ACM International Workshop on Social and Behavioral Networked Media Access  
(SBNMA’11)

Naeem Ramzan  
Queen Mary University of London  
London, United Kingdom  
naeem.ramzan@eecs.qmul.ac.uk

Fei Wang  
IBM T.J.Watson Research Center, USA  
fwang@us.ibm.com

Charalampos Z. Patrikakis  
Technological Education Institute of Piraeus, Greece  
bpatr@telecom.ntua.gr

Peng Cui  
Tsinghua University, China  
cui@tsinghua.edu.cn

Nikolaos Doulamis  
National Technical Univ. of Athens, Greece  
ndoulam@cs.ntua.gr

Shiqiang Yang  
Tsinghua University, China  
yangshq@tsinghua.edu.cn

Gordon Sun  
Tencent Technologies, China  
gordonsun@tencent.com

ABSTRACT
In an endeavour to speak and prevail over some of the open problems that obstruct efficient networked media, this workshop will fetch together folks from a number of research communities, including but not limited to Multimedia Distribution and Access, Social Network Analysis, Multimedia Content Analysis, Behavioral Analysis, User Modelling Adaptation and Personalization. It is our credence that a synergetic approach involving the above mentioned research areas can surpass their individual potentials, leading to improved networked media access. The main objective of this workshop is to provide a forum to disseminate work that explicitly exploit the synergy between multimedia content analysis, behavioral modelling, personalisation, and next generation networking and community aspects of social networks. This synergetic methodology could produce high quality of experience for personalized multimedia access in networking environment.

Categories and Subject Descriptors
H.5.1 Multimedia Information Systems

General Terms

Keywords
Social, behavioral, adaptation, personalization, interaction, access.

1. INTRODUCTION
In recent years, there has been an unprecedented increase in the creation and consumption of digital information. In addition, people are increasingly well connected on the Internet through some social network sites, blogs, wikis and mashups. Some popular social and networked media platforms such as Facebook, Flicker, YouTube and Windows Live provide virtual platforms for all users to share their ideas, photos, videos and music, and all users are collaborating together to contribute to media production and sharing. In these networked media platforms, a large volume of users actively interact with each other, which not only influence each individual’s behavior on information utilization, but also affect the system’s performance. To improve user experience and level of service, it is of ample importance to analyze the impact of social and human behavioural factors on networked media platforms. This environment presents clear challenges to conventional information retrieval research, which is based mainly on individual activities.

2. CONTRIBUTION
This workshop aims at presenting work focusing on two major topics: The first is to address the question on how multimedia content analysis can be combined with information derived from behavioural modelling and social interactions in order to improve personalised content distribution. The second relates to the way in which personalised services can be offered to users, in a real time, ubiquitous seamless and comprehensive way, through a comprehensive context framework that targets Quality of Experience rather than the traditionally sought Quality of Service.

The main goal of this workshop is to provide a forum to disseminate work that explicitly exploit the synergy among multimedia content analysis, behavioural modelling, personalisation, next generation networking and community aspects of social networks by reviewing pressing needs, examining challenging research issues, and showing the state-of-the-art research and development in networked media information retrieval environment.
Hence, this workshop is distinct from and complementary to previous initiatives in that it involves the integration of user’s behavioural modeling, multimedia content analysis, and seamless delivery techniques with information derived from networked communities and context awareness. The workshop will cover the following topics:

### 3. TOPICS OF INTEREST

This workshop is distinct from and complementary to previous initiatives in that it involves the integration of multimedia content analysis techniques with information derived from users, networked communities, and context awareness. The aim of the workshop is to cover the following topics:

- Techniques for robust and scalable distribution of multimedia content.
- Social behaviour methods in the area of multimedia search, content delivery and scalable streaming.
- Fundamental (theoretical and practical) challenges in behavioural modelling for networked media information retrieval.
- Recent advances in behavioral modeling on networked media.
- Robust distribution of multimedia services over heterogeneous networks and access technologies.
- QoE of the multimedia communication over social/P2P networks.
- Social networks analysis to multimedia content personalization and adaptation.

### 4. WORKSHOP SUMMARY

The workshop is featured with 2 keynote talks, 7 full research and 4 invited papers. Due to the high amount of accepted and invited papers, it was decided to allocate time slot of 30 mins for invited paper and 20 mins for full research papers. The papers accepted for presentations are given below.

A number of approaches have investigated the multimedia content personalization and adaptation by social network analysis. In the paper entitled "A Hierarchical Multimodal Approach for Placing videos on the Maps using Millions of Flickr Photographs" Kelm et al. present a hierarchical, multi-modal approach for placing non-geotagged Flickr video on the map. Zhang et al. explore an approach that exploits YouTube video co-watch data to improve the performance of video taxonomic classification system in their paper “Improving Video Classification via YouTube Video Co-Watch Data”. In “Using Media Related User Profiles to Personalize Multimedia Access over Social Networks” Argyriou et al. address the issue of personalized media access over the internet. Hu et al. in “Event Analytics via Social Media” analyze the response of users to public events through social media. Voulodimos et al. in “Employing Clustering Algorithms to Create User Groups for Personalized Context Aware Services Provision” present user preferences in the scope of social groups.

Accepted papers also focus on social behavior methods in the area of multimedia distribution. Brenner et al. in “Graph-Based Recognition in Photo Collections Using Social Semantics” show how to recognize people in consumer photo collections by employing a graphical model. Marian and Mathias in their paper “Pursuing the Holy Grail by interrelating user intentions and Bag of Visual Words to perform retrieval adaptation” present how the user’s search behavior enhance the multimedia retrieval and adaptation. Alan Gomes presents how to represent and measure social interactions as behavioral contingencies in his paper “Measuring Media-based Social Interactions Provided by Smartphones Applications in Social Networks”. Mirza et al. investigate in the use of eye tracking techniques in order to personalize image annotation and retrieval in “Gaze Movement Inference for User Adapted Image Annotation and Retrieval”.

Savas et al. in “Adaptive Multi-view Video Streaming over P2P Networks Considering Quality of Experience” present a mesh-based P2P streaming architecture for Multi-view video that employs rate adaptation according to the findings to deliver the best QoE under diverse network conditions. Finally, in "Social and Behavioral Media Access: A Survey" Ramzan gives his position on the expected evolution of social and behavioral networked media access.

### 5. ACKNOWLEDGEMENTS

This workshop was partially supported by its affiliated projects: Saracen (FP7-248474), 3DLife (FP7-247688), PetaMedia (FP7-216444) and SA EMC² (FP7-287663).
ABSTRACT
Intelligent Environments have the vision of enhancing our everyday environment and interaction with its objects by sensing, computing, and communication capabilities. The major characteristics of such environments are the increasing number of embedded intelligent devices (ubiquity) into the background (transparency). These devices are expected to disappear or blend into the background and will be invisible to the user. However, because of this transparency, users fail to develop an adequate mental model for interaction with such environments. The Ubiquitous Meta User Interfaces (Ubi-MUI) ACM workshop provides a venue for the development of highly intuitive, multimedia supported meta user interfaces that bring transparency, predictability, and control into intelligent environments.

Categories & Subject Descriptors: H. Information system: H.1.2 User/machine systems- Human Factors
General Terms: Design, Experimentation, Human Factors

INTRODUCTION
The nature of intelligent environments has transformed the way humans interact with the surrounding objects. Such advancement is calling for new types of user interfaces that create an overall system image for intelligent environments in order to help users to better understand and engage in the environment [1]. In this sense, representative user interfaces are Ubiquitous Meta User Interfaces (Ubi-MUI) that could increase the transparency and predictability of the whole system by visualizing the environment’s internal states, perception and decision making processes. Using Ubi-MUI, users could observe, analyze, understand, control, and customize the adaptive behavior and context-dependent interactions of their surrounding. In such environments, multimedia play two key roles: they support new ways of interaction that apply to multiple human senses, and they diffuse the content presentation in the user’s environment. Addressing the challenges of such research venue is a community-spanning effort, necessitating the pooling of resources and experiences of different research groups. While workshops related to Multimedia, Ubiquitous Computing and Human Factors exists, the unique requirements of interaction with and within intelligent environments require an interdisciplinary approach. The aim of this workshop thus is to provide a link (1) between these different research groups and (2) between the academic communities and industry groups.

BACKGROUND
Nowadays, there is a growing research community in the field of interactive systems that can be deployed for intelligent environments. The major characteristics of such environments are the increasing number of embedded intelligent devices (ubiquity) into the background (transparency). These environments will have profound impact on the type, content and functionality of the emerging products and services [2]. The research in the area of multimedia support that applies to the human senses in user interfaces is still in its infancy stage for such environments. The majority of user interface design research is based on modeling relative user interface related knowledge (e.g. [3]). While somehow multimedia editors are included into the user interface model and task interpretations for example the work in [4] [5] and [6] to name few, these designs need to bring transparency, predictability, and control into intelligent environments. The open literature describes some studies which focus on human centric user interfaces for smart environments. A survey of such studies can be found in [7]. The methods of user-centered design have been extended to other interaction paradigms as well, as described in [8]. While the interaction between the human and the computer is relatively low in legacy user interfaces, this interaction is much higher with smart environments. The human user can interact using her/his senses. The interaction can actually occupy the same space of the surrounding environment where the user is living. This interaction makes the transfer of the user-centered awareness to become very rich and highly context-dependent [9].
IMPORTANT TOPICS FOR UBI-MUI RESEARCH

This workshop intends to cover research that addresses representation of intelligent environments, concept of meta user interfaces, and implementation and evaluation aspects. Below are the details of such topics in relevant to the Ubi-MUI workshop.

Representing Intelligent Environments. This topic focuses on the state-of-the-art in areas related to visualization and animation of sensing activities, decisions, and implicit interactions of intelligent systems, which allows users to understand and predict the behavior of the system. It touches on awareness creation for implicit interaction concepts using multimedia artifacts. Using avatars, storytelling, or gaming to introduce the functional capabilities and adaptive behavior of intelligent environments to novice users is also included in the scope of this topic.

Concepts for Meta User Interfaces. Currently there is a momentum for multimodal multimedia interfaces to control and modify implicit interactions and the environment’s responsive activities. This workshop extends its coverage to include user interfaces to control and modify system behavior. In this context, the emphasis is given to supportive user interfaces that create a system face for intelligent environments, natural meta interaction concepts allowing users an easy access to multimedia environments (search, exploration, manipulation and control of media and devices), touch and gesture based interfaces, 3D displays and audio immersive systems for augmented-reality. Also, topics such as metaphors and coordination algorithms for distributed conflict management are considered.

Implementation, and Evaluation Aspects. Especially in user-centric interfaces, it is important to provide methods to evaluate the added-value of multimedia assisted meta user interfaces. That is, including user studies related to mental models for human-environment-interaction and meta user interfaces. The workshop seeks for novel approaches to effectively manage the complexity of development, real-time, de-centralized media-processing architectures, middleware architectures for sensor and multimedia integration. Implementation and evaluation that encompasses activity analysis and domain observations related to phenomena such as loss of control and over-automation Smart Multimedia sensors are particularly important.

CONCLUSION

This workshop extends the trend of using multimedia supported meta user interfaces in smart environments. The areas of the research topics described in the previous section are very relevant to Ubi-MUI and expected to increase the richness of the interaction with the smart environments. The main goal of the workshop is to foster collaboration among researchers in their areas of interest related to Ubiquitous Meta User Interfaces and allow a face-to-face exchange of ideas.

REFERENCES

ABSTRACT
The ACM AIEMPro 2011 workshop presents research on automated media content analysis and production for, amongst others, the development of novel TV services. The program of the workshop has two sessions. The first one is composed of three papers on video and TV content structuring and indexing. The second session is also composed of three papers on media production and retrieval systems and applications.

Categories and Subject Descriptors
H.4.0 [Information Systems Applications]: General

General Terms
Algorithm, Design, Experimentation, Performance, Theory

Keywords
Automatic media content analysis and indexing, media production, TV services.

1. INTRODUCTION
The workshop aims at exploring automated information extraction techniques and audiovisual content analysis tools to support future media production for novel TV services. This workshop follows in a series of previous events: the first one in conjunction with DEXA 2008 (AIEMPro08), the second one in conjunction with WIAMIS 2009 (AIEMPro09), and the third one in conjunction with ACM Multimedia 2010 (AIEMPro10). This year, the workshop is the result of a successful merge with the CBTV workshop (held in 2009 in conjunction with the IEEE International Symposium on Multimedia).

The significant increase in the amount of digital video content (TV channels in particular), and the diversification of broadcast possibilities and storage devices, have recently given rise to the emergence of many new services and novel TV programs consumption schemes and usage trends. These new services are aimed at making broadcast content available to consumers according to their needs, i.e. without any constraint on location and/or time and with the possibility to cherry pick and navigate according to the viewers desire. These services have been proven to be very effective in significantly increasing the content audience and open new niches for profitable use of broadcast content. Examples of such services are TV-on-demand, interactive TV, Catch-Up TV... This explosion of new media distribution and consumption paradigms and the corresponding new production workflows based on digital computer-based tools require an immediate revision of the traditional ways of making business in media industry. These evolutions are progressively substituting the traditional one-to-many broadcasting model. To cope with these trends, broadcasters are revolutionizing their point of view, trying to embrace these new models into their production and distribution facilities.

Leading experts all agree in recognizing automation of processes as one of the key for success in this scenario, because of the potential costs reduction introduced by it. This concept is grounded on several concurring factors in the media sector: a) switch over to Digital Television introduces more channels (meaning ore content items are produced/published); b) cross media production (web TV, Hybrid Broadcast/Broadband, multi-screen applications) means reuse of the same material in many different ways; c) improvements in infrastructure (IT) allows better content accessibility; d) recovery of Cultural Heritage implies archive digitization, annotation and multi-platform dissemination; e) budget limitations and shorter time-to-market mean number of media professionals cannot increase.

The cumulative effect of these factors can be summarized as follows: a lot more digital media items need to be produced by the same or reduced staff, and in a quicker and more cost-efficient way. As a consequence, the adoption of tools for intelligent analysis and synthesis of multimedia data are seen as substantial enabling factors in making interactive, multi-channel and multi-purpose productions value-returning.

The workshop aims thus at catalyzing the migration towards new ways of producing, broadcasting, and presenting media content, via the introduction of tools for automated multimedia analysis and understanding. At the same time, the workshop aims at helping academic researchers better understand the real-life key requirements that enable higher impact and wider adoption of these methods.
The program of AIEMPro 2011 includes 6 papers, accepted out of 11 submissions, grouped into two sessions: (1) Media Content Structuring and Indexing, and (2) Media Production and Retrieval Systems and Applications. Though the range of topics is quite diverse, including news story clustering, concept detection in live video streams, TV program structuring, video copy detection, speech recognition, media retrieval systems and architecture. The overall delivered picture is pointing very clearly to a future scenario in which automated analysis of content plays a crucial and indispensable role.

In fact, the careful analysis of this year’s program as well as past editions’ ones [6], makes foresee that future media production and distribution facilities will be integrated with a highly distributed and intelligent content analysis framework capable of supporting all the phases of the value chain. This will mainly allow broadcasters and media producers to make production processes and new services more efficient to be designed, deployed and maintained.

However, what is still missing is an overall and comprehensive view on how these technologies can collaborate with each other in the context of standard frameworks and architectures, and a robust analysis of performances that can be obtained in real use cases, i.e. where content differs from items used for developing algorithms and where new content and service paradigms arise on almost a daily basis.

These latter issues will be object of future efforts by the AIEMPro community at large.

2. WORKSHOP CONTENT SUMMARY

This Section gives some more detail about each of the accepted papers.

Session 1 presents research on Media Content Structuring and Indexing. Paper [2] proposes a sequence-based kernel approach for on-line high-level concept detection on live video streams. The technique efficiently computes the kernel function for an updated value for the current sliding window position based on information about matching elements from previous positions. The approach is tested on the TRECVID 2007 by analyzing, for distinct kernel functions, the required information needed to efficiently evaluate the kernel in the next sliding window, the memory and runtime complexity.

Paper [3] proposes a clustering approach for news stories to be used for news TV channels. It consists of several steps: shot boundary detection, commercial break detection/filtering, anchorperson detection, story aggregation and clustering. The bag of visual-words model and the bag of trajectory model are then used in order to describe objects present in news stories.

Paper [1] addresses the problem of unsupervised detection of recurrent audio segments in TV programs toward program structuring. Recurrent segments are the key elements in the process of program structuring. The paper focuses on a repeated sequence detection method and a set of basic audio descriptors. The objective of the study is to evaluate the effectiveness of considered method and descriptors and their ability to handle the wide range of existing audio recurrences. The idea is provide readers with a synthesis that allows them choosing the appropriate method for each kind of recurrences.

Session 2 deals with Media Production and Retrieval systems and applications. In Paper [5], a novel search environment is presented which collects as much metadata as possible in one single retrieval system. Authors explain how speech recognition can facilitate unlocking the archive by illustrating experiments and measuring the keyword recognition rate, rather than the pure word error rate. Conclusions are that keyword recognition rate is sufficient for efficient media retrieval in a search application.

Paper [7] deals with the picture-in-picture copy detection problem. It proposes a generalized spatial coding representation in which both the relative position and relative orientation is embedded in the spatial code of images key points. A novel formulation for spatial verification problem is also discussed and polynomial and non-polynomial algorithms to efficiently address the spatial verification problem are also introduced. Experiment results on TRECVID and MSRA datasets shows a 20% improvement over the classical hierarchical bag-of-words approach.

Paper [4] describes supporting most aspects of a media provider real workflows such as production, distribution, content description, archiving, and re-use of video items. Issues like lack of human resources, necessity of parallel media distribution, and retrieving previously archived content through editors or consumers are discussed.

3. REFERENCES


