UiT – The Arctic University of Norway), Håvard Johansen (UiT – The Arctic University of Norway), Pål Halvorsen (Simula Research Laboratory and University of Oslo)

14:00-15:30 Topics in Multimedia II (Zaal 3)

Session chair: Matt Cooper (FXPAL)

- 195 **Micro Tells Macro: Predicting the Popularity of Micro-Videos via a Transductive Model**
Jingyuan Chen (National University of Singapore), Xuemeng Song (National University of Singapore), Liqiang Nie (National University of Singapore), Xiang Wang (National University of Singapore), Hanwang Zhang (National University of Singapore), Tat-Seng Chua (National University of Singapore)
- 196 **Leveraging Contextual Cues for Generating Basketball Highlights**
Vinay Bettadapura (Google Inc.), Caroline Pantofaru (Google Inc.), Irfan Essa (Georgia Institute of Technology)
- 197 **Server Allocation for Multiplayer Cloud Gaming**
Yunhua Deng (Nanyang Technological University), Yusen Li (Nankai University), Xueyan Tang (Nanyang Technological University), Wentong Cai (Nanyang Technological University)
- 198 **Share-and-Chat: Achieving Human-Level Video Commenting by Search and Multi-View Embedding**
Yehao Li (Sun Yat-Sen University), Ting Yao (Microsoft Research Asia), Tao Mei (Microsoft Research Asia), Hongyang Chao (Sun Yat-Sen University), Yong Rui (Microsoft Research Asia)

14:00-15:30 Events and Context (Zaal 4)

Session chair: Guillaume Gravier (IRISA)

- 199 **Context-aware Image Tweets Modelling and Recommendation**
Tao Chen (National University of Singapore), Xiangnan He (National University of Singapore), Min-Yen Kan (National University of Singapore)
- 200 **Semantic Image Profiling for Historic Events: Linking Images to Phrases**
Jia Chen (Carnegie Mellon University), Qin Jin (Renmin University of China), Yifan Xiong (Renmin University of China)
- 201 **Audio Event Detection using Weakly Labeled Data**
Anurag Kumar (Carnegie Mellon University), Bhiksha Raj (Carnegie Mellon University)
- 202 **Event Localization in Music Auto-tagging**
Jen-Yu Liu (National Taiwan University), Yi-Hsuan Yang (Academia Sinica)

15:30 – 16:00 Break

16:00 – 17:30 Multimedia Grand Challenge (Zaal 1)

Session chair: Xavier Anguera (ELSA)

- 203 **Face Recognition via Active Annotation and Learning**
Hao Ye (Shanghai Advanced Research Institute, Chinese Academy of Sciences), Weiyuan Shao (Shanghai Advanced Research Institute, Chinese Academy of Sciences), Hong Wang (Shanghai Advanced Research Institute, Chinese Academy of Sciences), Jianqi Ma (Fudan University), Li Wang (Fudan University), Yingbin Zheng (Shanghai Advanced Research Institute, Chinese Academy of Sciences), Xiangyang Xue (Fudan University)
- 204 **Deep Convolutional Neural Network with Independent Softmax for Large Scale Face Recognition**
Yue Wu (Northeastern University), Jun Li (Northeastern University), Yu Kong (Northeastern University), Yun Fu (Northeastern University)
- 205 **Robust Face Recognition with Deep Multi-View Representation Learning**
Jianshu Li (National University of Singapore), Jian Zhao (National University of Singapore), Fang Zhao (National University of Singapore), Hao Liu (Hefei University of Technology), Jing Li (National University of Singapore), Shengmei Shen (Panasonic R&D Center Singapore), Jiashi Feng (National University of Singapore), Terence Sim (National University of Singapore)

206 Frame- and Segment-Level Features and Candidate Pool Evaluation for Video Caption Generation

Rakshith Shetty (Aalto University), Jorma Laaksonen (Aalto University)

207 Contextual Enrichment of Remote-Sensed Events with Social Media Streams

Benjamin Bischke (German Research Center for Artificial Intelligence), Damian Borth (German Research Center for Artificial Intelligence), Christian Schulze (German Research Center for Artificial Intelligence), Andreas Dengel (German Research Center for Artificial Intelligence)

208 Early Embedding and Late Reranking for Video Captioning

Jianfeng Dong (Zhejiang University), Xirong Li (Renmin University of China), Weiyu Lan (Renmin University of China), Yujia Huo (Renmin University of China), Cees G. M. Snoek (University of Amsterdam)

209 Describing Videos using Multi-modal Fusion

Qin Jin (Renmin University of China), Jia Chen (Carnegie Mellon University), Shizhe Chen (Renmin University of China), Yifan Xiong (Renmin University of China), Alexander Hauptmann (Carnegie Mellon University)

210 Multimodal Video Description

Vasili Ramanishka (University of Massachusetts), Abir Das (University of Massachusetts), Dong Huk Park (University of California), Subhashini Venugopalan (University of Texas), Lisa Anne Hendricks (University of California), Marcus Rohrbach (University of California), Kate Saenko (University of Massachusetts)

211 Tracking Natural Events through Social Media and Computer Vision

Jingya Wang (Indiana University), Mohammed Korayem (Indiana University), Saul Blanco (Indiana University), David Crandall (Indiana University)

212 ConTagNet: Exploiting User Context for Image Tag Recommendation

Yogesh Singh Rawat (National University of Singapore), Mohan S Kankanhalli (National University of Singapore)

213 Image Captioning with both Object and Scene Information

Xiangyang Li (Institute of Computing Technology, Chinese Academy of Sciences), Xinhang Song (Institute of Computing Technology, Chinese Academy of Sciences), Luis Herranz (Institute of Computing Technology, Chinese Academy of Sciences), Yaohui Zhu (Institute of Computing Technology, Chinese Academy of Sciences), Jiang Shuqiang (Institute of Computing Technology, Chinese Academy of Sciences)

214 Generating Affective Captions using Concept And Syntax Transition Networks

Tushar Karayil (University of Kaiserslautern), Philipp Blandfort (University of Kaiserslautern), Damian Borth (German Research Center for Artificial Intelligence (DFKI)), Andreas Dengel (German Research Center for Artificial Intelligence (DFKI)), University of Kaiserslautern

17:30 – 19:30 Social event, departure from Tuschinski (Canal Cruise)

19:30 – 22:00 Conference dinner at I-Dock

Topics and Meetings

Wednesday, October 19

Location: Tuschinski

09:00 – 10:30 **Keynote II: Jack van Wijk (Zaal 1)**

Session chair: Marcel Worring (University of Amsterdam)

Visual Analytics for Multimedia: Challenges and Opportunities

Jack van Wijk (Eindhoven University of Technology)

10:30 – 11:00 **Break**

11:00 – 12:30 **Open Source Software Competition (Zaal 1)**

Session chair: Tao Mei (Microsoft Research Asia)

215 **LightNet: A Versatile, Standalone Matlab-based Environment for Deep Learning**

Chengxi Ye (Univ of Maryland), Chen Zhao (Univ of Maryland), Yezhou Yang (Univ of Maryland), Cornelia Fermüller (Univ of Maryland), Yiannis Aloimonos (Univ of Maryland)

216 **Morph: A Fast and Scalable Cloud Transcoding System**

Guanyu Gao (Nanyang Technological University), Yonggang Wen (Nanyang Technological University)

217 **Smart Beholder: An Extensible Smart Lens Platform**

Chun-Ying Huang (National Chiao Tung University), Ching-Ling Fan (National Tsing Hua University), Chih-Fan Hsu (Academia Sinica), Hsin-Yu Chang (National Tsing Hua University), Tsung-Han Tsai (Academia Sinica), Kuan-Ta Chen (Academia Sinica), Cheng-Hsin Hsu (National Tsing Hua University)

218 **A Platform for Building New Human-Computer Interface Systems that Support Online Automatic Recognition of Audio-Gestural Commands**

Nikolaos Kardaris (National Technical University of Athens), Isidoros Rodomagoulakis (National Technical University of Athens), Vassilis Pitsikalis (National Technical University of Athens), Antonis Arvanitakis (National Technical University of Athens), Petros Maragos (National Technical University of Athens)

219 **madmom: a new Python Audio and Music Signal Processing Library**

Sebastian Böck (Johannes Kepler University), Filip Korzeniowski (Johannes Kepler University), Jan Schlüter (Austrian Research Institute for Artificial Intelligence), Florian Krebs (Johannes Kepler University), Gerhard Widmer (Johannes Kepler University)

220 **Kvazaar: Open-Source HEVC/H.265 Encoder**

Marko Viitanen (Tampere University of Technology), Ari Koivula (Tampere University of Technology), Ari Lemmetti (Tampere University of Technology), Arttu Ylä-Outinen (Tampere University of Technology), Jarno Vanne (Tampere University of Technology), Timo D. Hämmäläinen (Tampere University of Technology)

221 **vitivr – A Flexible Retrieval Stack Supporting Multiple Query Modes for Searching in Multimedia Collections**

Luca Rossetto (University of Basel), Ivan Giangreco (University of Basel), Claudiu Tanase (University of Basel), Heiko Schuldt (University of Basel)

222 **Kurento: the WebRTC Modular Media Server**

Luis López (URJC), Miguel París (URJC), Santiago Carot (ETSIT), Boni García (URJC), Micael Gallego (URJC), Francisco Gortázar (URJC), Raul Benítez (URJC), Jose A Santos (Naevatec),

David Fernández (Naevatec), Radu T Vlad (Naevatec), Iván Gracia (Naevatec), Francisco Javier López (Naevatec)

223 **Modular Parallelization Framework for Multi-Stream Video Processing**

Tim Lenertz (Université Libre de Bruxelles), Gauthier Lafruit (Université Libre de Bruxelles)

224 **OpenVQ**

Kristian Skarseth (LABO Mixed Realities), Henrik Bjørlo (Sopra Steria), Pål Halvorsen (Simula Research Laboratory & University of Oslo, Norway), Michael Riegler (Simula Research Laboratory & University of Oslo, Norway), Carsten Griwodz (Simula Research Laboratory & University of Oslo, Norway)

225 **CNNdroid: GPU-Accelerated Execution of Trained Deep Convolutional Neural Networks on Android**

Seyyed Salar Latifi Oskoui (Sharif University of Technology), Hossein Golestani (Sharif University of Technology), Matin Hashemi (Sharif University of Technology), Soheil Ghiasi (University of California, Davis)

226 **Tamp: A Library for Compact Deep Neural Networks with Structured Matrices**

Bingchen Gong (Zhejiang University), Brendan Jou (Columbia University), Felix Yu (Columbia University), Shih-Fu Chang (Columbia University)

227 **Barrista — Caffè Well-Served**

Christoph Lassner (Bernstein Center for Computational Neuroscience & MPI for Intelligent Systems), Daniel Kappler (Max-Planck Institute for Intelligent Systems), Martin Kiefel (Bernstein Center for Computational Neuroscience & MPI for Intelligent Systems), Peter Gehler (Bernstein Center for Computational Neuroscience & MPI for Intelligent Systems)

228 **Pyo, the Python DSP toolbox**

Olivier Belanger (University of Montreal)

229 **SenseCap: Synchronized Data Collection with Microsoft Kinect2 and LeapMotion**

Julian F P Kooij (Delft University of Technology & Leiden University Medical Center)

230 **MP3DG-PCC, Open Source Software Framework for Implementation and Evaluation of Point Cloud Compression**

Rufael Mekuria (CWI, Unified Streaming), Pablo Cesar (CWI)

11:00 – 12:30 **Topics in Multimedia III (Zaal 2)**

Session chair: Symeon Papadopoulos (CERTH-ITI)

231 **V3I-STAL: Visual Vehicle-to-Vehicle Interaction via Simultaneous Tracking and Localization**

Xiaobai Liu (San Diego State University)

232 **Are Safer Looking Neighborhoods More Lively? A Multimodal Investigation into Urban Life**

Marco De Nadai (FBK, University of Trento), Radu Laurentiu Vieriu (University of Trento), Gloria Zen (University of Trento), Stefan Dragicevic (TIM and University of Trento), Nikhil Naik (MIT Media Lab), Michele Caraviello (TIM), Cesar Augusto Hidalgo (MIT Media Lab), Nicu Sebe (University of Trento), Bruno Lepri (FBK)

233 **Detecting Sarcasm in Multimodal Social Platforms**

Rossano Schifanella (University of Turin), Paloma de Juan (Yahoo), Joel Tetreault (Yahoo), LiangLiang Cao (Yahoo)

234 **User Redirection and Direct Haptics in Virtual Environments**

Cristiano Carvalheiro (FEUP), Rui Nóbrega (FEUP/INESC TEC), Hugo da Silva (FEUP), Rui Rodrigues (FEUP/INESC TEC)

11:00 – 12:30 Learning & Hashing (Zaal 3)

Session chair: Xavier Giró-i-Nieto (Universitat Politècnica de Catalunya)

- 235 **Human Pose Estimation from Still Depth Image via Inference Embedded Multi-task Learning**
Keze Wang (Sun Yat-sen University), Shengfu Zhai (Sun Yat-sen University), Hui Cheng (Sun Yat-sen University), Xiaodan Liang (Sun Yat-sen University), Liang Lin (Sun Yat-sen University)
- 236 **Cross-batch Reference Learning for Deep Classification and Retrieval**
Huei-Fang Yang (Academia Sinica), Kevin Lin (Academia Sinica), Chu-Song Chen (Academia Sinica)
- 237 **Binary Optimized Hashing**
Qi Dai (Fudan University), Jianguo Li (Intel Labs China), Jingdong Wang (Microsoft Research Asia), Yu-Gang Jiang (Fudan University)
- 238 **Linear Distance Preserving Pseudo-Supervised and Unsupervised Hashing**
Min Wang (University of Science and Technology of China), Wengang Zhou (University of Science and Technology of China), Qi Tian (University of Texas at San Antonio), Zhengjun Zha (University of Science and Technology of China), Houqiang Li (University of Science and Technology of China)

11:00 – 12:30 Transport & Experience (Zaal 4)

Session chair: Wenwu Zhu (Tsinghua University)

- 239 **A Pragmatically Designed Adaptive and Web-compliant Object-based Video Streaming Methodology – Implementation and Subjective Evaluation**
Maarten Wijnants (Hasselt University – tUL – iMinds), Gustavo Rovelo (Hasselt University – tUL – iMinds), Peter Quax (Hasselt University – tUL – iMinds), Wim Lamotte (Hasselt University – tUL – iMinds)
- 240 **A Perceptual Quality Metric for Videos Distorted by Spatially Correlated Noise**
Chao Chen (Google Inc.), Mohammad Izadi (Google Inc.), Anil Kokaram (Google Inc.)
- 241 **Zero-Shot Hashing via Transferring Supervised Knowledge**
Yang Yang (University of Electronic Science and Technology of China), Yadan Luo (University of Electronic Science and Technology of China), Weilun Chen (University of Electronic Science and Technology of China), Fumin Shen (University of Electronic Science and Technology of China), Jie Shao (University of Electronic Science and Technology of China), Heng Tao Shen (The University of Queensland)
- 242 **SDNDASH: A Novel Architecture For Improving QoE of Dynamic Adaptive video Streaming over HTTP using Software Defined Networking**
Abdelhak Bentaleb (National University of Singapore), Ali C. Begen (Ozyegin University), Roger Zimmermann (National University of Singapore)

12:30 – 14:15 SIGMM Business Meeting & ACMM M16 Awards (Zaal 1)

Session chair: Shih-Fu Chang (Columbia University)

14:15 – 15:45 Topics in Multimedia IV (Zaal 1)

Session chair: Qi Tian (University of Texas at San Antonio)

- 243 **Query Adaptive Instance Search using Object Sketches**
Sreyasee Das Bhattacherjee (Nanyang Technological University), junsong Yuan (Nanyang technological University), Weixiang Hong (Nanyang technological University), Xiang Ruan (Tiwaki Co. Ltd)
- 244 **Key Color Generation for Affective Multimedia Production : An Initial Method and Its Application**
EunJin Kim (KAIST), Hyeon-Jeong Suk (KAIST)

- 245 **Academic Coupled Dictionary Learning for Sketch-based Image Retrieval**
Dan Xu (DISI, University of Trento), Xavier Alameda-Pineda (DISI, University of Trento), Jingkuan Song (DISI, University of Trento), Elisa Ricci (Fondazione Bruno Kessler (FBK) & University of Perugia), Nicu Sebe (DISI, University of Trento)
- 246 **Time Matters: Multi-scale Temporalization of Social Media Popularity**
Bo Wu (Institute of Computing Technology, Chinese Academy of Sciences), Wen-Huang Cheng (Research Center for Information Technology Innovation, Academia Sinica), Yongdong Zhang (Institute of Computing Technology, Chinese Academy of Sciences), Tao Mei (Microsoft Research)

14:15 – 15:45 Analysis & Middleware (Zaal 3)

Session chair: Lei Zhang (Microsoft Research)

- 247 **Transform Invariant Convolutional Neural Networks For Image Classification And Search**
Xu Shen (University of Science and Technology of China), Xinmei Tian (University of Science and Technology of China), Anfeng He (University of Science and Technology of China), Shaoyan Sun (University of Science and Technology of China), Dacheng Tao (University of Technology, Sydney)
- 248 **PL-ranking: A Novel Ranking Method for Cross-Modal Retrieval**
Liang Zhang (University of Chinese Academy of Sciences), Bingpeng Ma (University of Chinese Academy of Sciences), Guorong Li (University of Chinese Academy of Sciences), Qingming Huang (University of Chinese Academy of Sciences), Qi Tian (University of Texas at San Antonio)
- 249 **Video eCommerce: Towards Online Video Advertising**
Zhi-Qi Cheng (Southwest Jiaotong University), Yang Liu (Alibaba Group), Xiao Wu (Southwest Jiaotong University), Xian-Sheng Hua (Alibaba Group)
- 250 **Affective Contextual Mobile Recommender System**
Chao Wu (Tsinghua University), Jia Jia (Tsinghua University), Wenwu Zhu (Tsinghua University), Xu Chen (University of Goettingen), Bowen Yang (Tsinghua University), Yaoxue Zhang (Central South University)

14:15 – 15:45 Emotions, People and Faces (Zaal 1)

Session chair: Lei Zhang (Microsoft Research)

- 251 **Predicting Personalized Emotion Perceptions of Social Images**
Sicheng Zhao (Harbin Institute of Technology), Hongxun Yao (Harbin Institute of Technology), Yue Gao (Tsinghua University), Rongrong Ji (Xiamen University), Wenlong Xie (Harbin Institute of Technology), Xiaolei Jiang (Harbin Institute of Technology), Tat-Seng Chua (National University of Singapore)
- 252 **StressClick: Sensing Stress from Gaze-Click Patterns**
Michael Xuelin Huang (Hong Kong Polytechnic University), Jiajia Li (Hong Kong Polytechnic University), Grace Ngai (Hong Kong Polytechnic University), Hong Va Leong (Hong Kong Polytechnic University)
- 253 **Ensemble of Sparse Cross-Modal Metrics for Heterogeneous Face Recognition**
Jing Huo (Nanjing University), Yang Gao (Nanjing University), Yinghuan Shi (Nanjing University), Wanqi Yang (Nanjing Normal University & Nanjing University), Hujun Yin (The University of Manchester)
- 254 **Shorter-is-Better: Venue Category Estimation from Micro-Video**
Jianglong Zhang (Communication University of China), Liqiang Nie (Shandong University), Xiang Wang (National University of Singapore), Xiangnan He (National University of Singapore), Xianglin Huang (Communication University of China), Tat Seng Chua (National University of Singapore)

14:15 – 15:45 Exchange meeting (VIP)
On invitation only

15:45 – 16:15 Break

16:15 – 17:45 SIGMM Rising Star Symposium (Zaal 1)

Session chair: Susanne Boll (University of Oldenburg), Shih-Fu Chang (Columbia University)

255 **Opportunities and Challenges of Social Media in Personal and Societal Well-Being**

Munmun De Choudhury (Georgia Institute of Technology)

Commentator: Klara Nahrstedt

256 **Being Moved by Motion: How Social Science Inspires Multimedia Analysis in the Wild**

Hayley Hung (Delft University of Technology)

Commentator: Lynn Wilcox

257 **Tag Embeddings for Multimedia Retrieval and Description**

Xirong Li (Renmin University of China)

Commentator: Mohan Kankanhalli

258 **Purpose and Repurpose: Lessons from the Concert Stage**

Cynthia Liem (Delft University of Technology)

Commentator: Gerald Friedland

259 **Towards "Wow!" Multimedia Quality of Experience: if you can measure it, you can achieve it**

Judith Redi (Delft University of Technology and CWI)

Commentator: Carsten Griwodz

260 **About Multimedia Presentation Generation and Multimedia Metadata: From Synthesis to Analysis, and Back?**

Ansgar Scherp (ZBW — Leibniz-Information Centre for Economics and Kiel University)

Commentator: Arnold Smeulders

Sponsors and supporters
ACM Multimedia 2016



