



## Call for Full and Short Papers

The 23<sup>rd</sup> **ACM** International Conference on **Multimedia**

Brisbane, Australia | October 2015

### Overview

ACM Multimedia 2015, to be held on the Southern Hemisphere for the first time, calls for **full** papers presenting novel theoretical and algorithmic solutions addressing problems across the domain of multimedia and related applications. The conference also calls for **short** papers presenting novel, thought-provoking ideas and promising (preliminary) results in realizing these ideas.

Submissions are invited in the following **15** topic areas grouped into 4 themes:

### Systems

Research in multimedia systems is generally concerned with understanding fundamental tradeoffs between competing resource requirements, developing practical techniques and heuristics for realizing complex optimization and allocation strategies, and demonstrating innovative mechanisms and frameworks for building large-scale multimedia applications. Within this theme, we have focussed on four target topic areas:

- Mobile Multimedia
- Multimedia Systems and Middleware
- Multimedia Telepresence and Virtual Reality
- Multimedia Transport and Delivery

### Experience

One of the core tenants of our research community is that multimedia data contributes to the user experience in a rich and meaningful manner. The topics organized under this theme are concerned with innovative uses of multimedia to enhance the user experience, how this experience is manifested in specific domains, and metrics for qualitatively and quantitatively measuring that experience in useful and meaningful ways. Specific topic areas addressed this year include:

- Multimedia for Collaboration in Education and Distributed Environments
- Multimedia HCI and QoE
- Music, Speech, and Audio Processing
- Multimedia Authoring and Enrichment

### Engagement

The engagement of multimedia with society as whole requires research that addresses how multimedia can be used to connect people with multimedia artifacts that meet their needs in a variety of contexts. The topic areas included under this theme include:

- Multimedia Art, Entertainment and Culture
- Multimedia Search and Recommendation
- Social Multimedia
- Emotional and Social Signals In Multimedia

### Understanding

Multimedia data types by their very nature are complex and often involve intertwined instances of different kinds of information. We can leverage this multi-modal perspective in order to extract meaning and understanding of the world, often with surprising results. Specific topics addressed this year include:

- Deep Learning for Multimedia
- Multimodal Analysis and Description

- Multimedia and Vision

Full/short papers submissions must conform with a **"Double-blind Review"** process, such that authors should not know the names of the reviewers of their papers, and reviewers should not know the names of the authors. Please prepare your paper in a way that preserves anonymity of the authors.

- Do not put your names under the title.
- Avoid using phrases such as "our previous work" when referring to earlier publications by the authors.
- Remove information that may identify the authors in the acknowledgments (e.g., co-workers and grant IDs).
- Check supplemental material (e.g., titles in the video clips, or supplementary documents) for information that may identify the authors identity.
- Avoid providing links to websites that identify the authors.

Submission Website: <https://cmt.research.microsoft.com/acmmm2015>

## Important Dates

Manuscript for Full Papers	March 31, 2015
Manuscript for Short papers	April 30, 2015
Initial reviews to authors (only for Full Papers)	May 31, 2015
Rebuttal period (only for Full Papers)	June 1-10, 2015
Author-to-Author's Advocate contact period	June 1-6, 2015
Notification of Acceptance	July 6, 2015
Camera-ready submission	August 15, 2015

**Important note for the authors:** The official publication date is the date the proceedings are made available in the ACM Digital Library. This date may be up to two weeks prior to the first day of the conference. The official publication date affects the deadline for any patent filings related to published work.

## Contact

### Chairs :

- **Dick Bulterman**, FXPal, USA
- **Ketan Mayer-Patel**, The University of North Carolina, USA
- **Heng Tao Shen**, The University of Queensland, Australia
- **Shuicheng Yan**, National University of Singapore, Singapore

For any questions, email [technical.program@acmmm.org](mailto:technical.program@acmmm.org)

## Areas of Interest

### Systems

#### Mobile Multimedia

With a multitude of sensors, including accelerometer, GPS-location, multiple video and still cameras, microphone and speakers, mobile devices are arguably the truest embodiment of multimedia. Furthermore, as processing power, camera quality, and display resolution continue to improve at an exponential pace, the potential of these devices is seemingly unbounded. At the very same time, however, much more limited growth in battery life and communication capacity create distinct systems-level challenges that must be addressed to tap that potential.

Topics include but are not limited to:

- Media streaming to/from mobile devices
- Innovative uses of location and accelerometer sensors within multimedia applications
- Mobility models and simulation in service of multimedia research
- Performance and quality-of-experience studies of multimedia applications in a mobile context
- Peer-to-peer mobile media
- Low-power and power-aware media coding and multimedia applications
- Real-time interactive mobile-to-mobile conferencing
- Crowdsourcing within mobile multimedia applications
- Touch-based interfaces for multimedia applications
- Vehicle-based multimedia systems

#### Multimedia Systems and Middleware

This area targets applications, mechanisms, algorithms, and tools that enable the design and development of efficient, robust and scalable multimedia systems. In general, it includes solutions at various levels in the software and hardware stack.

We call for submissions that explore the design of architectures and software for mobile multimedia, multimedia in pervasive computing applications and large-scale and/or distributed multimedia systems. This includes tools and middleware to build multimedia applications, like content adaptation and transcoding, stream processing, scheduling and synchronization and cloud multimedia systems.

Multimedia architectures and systems are continuously evolving and hardware technology changes influence middleware and applications. We therefore also solicit submissions to new research on host-device interaction in heterogeneous systems, applications for transactional memory, multi-level memory architectures (e.g., RAM - SSDs - spinning disks), for operating systems and storage functions.

Topics of interest include, but are not limited to:

- Efficient implementations of and processing frameworks for multimedia workloads
- System and middleware implementation with graphics processing units (GPUs), network processors and field-programmable gate arrays (FPGAs)
- Multimedia systems over new generation computing platforms, e.g., social, cloud
- Wireless mobile multimedia systems
- Real-time multimedia systems and middleware
- Embedded multimedia systems and middleware
- QoS of multimedia systems and middleware
- System prototypes and deployment
- Energy-efficiency for multimedia systems and middleware
- Multimedia storage systems
- Immersive and virtual world systems and middleware

Submitted contributions should either provide new and interesting theoretical or algorithmic insights of specific aspects of multimedia systems, or they may contain a comprehensive evaluation of an overall multimedia system.

#### Multimedia Telepresence and Virtual Reality

Telepresence and virtual reality have long been grand challenges for researchers and industry alike. High-resolution, 3D telepresence can dramatically improving sense of presence for interpersonal and group communication which is paramount for supporting non-verbal and subconscious communication that is currently lost in video and audio conferencing environments. Realistic virtual reality enables a wide spectrum of important applications including tele-medicine, training for hazardous situations, scientific visualization, and engineering prototyping. Addressing the challenges of telepresence and

virtual reality requires the development of new media representation and streaming techniques as well as innovations in human-computer interaction.

Topics include but are not limited to:

- Multi-camera coding and streaming
- 3D video coding
- Image-based rendering for virtual environments
- Virtual reality user interface design and evaluation
- Haptic interfaces for virtual reality
- Virtual-world design and authoring tools
- 3D sound rendering in virtual environments
- Multi-viewpoint stereo for group telepresence
- Automated group telepresence capture and control
- Distributed multi-user virtual reality systems
- Real-time bandwidth adaptation for VR and telepresence.
- Innovative VR and telepresence applications
- Quality of experience models and evaluation for VR and telepresence.

## Multimedia Transport and Delivery

The area of Media Transport and Delivery (MT&D) invites research that is concerned with the mechanisms that are used to move information through public networks like the Internet, as well as the placement and movement of information within CDNs, P2P networks, Clouds, clusters, or even within a single computer, with the goal of enabling multimedia applications. Such information movement is today no longer content-agnostic: mechanisms may adapt, filter or combine content, and they may even organize movement based on content types or content semantics.

Topics of interest include, but are not limited to:

- New, innovative, and disruptive research results in the application of information centric networks, named data networking, publish/subscribe information networking, opportunistic networking and disruption tolerant networking
- New deployment concepts, such as network function virtualization and software defined networking
- Adaptive HTTP streaming and Web RTC drive protocol
- New research with a view towards the fight against bufferbloat and new congestion management methods including AQM strategies
- Performance improvement due to new forms of host-device interaction, including the benefits of new interconnects, transactional memory, and SSD controllers allowing random memory access
- Transport of multimodal multimedia data and other interactive applications, including the relation of MT&D with scalable media adaptation, scaling, compression, and coding
- Multimedia content-aware pre-fetching and caching, multimedia analysis and recommendations for media distribution and caching purposes

The proposed contributions are expected to provide new theoretical or experimental insights into individual MT&D mechanisms, enhancements to one or more system components, complete systems, or applications with a strong contribution to understanding MT&D.

We encourage submissions that touch upon the MT&D needs of hot application areas such as immersive systems, multimodal sensor networks, office of the future, healthcare, virtual or augmented reality, 3D video or graphics, networked games, collaboration systems, social networks, multimedia sharing services, Massive Open Online Courses (MOOC), mobile multimedia and cloud-based systems.

## Experience

### Multimedia for Collaboration in Education and Distributed Environments

On-line and distance collaboration platforms and applications have taken center stage as people are trying to find efficient ways to deliver materials to distributed audiences. This is especially true in the area of Education, but also touches many forms of spatially (and temporally) distributed information sharing.

The submissions in this area should look specifically at methods and systems where remote collaboration support provides a fundamental technology. This can be for any relevant area, with an emphasis on tools to support education in its broadest sense. Examples include applications to address the challenge of mass delivery of basic education in the world's poorest and most rural countries. Multimedia educational resources have the potential to provide interactive and engaging instruction that adapts to students as they learn at both ends of this spectrum.

Topics include but are not limited to:

- Multimedia for distance learning
- Real-time multimedia to support educators
- Multimedia analytics of classrooms
- Automated meeting capture and authoring
- Multimedia presentation tools for remote sharing
- User studies of multimedia education or distributed communications applications
- Multimedia technologies in support of remote collaboration in poor and rural environments

### Multimedia HCI and QoE

There is a growing evolution of media towards interactivity, which is prompted by the intrinsically interactive nature of devices used for media consumption, as well as progress in media content description that makes it amenable to direct access and manipulation. The same advances also provide a basis for capturing and understanding user experience.

The Multimedia HCI and QoE area follows on from similar areas in previous editions of the conference, which have been dedicated to Human-Centered Multimedia, Media Interactions, or Multimedia HCI. The specific inclusion of Quality of Experience (QoE) recognizes the need for user studies to focus on the most specific aspects of media consumption.

Topics of interest include, but are not restricted to:

- Design and implementation of novel interactive media: interactive films and narrative, storyfication of social media
- Human-centred multimedia, including immersive multimedia
- Novel interaction modalities for accessing media content, including multimodal, affective, and brain-computer interfaces
- Systems and architectures for multimodal and multimedia integration
- User interaction for media content authoring or media production
- Subjective assessment methodologies to estimate the QoE in multimedia systems
- Influencing parameters, models and objective metrics to measure QoE in multimedia systems
- System design and implementations taking advantage of a direct multimedia and multimodal QoE measurements
- Datasets, benchmarks and validation of multimedia quality of experience

We expect all papers to include a substantial media element, characterized by either media content, production or consumption modalities. For instance, while we welcome contributions on immersive media, papers reporting virtual and augmented reality systems will only be eligible if their informational content is media-based. Similarly, work based on mainstream interactive media, such as computer games, will have to demonstrate originality and added value in terms of media production or consumption.

As in previous years, we expect papers to include evaluation studies. However, to avoid stifling creativity and innovation, such studies should be adapted to the paper's objectives and focus. Papers centered on user experience are expected to include user studies with rigorous methodology: sample size can however be adapted to the nature of the study (usability or psychological findings) and real-world constraints (when subjects' recruitment is limited, e.g. domain experts, artists...). Papers describing new interaction techniques can include a mix of performance evaluations and usability studies. Papers taking an integrated systems approach can include an evaluation primarily based on performance, enhanced with narrative user feedback.

## Music, Speech and Audio Processing in Multimedia

As a core part of multimedia data, the acoustic modality is of great importance as a source of information that is orthogonal to other modalities like video or text. Incorporating this modality in multimedia systems allows for richer information to be extracted when performing multimedia content analysis and provides richer means for interacting with multimedia data collections.

In this area, we call for strong technical submissions revolving around music, speech and audio processing in multimedia. These submissions may address analysis of the acoustic signals in order to extract information from multimedia content (e.g. what notes are being played, what is being said, or what sounds appear), or from the context (e.g. language spoken, age and gender of the speaker, localization using sound). They may also address the synthesis of acoustic content for multimedia purposes (e.g. speech synthesis, singing voices, acoustic scene synthesis), or novel ways to represent acoustic data as multimedia, for example, by combining audio analysis with recordings of gestures in the visual channel. We are also interested in the submissions addressing novel multimedia interfaces and interaction concepts enabled by the inclusion of acoustics, as well as the changing paradigms of analyzing, indexing, organizing, searching, recommending, consuming and enjoying music by taking into account contextual, social and affective aspects and content from other modalities or other information sources in general.

While the acoustic modality is central in this area, the submissions should consider this modality in the multimedia context, for instance by developing methods and approaches addressing multimedia items, applications or systems, or by explicitly considering information sources from different modalities.

Topics of interest include, but are not restricted to:

- Multimodal approaches to audio analysis and synthesis
- Multimodal approaches to audio indexing, search, and retrieval
- Multimodal and multimedia context models for music, speech, and audio
- Computational approaches to music, speech, and audio inspired by other domains (e.g. computer vision, information retrieval, musicology, psychology)
- Multimedia localization using acoustic information
- Social data, user models and personalization in music, speech, and audio
- Music, audio, and aural aspects in multimedia user interfaces
- Multimedia and/or interactive musical instruments and systems
- Multimedia applications around music, speech and audio
- Music, speech, and audio coding, transmission, and storage for multimedia applications

## Multimedia Authoring and Enrichment

Multimedia content can be created and integrated in many ways. Isolated bits of content can be created (and published) by media capture software, but complex combinations of media can also be created using authoring systems and tools. Existing media can be enhanced and enriched over time, defining another method for rich content creation and management.

In this area, we call for strong technical submissions that define innovative ways of explicitly creating, extending, enhancing and enriching multimedia content. These submissions may highlight dedicated media-centric tools or they may describe more lifecycle-centric system that allow content to be extended as a function of use or sharing.

The content in question should either be multi-modal media content in the most general sense or systems that take single modalities and enrich them (possibly over time) with additional content types.

Topics of interest include, but are not restricted to:

- Multimodal media combination systems
- Multimodal media personalization systems, where personalization is considered an authoring activity
- Systems for generating new versions of existing content as a function of (social) use
- Systems for content use analysis, if this leads to incremental content creation
- Systems for enriching content, either for all viewers or for dedicated sub-communities
- Systems for tracking and updating content based on analysis of use



# Engagement

## Multimedia Art, Entertainment and Culture

The focus of this area is on the innovative use of digital multimedia technology in arts, entertainment and culture, to support the creation of multimedia content, artistic interactive and multimodal installations, the analysis of media consumption and user experience, or cultural preservation. We seek full and short papers in a broad range of integrated artistic and scientific statements describing digital systems for arts, entertainment, and culture. Successful papers should achieve a balance between sophisticated technical content and artistic or cultural purpose.

Topics of interest include, but are not limited to:

- Models of interactivity specifically addressing arts and entertainment
- Active experience of multimedia artistic content by means of socio-mobile multimodal systems
- Analysis of spectator experience in interactive systems or digitally-enhanced performances
- Virtual and augmented reality artworks, including hybrid physical/digital installations
- Dynamic, generative and interactive multimedia artworks
- Creativity support tools
- Computational aesthetics in multimedia and multimodal systems
- Tools for or case studies on cultural preservation or curation

Papers addressing entertainment applications should clearly advance the state-of-the-art in multimedia technologies and report original forms of media consumption, extending beyond current practice in digital entertainment and computer games. We welcome papers in all areas of multimedia and multimodal systems for art or cultural engagement characterized by innovative multimodal interaction and multimedia content processing.

For those papers describing fully implemented systems, extensive user evaluation is not a strict condition for acceptance provided the levels of performance achieved can be clearly stated. Papers centered on user experience should follow rigorous standards of evaluation.

In general, the following evaluation criteria will be followed for the papers submitted in this area:

- Novel computational and/or technology approach
- Original concept and execution
- Clear description of societal impact and/or audience reception
- High aesthetic quality

Authors are encouraged to critically examine the impact of their work, revealing challenges and opportunities of rich societal significance, in arts, entertainment and culture, including cross-fertilization with art and multimedia.

## Multimedia Search and Recommendation

In the past decade, there has been an explosive growth of multimedia contents on the Web, desktops, and mobile devices. The deluge of multimedia leads to "information overload" and poses new challenges and requirements for effective and efficient access to multimedia content. Multimedia search and recommendation techniques are essential in order to provide information relevant to users' information needs.

This area calls for contributions on reporting novel problems, solutions, models, and/or theories that tackle the key issues in searching, recommending, and discovering multimedia content, as well as a variety of multimedia applications based on search and recommendation technologies.

Topics of interest include, but are not restricted to:

- Large-scale multimedia indexing, ranking, and re-ranking
- Novel representation and scalable quantization for efficient multimedia retrieval
- Interactive and collaborative multimedia search
- Search and recommendation for social media data
- User intent modeling, query suggestion and feedback mechanisms
- Multimedia search in specific domains (e.g., scientific, enterprise, social, fashion)
- Summarization and organization of multimedia collections
- Knowledge discovery from massive multimedia data
- Data representations for recommendation tasks
- Multimedia-focused recommendation models
- Cross-modal recommendations
- Personalization in recommendation

- New interaction strategies for recommendation
- Diversity in search and recommendation
- Generalization of recommendations to new users and/or new content (cold start)

## Social Multimedia

This area seeks novel contributions investigating online social interactions around multimedia systems, streams, and collections. Social media (such as Facebook, Twitter, Flickr, YouTube etc.) has substantially and pervasively changed the communication among organizations, communities, and individuals. Sharing of multimedia objects, such as images, videos, music, and associated text messages, constitutes a prime aspect of many online social systems nowadays. This gives us valuable opportunities to understand user-multimedia interaction mechanisms, to predict user behavior, to model the evolution of multimedia content and social graphs, or to design human-centric multimedia applications and services informed by social media.

The submissions in this area should look specifically at methods and systems wherein social factors, such as user profiles, user behaviors, social relations are organically integrated with online multimedia data to understand media content and media use in an online social environment. Or they should leverage the socially created data to solve challenging problems in traditional multimedia computing, or address new research problems emergent in the online social media scenario.

The proposed contributions are expected to scale up to serve large online user communities. They should exploit massive online collective behavior by looking at e.g., large-group online interactions and group sentiments aggregated across many users in an online community. They should also be able to handle large, heterogeneous and noisy multimedia collections typical for social media environments. Special emphasis is put on multimodal approaches leveraging multiple information sources and modalities to be found in the social media context.

Topics of interest include, but are not restricted to:

- Social media data collection, filtering, and indexing
- Social media data representation and understanding
- User profiling from social media
- Modeling collective behavior in social media
- Multimedia propagation in online social environments
- Spatial-temporal context analysis in social media
- Prediction and forecasting with social media
- Multimedia-enabled social sharing of information
- Detection and analysis of emergent events in social media collections
- Evaluation of user engagement around shared media
- Systems and analysis of location-based social media
- Network theory and algorithms in social multimedia systems
- Models for the spatio-temporal characteristics of social media
- Models and systems for analyzing large-scale online sentiments



## Emotional and Social Signals In Multimedia

A lot of multimedia systems capture human behavior, and in particular, social and emotional signals. These systems would therefore benefit from the ability to automatically interpret and react to the social and emotional context. The interpretation, analysis, and synthesis of social and emotional signals requires a different expertise that draws from a combination of signal processing, machine learning, pattern recognition, and behavioral and social psychology. Analyzing multimedia content, where humans spontaneously express and respond to social or affect signals, helps to attribute meaning to users' attitudes, preferences, relationships, feelings, personality, etc., as well as to understand the social and affective context of activities occurring in people's everyday life.

This area focuses on the analysis of emotional and interactive social behavior in the spectrum of individual to small group settings. It calls for novel contributions with a strong human-centered focus specializing in supporting or developing automated techniques for analyzing, processing, interpreting, synthesizing, or exploiting human social and affective signals for multimedia applications. Special emphasis is put on multimodal approaches leveraging multiple streams when analyzing the verbal and/or non-verbal social and emotional signals during interactions. These interactions could be remote or co-located, and can include e.g. interactions between multiple people, humans with computer systems/robots, or humans with conversational agents.

Topics of interest include, but are not restricted to:

- Human social, emotional, and/or affective cue extraction
- Cross-media and/or multimodal fusion of interactive social and/or affective signals
- The analysis of social and/or emotional behavior
- Novel methods for the interpretation of interactive social and/or affective signals
- Novel methods for the classification and representation of interactive social and/or emotional signals
- Real-time processing of interactive social and emotional signals for interactive/assistive multimedia systems
- Emotionally and socially aware dialogue modeling
- Affective (emotionally sensitive) interfaces
- Socially interactive and/or emotional multimedia content tagging
- Social interactions and/or affective behavior for quality of delivery of multimedia systems
- Collecting large scale affective and/or social signal data
- Multimedia tools for affective or interactive social behavior
- Facilitating and understanding ecological validity for emotionally and socially aware multimedia
- Annotation, evaluation measures, and benchmarking
- Dyadic or small-group interaction analysis in multimedia

## Understanding

### Multimedia and Vision

Advances in computer vision coupled with increased multicore processor performance have made real-time vision-based multimedia systems feasible. The interaction of commodity off-the-shelf components and codecs adapting to real-world bandwidth limits with the performance and stability of vision algorithms remains an important research area. Next generation codecs need to be investigated and designed to support the needs of vision-based systems which may be sensitive to different kinds of noise than human consumers. Another rich avenue of research is the integration of related non-visual media such as audio, pose estimates, and accelerometer data into vision algorithms.

Topics include but are not limited to:

- Video coding and streaming in support of vision applications
- Compression tradeoffs and artifacts within vision algorithms
- Integration of non-video information sources with vision
- Vision-directed compression
- Vision-based video quality models
- Multimedia signal processing for vision
- Vision-based multimedia analytics
- Distributed and coordinated surveillance systems
- Inter-media correspondance and geotagging
- Multimedia vision applications and performance evaluation

### Deep Learning for Multimedia

Deep Learning is an emergent field of Machine Learning focusing on learning representations of data. Deep Learning has recently found success in a variety of domains, from computer vision to speech recognition, natural language processing, web search ranking, and even online advertising. Deep Learning's power comes from learning rich representations of data that can be tuned for the task of interest. The ability of Deep Learning methods to capture the semantics of data is however limited by both the complexity of the models and the intrinsic richness of the input to the system. In particular, current methods only consider a single modality leading to an impoverished model of the world. Sensory data are inherently multimodal instead: images are often associated with text; videos contain both visual and audio signals; text is often related to social content from public media; etc. Considering cross-modality structure may yield a big leap forward in machine understanding of the world.

Learning from multimodal inputs is technically challenging because different modalities have different statistics and different kinds of representation. For instance, text is discrete and often represented by very large and sparse vectors, while images are represented by dense tensors that exhibit strong local correlations. Fortunately, Deep Learning has the promise to learn adaptive representations from the input, potentially bridging the gap between these different modalities.

In this track, we encourage submissions that effectively deploy Deep Learning to advance the state of the art across the domain of multimedia and related applications.

Topics of interest include, but are not restricted to:

- Deep learning application involving multiple modalities, such as images, videos, audio, text, clicks or any other kind of (social) content and context
- Deploying deep learning to learn features from multimodal inputs
- Deploying deep learning to generate one modality from other modalities
- Deep learning based methods that leverage multiple modalities and also account for temporal dynamics
- Deploying deep learning to increase the robustness to missing modalities

## Multimodal Analysis and Description

Analysis of multimedia content enables us to better understand what the content is about in order to improve its indexing, representation, and consumption for the purpose of retrieval, content creation/enhancement and interactive applications. Research so far has mostly focused on mono-modal analysis of multimedia content, such as looking only into images, only into text, or only into video, but ignoring other modalities like the text floating around an image on a web page or the audio accompanied with the video.

The goal of this area is to attract novel multimodal analysis research that takes multiple modalities into account when it comes to multimedia content analysis and better description of the multimedia content. The different modalities may be temporally synchronized (e.g., video clips and corresponding audio transcripts, animations, multimedia presentations), spatially related (images embedded in text, object relationships in 3D space), or otherwise semantically connected (combined analysis of collections of videos, set of images created by one's social network).

This area calls for submissions that reveal the information encoded in different modalities, combine this information in a non-trivial way and exploit the combined information to significantly expand the current possibilities for handling and interacting with multimedia content. In addition, the submitted works are expected to support effective and efficient interaction with large-scale multimedia collections and to stretch across mobile and desktop environments in order to address changing demands of multimedia consumers.

Topics of interest include, but are not restricted to:

- Novel strategies for multimodal analysis of multimedia data
- Multimodal feature extraction and fusion
- Multimodal semantic concept detection, object recognition and segmentation
- Multimodal approaches to detecting complex activities
- Multimodal approaches to event analysis and modeling
- Multimodal approaches to temporal or structural analysis of multimedia data
- Machine learning for multimodal analysis
- Scalable processing and scalability issues in multimodal content analysis
- Advanced descriptors and similarity metrics exploiting the multimodal nature of multimedia data