

Reports on CBMI 16 and ICME 16

CBMI 16

Bogdan Ionescu
*University Politehnica of
Bucharest*

Henning Müller
*University of Applied
Sciences Western
Switzerland*

Yiannis Kompatsiaris
*Centre for Research and
Technology-Hellas*

Guillaume Gravier
IRISA

ICME 16
Anthony Vetro
*Mitsubishi Electric
Research Labs*

This issue, we have not one but two conference reports. The first covers the 14th International Workshop on Content-Based Multimedia Indexing (CBMI 16), while the second (starting on page 91) covers the 2016 IEEE International Conference on Multimedia and Expo (ICME 16).

—Susanne Boll, Department Editor

The 14th International Workshop on Content-Based Multimedia Indexing

Bogdan Ionescu, Henning Müller, Yiannis Kompatsiaris, and Guillaume Gravier

The International Workshop on Content-Based Multimedia Indexing (CBMI) aims to annually bring together the various communities involved in all aspects of content-based multimedia indexing—from retrieval and browsing to visualization and analytics. The 14th edition of CBMI (<http://cbmi2016.upb.ro>) achieved its goal, maintaining the tradition of the previous successful events, the first of which was held at the Paul Sabatier University, Toulouse, back in 1999. This year's workshop was organized by the University Politehnica of Bucharest in cooperation with University of Applied Sciences Western Switzerland (HES-SO), in Bucharest, Romania from 15–17 June. Intel Software Romania and Technicolor France cosponsored the event, and the IEEE Signal Processing Society, ACM Special Interest Group on Multimedia, and European Association for Signal Processing were technical cosponsors. Although the workshop was primarily a European event, contributions came from authors in 27 different countries, including the US, Singapore, Japan, China, India, and Brazil (see Figure 1).

CBMI 2016 welcomed submissions as regular papers, short papers, special sessions, and demo contributions/presentations. The highlighted topics ranged from multimedia analytics, event-based media processing, and data representation and coding, to learning (including deep learning) and indexing (cross-modal and large-scale indexing), to multimedia presentation, visualization techniques, and bench-

marking activities. The special sessions covered the domain-specific applications of multimedia retrieval in medicine and e-learning.

Best Papers

Among the main-track submissions, two were given the Best Paper award. The first paper was “Experimenting with Musically Motivated Convolutional Neural Networks,” by Jordi Pons, Thomas Lidy, and Xavier Serra. The authors explored, in an innovative way, the various architectural choices of deep networks for music signal classification tasks. They analyzed different convolutional filters of varying shapes to determine to what degree they can fit the specificity of musical concepts. Then, they introduced several musically motivated architectures and successfully validated the architectures through experiments.

The second Best Paper award went to “A Scattering Transform Combination with Local Binary Patterns for Texture Classification,” by Vu-Lam Nguyen, Ngoc-Son Vu, and Philippe-Henri Goselin. The authors addressed the problem of texture classification, introducing a combined feature approach that exploits local structure information (via local binary patterns) and global information (via the scattering transform). Extensive experimental tests on many specific benchmarking datasets proved the effectiveness of this approach for achieving state-of-the-art results.

Special Sessions

Apart from the main track, the workshop hosted three special sessions. The first,



Figure 1. Participants at the 14th International Workshop on Content-Based Multimedia Indexing (CBMI 2016) at the University Politehnica of Bucharest Conference Center.

“Content-Based Image and Multimedia Analysis and Indexing for Healthcare,” focused on applications using multimedia data in health. Medical institutions are indeed large producers of multimedia data. It’s often difficult to share such datasets, because personally identifying information might be included. However, as illustrated during this special session, an increasing number of large and partly annotated datasets have become available.

The second special session, “Multimedia Indexing for eLearning,” was a novel initiative. With the arrival of Massive Open Online Courses—and other distance-learning activities—the management of multimedia data for learning is becoming an important domain. This trend was explored in detail during this session.

The last special session was “Deep Learning for Multimedia Indexing.” Deep learning—that is, the use of multilayer neural networks—captured the attention of the research community during the ImageNet benchmarks as well as other scientific challenges in which deep-learning frameworks achieved very good performance. In some cases, accuracy improved more than 10 percentage points over current state-of-the-art techniques. With tools such as Caffe or TensorFlow being openly accessible, it seems important to explore the use of deep-learning frameworks and provide experience reports on using them for content-based analysis and indexing.

Keynotes and Panel Session

Each of the three days of the workshop featured a position talk or panel that discussed some of the hot topics in the field. Alexander G. Hauptmann from Carnegie Mellon University focused on the problem of Web-scale video searches.

Indeed, high-quality systems are often slow in practical use, so for interactive systems, it’s important to weigh speed versus accuracy for all system components. Participants discussed experiments that had been performed on existing systems to evaluate the various aspects of video search performance.

Daniel Gatica-Perez from the Idiap Research Institute addressed civic multimedia, crowdsourcing, and the public good. He explored how young people can be engaged in tasks that include sharing multimedia content on civic tasks.

Finally, a panel discussion, called “Multimedia Analysis Out of the Box: New Applications and Domains,” was led by Martha Larson from the Delft University of Technology, who brought together experts such as Hauptmann, Guillaume Gravier (IRISA), and Bernard Merialdo (EURECOM). The panelists engaged the audience in an enlightening discussion about multimedia analysis applications and domains (see Figure 2a). The panel was structured into three rounds related to the past, present, and future of multimedia analysis research; participants discussed, respectively,

- past visions (from the last 20 years) for multimedia analysis applications that came to be,
- current (widely held) visions for multimedia analysis applications that are doomed, and
- future visions for which we should strive.

A detail overview of the gained insights is available at <http://ngrams.blogspot.ro/2016/06/multimedia-analysis-out-of-box-new.html> (see Figure 2).



(a)



(b)

Figure 2. Conference photos: (a) the “Multimedia Analysis Out of the Box: New Applications and Domains” panel session and (b) the social dinner at Crama Domnească restaurant.

CBMI 2016 had an extremely interesting program that stimulated further discussions in the community in a dynamic environment. In fact, lively conversations occurred at the social dinner on the second night of the workshop (see Figure 2b). We hope the research and discussions continue, and we look forward to the 2017 edition of CBMI.

Bogdan Ionescu is a professor at the University Politehnica of Bucharest, Romania. He was a general chair for CBMI 2016. Contact him at bionescu@alpha.imag.pub.ro.

Henning Müller is a professor at the University of Applied Sciences Western Switzerland (HES-SO). He was a general chair for CBMI 2016. Contact him at henning.mueller@hevs.ch.

Yiannis Kompatsiaris is a senior researcher at the Centre for Research and Technology-Hellas, Greece. He was a technical program chair for CBMI 2016. Contact him at ikom@iti.gr.

Guillaume Gravier is a senior researcher at IRISA, the Research Institute of Computer Science and Random Systems, in France. He was a technical program chair for CBMI 2016. Contact him at guig@irisa.fr.

Highlights from ICME 2016

Anthony Vetro

The 2016 IEEE International Conference on Multimedia and Expo (ICME 2016) was held in Seattle, Washington, from 11–15 July 2016. ICME is a unique event that brings together multimedia researchers from the four IEEE sponsoring societies: the IEEE Circuits and Systems Society (CAS), IEEE Communications Society (ComSoc), IEEE Computer Society (CS), and IEEE Signal Processing Society (SPS). This year, the conference received a total of 512 submissions, and 152 papers (30 percent) were accepted.

Key Themes

Spearheading the conference were keynotes given by leading authorities in three of the most interesting and timely areas of multimedia today. Mark Billinghurst (University South Australia) spoke about wearable computing and augmented and virtual reality (AR/VR), Fei-Fei Li (Stanford University) covered image understanding and deep learning (see Figure 3), and Dariu Gavrilă (Daimler R&D) discussed the future prospects of autonomous vehicles.

Many elements of the technical program revolved around these themes. For instance, in the area of augmented and virtual reality, a panel discussion on “Making the Virtual Real: The Future of Augmented and Virtual Reality” was moderated by Billinghurst and featured key players in this area, including Steve Seitz (University of Washington and Google), Jeremy Selan (Valve), and Michael Gourlay (Microsoft). AR/VR was also addressed during the seventh edition of the Hot3D Workshop, which featured keynotes by Fernando Pereira (Instituto Superior Técnico) and Bernard Kress (Microsoft). Pereira discussed recent advances and challenges in plenoptic representation and coding, while Kress argued that AR/VR will be the next computing platform. Kress then spoke about the difficulties that must be overcome to realize such systems. ICME 16 also hosted a related grand challenge on light-field image compression. Participants presented novel solutions for efficiently compressing and manipulating images captured with an array of microlenses using the Lytro ILLUM camera.

In the area of media and scene understanding, the keynotes by Li and Gavrilă were com-

plemented by a special session on “Deep Learning for Multimedia Computing,” as well as regular sessions and workshops reporting on recent advances in content analysis, object detection/tracking, and media computing. In addition to a tutorial on situation recognition from multimodal data, the conference also featured an ongoing grand challenge on image recognition, sponsored by Microsoft Research, with this year’s challenge focusing on the recognition of celebrity faces in large-scale image databases.

Multimedia compression and delivery continues to be an important theme at ICME. For the first time, the Packet Video Workshop was hosted at ICME, featuring a keynote by Phil Chou (Microsoft) on coding for augmented and virtual reality, which was followed by a series of papers on immersive video and media streaming. The conference also included a tutorial on High Efficiency Video Coding (HEVC) and Graph Signal Processing for Compression, as well as a grand challenge on media streaming, organized by Bitmovin, which explored techniques for improving the quality of experience in end-to-end delivery systems based on the DASH (Dynamic Adaptive Streaming over HTTP) standard.

Industry Program

To better engage industry, ICME 16 introduced a new program called the “Hands-On Expo,” which paired companies that have ground-breaking new media-related products with world-class multimedia researchers and technologists. Participants were given the opportunity to get hands-on experience and technical information from the creators of breakthrough products, while companies could promote their products and develop strong ties with key technical influencers. The inaugural event featured expo sessions by Microsoft on cognitive services and by Intel on RealSense depth sensors. The sessions were very successful, with rooms at capacity. The sessions appealed to students, senior researchers, and professors alike (see Figure 4), and presented a unique opportunity for in-depth engagement between industry and academia.



Figure 3. Keynote by Fei-Fei Li (Stanford University) on “A Quest for Visual Intelligence in Computers.”

delivering HDR content to the home, including legacy issues; and Nandhakumar outlined the latest advances in HDR display technology. The forum was not only educational; it also provided a glimpse into high-quality home entertainment services on the horizon.

Student Program

There were two key aspects of the Student Program: a three-minute thesis competition, followed by a student-mentor luncheon. The program was funded by the National Science Foundation, and both the competition and luncheon had strong participation. Seventeen students succinctly presented their work as part of the competition, with the award going to Caglayan Dicle of Northeastern University for his work on characterizing dynamical systems. The luncheon included 22 students, paired with 18 mentors, providing students and mentors with the opportunity to discuss their technical work and career plans. Feedback from the student program was overwhelmingly positive.



Figure 4. The new Hands-On Expo, which brought together companies that have ground-breaking new media-related products and world-class multimedia researchers and technologists, including Touradj Ebrahimi and Fernando Pereira (shown here).

Perspectives on Multimedia Research

ICME 16 also featured a panel discussion on “Multimedia Research and Products: Increasing Impact,” which was moderated by Adriana Dumitras (Skype) and included leaders from each of the four IEEE societies: Touradj Ebrahimi of EPFL (representing IEEE SPS), Jenq-Neng Hwang of the University of Washington (IEEE CAS), Haohong Wang of TCL Research America (IEEE ComSoc), and John R. Smith of IBM Watson (IEEE CS). The panelists offered their views on the most significant advances in multimedia over the past 5–10 years and discussed what can we do as a community to further advance multimedia research and its impact on society. The future impact of machine learning was discussed extensively, and all panelists agreed that making more data available to the research community was critical to its success.

Best Paper Awards

Overall, there were many interesting research papers presented—too many to cover here—and the Technical Program Chairs and the ICME 2016 Paper Award Committee followed a rigorous evaluation process to select the following Best Paper award recipients.

The Best Paper award was given to “Phonetic Posteriorgrams for Many-to-One Voice Conversion without Parallel Data Training,” by Lifa

Another highlight of the industry program was a forum discussion on “Ultra HD—Roadmap of High Quality A/V Content to the Home,” which was moderated by Dolby’s Vice President of Technology, Patrick Griffis. Panelists included Don Eklund (the Executive Vice President of Advanced Technologies at Sony Pictures), Jens Peter Wittenburg (Director at Technicolor), and Nandhu Nandhakumar (Senior Vice President at LG Electronics). Eklund spoke about the production of HDR content; Wittenburg discussed challenges in



(a)



(b)

Figure 5. The technical program was coupled with an excellent social program, which included (a) a welcome reception at the Space Needle (here I'm shown on the left, with Wenwu Zhu on the right) and (b) an awards banquet, where, among the various awards presented, Yong Rui gave this year's Best Department Article award for *IEEE MultiMedia* to Ramesh Jain for "Let's Weave the Visual Web."

Sun, Kun Li, Hao Wang, Shiyin Kang, and Helen Meng. The paper proposed an approach to voice conversion (converting the speech of one speaker to another speaker's voice) using an unpaired source speaker and target-speaker training data. The algorithm bridges between speakers using Phonetic PosteriorGrams (PPGs), obtained from a speaker-independent automatic speech recognition system. The idea is novel and can be applied to a variety of applications, including personalized speaking aids and movie dubbing.

There were two Best Student Paper awards. One went to "Large-Scale Vehicle Re-Identification in Urban Surveillance Videos," by Xinchun Liu, Wu Liu, Huadong Ma, and Huiyuan Fu. This paper examined the problem of vehicle re-identification (Re-Id) across a large number of traffic cameras. One of the major contributions is a new large-scale benchmark dataset for vehicle Re-Id, which contains more than 40,000 bounding boxes of 619 vehicles captured by 20 cameras in unconstrained traffic scenes. Each vehicle was captured by 2–18 cameras in different viewpoints, illuminations, and resolutions to provide a high recurrence rate. The paper also presented a vehicle Re-Id method that combines the texture, colors, and high-level attributes information as a baseline on the dataset.

The other Best Student Paper award went to "Blind Quality Assessment of Compressed Images via Pseudo Structural Similarity," by Xionghuo Min, Guangtao Zhai, Ke Gu, Yuming Fang,

Xiaokang Yang, Xiaolin Wu, Jiantao Zhou, and Xianming Liu. This paper explored the pseudo structures of images compressed using block-based methods. It presented an interesting way to evaluate the quality of compressed images via the similarity between the pseudo structures of two images. The proposed pseudo structural similarity (PSS) model works well not only on natural scene images but also on screen content images.

The outstanding technical program was coupled with an excellent social program, including a welcome reception at the Space Needle, which gave attendees a great opportunity to interact with fellow colleagues while enjoying panoramic views of Seattle, as well as an awards banquet that recognized outstanding work of our community (just described), as well as Best Paper awards from *IEEE Transactions on Multimedia* and *IEEE MultiMedia* (see Figure 5).

The ICME 2016 conference proceedings are available on IEEE Xplore. Next year's event will be held in Hong Kong—you're warmly invited to submit your work and get involved. Please visit the website for the latest information on paper submission deadlines and other important dates (www.icme2017.org). **MM**

Anthony Vetro is a deputy director at Mitsubishi Electric Research Labs and served as one of the technical program co-chairs of ICME 2016. Contact him at avetro@merl.com.