

Introduction to the special issue on multimedia intelligent services and technologies

Zhiwen Yu · Artur Lugmayr · Konstantinos Chorianopoulos · Tao Mei

Published online: 20 June 2010
© Springer-Verlag 2010

The rapid development of digital media technologies enables the emergence of novel media content types for eCommerce, eEducation, and digital entertainment. On the other hand, the advances in communication and micro-electronics have led to a transition from traditional personal computer-centric to more intuitive human-centric information access modes and the embedment of computer systems throughout the natural environment. This type of computation is generally known as *Pervasive or Ubiquitous Computing*. It can also be referred to as ambient intelligence, or talking about media environments as ambient media. It allows a person to use a variety of devices and sensor networks seamlessly embedded throughout our daily life, such as personal digital assistant (PDA), smart electronics, sensors, as well as personal computer, to access those media contents. The combination of these two trends (emergence of media and pervasive computing) holds the potential of providing a user with

seamless and ubiquitous access to rich and dynamic multimedia resources from anywhere and anytime.

With the emergence of ubiquitous and pervasive computation, distributed devices embedded in the natural human environment are getting more and more intelligent. This enables more advanced multimedia services far beyond simple video streaming and communication services. These services are currently on the brink of emergence, and include, e.g., personalized media environments, smart homes, semantic locative media, and context aware computer systems. There are several new technical challenges for effective and appropriate content delivery to pervasive user-terminals. First, different users may have different needs. Nowadays, multimedia content is absolutely overabundant, while the portion of which a user really desired is rather small. Smart personalization techniques, where content is aggregated, either delivered in push and/or pull schemes are required. Second, pervasive end devices usually have various capabilities (e.g., screen size, color depth, video/audio codec, memory, CPU speed, and electrical power), and network connection and transmission bandwidth might be changing dynamically. Therefore, intelligent multimedia services are expected for efficient storage, search, filtering, adaptation, and presentation of media content, to deliver a ubiquitous and personalized media experience to end users. Multimedia intelligent services and technologies have attracted increasing interest in both industry and academia over the last decade.

Submissions to this special issue come from an open call for papers as well as from selected papers presented at the 6th International Conference on Ubiquitous Intelligence and Computing (UIC-09) held at Brisbane, Australia, July 7–9, 2009. A large number of reviewers assisted us in the review process. In order to ensure high reviewing

Z. Yu (✉)
School of Computer Science,
Northwestern Polytechnical University, Xi'an, China
e-mail: zhiwenyu@nwpu.edu.cn

A. Lugmayr
EMMi Laboratory, Department of Business Information Systems
and Logistics, Tampere University of Technology,
Tampere, Finland
e-mail: lartur@acm.org

K. Chorianopoulos
Department of Informatics, Ionian University, Corfu, Greece
e-mail: choko@ionio.gr

T. Mei
Microsoft Research Asia, Beijing, China
e-mail: tmei@microsoft.com

standards, three to five reviewers evaluated each paper. After two rounds of rigorous reviews, we accepted nine papers in this special issue. Accordingly, those selected papers focus on multimedia intelligent services and technologies including three papers on multimedia recommendation, two papers on pervasive multimedia, and the other four on intelligent services of virtual reality, image analysis, 3D graphics retrieval, and video advertising.

Frank Hopfgartner and Joemon M. Jose investigate the problem of the explosion of news materials available through broadcast and other channels that precipitates a need for personalized news video retrieval. They introduce a semantic-based user modeling technique to capture users' evolving information needs. The approach exploits implicit user interaction to capture long-term user interests in a profile. The organized interests are used to retrieve and recommend news stories to the users. Specifically, the Linked Open Data Cloud is adopted to identify similar news stories that match the users' interest. Various recommendation parameters were tested using a simulation-based evaluation scheme.

In their paper "The effects of recommendations' presentation on persuasion and satisfaction in a movie recommender system", Theodora Nanou and George Lekakos argue that the selection of appropriate recommendation content and the presentation of information are important in creating successful recommender applications. They present a review of previous research approaches and popular recommender systems with focus on user persuasion and satisfaction. Experiments are conducted to compare different presentation methods in terms of recommendations' organization in a list (i.e., top N-items list and structured overview) and recommendation modality (i.e. simple text, combination of text and image, combination of text and video). Results showed that the most efficient presentation methods, regarding user persuasion and satisfaction, are the "structured overview" and the "text and video" interfaces, while a strong positive correlation was also found between user satisfaction and persuasion, in all experimental conditions.

To realize advanced TV personalization, e.g., TV content group recommendation and similar TV content retrieval, Zhiwen Yu et al. propose a hybrid approach for TV content similarity measure. It combines vector space model and category hierarchy model and leverages the advantages of both the methods. Similar TV contents are defined to be those with similar semantic information, e.g., plot, background, genre. The approach measures TV content similarity from the semantic level other than the physical level. Furthermore, they propose an adaptive strategy for setting the combination parameters. The experimental results showed that using the hybrid similarity measure proposed is superior to using either alone for TV content clustering and example-based retrieval.

Peizhao Hu et al. propose the MeshVision system that incorporates wireless mesh network functionality directly into the cameras. Video streams can be pulled from any camera within a network of MeshVision cameras, irrespective of how many hops away that camera is. To manage the trade-off between video stream quality and the number of video streams that could be concurrently accessed over the network, MeshVision uses a Bandwidth Adaptation Mechanism. This mechanism monitors the wireless network looking for drops in link quality or signs of congestion and adjusts the quality of existing video streams in order to reduce that congestion. A significant benefit of the approach is that it is low cost, requiring only a software upgrade of the cameras.

Jung-Hyun Kim et al. propose an advanced location awareness-based intelligent multi-agent technology that allows multiple users to share various user-centric mobile multimedia contents. Their main contributions include (1) presenting a mobile station-based mixed-web map module via mobile mash-up technology, (2) designing a new location-based mobile multimedia technology using ubiquitous sensor Net.-based five senses content, and (3) proposing a location awareness-based intelligent multi-agent technology that includes a location-based integrated retrieval agent, a Mobile Social Network-based multi-user detection agent, and a user-centric automatic mobile multimedia recommender agent.

In "Asynchronous reflections: theory and practice in the design of multimedia mirror systems", Wei Zhang et al. present a theoretical framing of the functions of a mirror by breaking the synchrony between the state of a reference object and its reflection. This framing provides a new conceptualization of the uses of reflections for various applications. They describe the fundamental technical components of such systems and illustrate the technical challenges in two different forms of electronic mirror systems for apparel shopping, the Responsive Mirror and the Countertop Responsive Mirror. The instantiations of the mirror systems in fitting room and jewelry shopping scenarios are described, focusing on the system architecture and the intelligent computer vision components.

Frode Eika Sandnes proposes a strategy for inferring approximate geographical information from the exposure information and temporal patterns of outdoor images in image collections. The basic idea is that image exposure is reliant on light and most photographs are therefore taken during daylight which again depends on the position of the sun. The sun results in different lighting conditions at different geographical location at different times of the day and hence the observed intensity patterns can be used to deduce the approximate location of the photographer at the time the photographs were taken. The approach proposed is efficient as it only considers meta information and not

image contents. Large databases can therefore be indexed efficiently. Experimental results demonstrated that the current approach yields a longitudinal error of 15.7° and a latitudinal error of 30.5° for authentic image collections comprising a mixture of outdoor and indoor images.

Zhenbao Liu et al. built an efficient 3D shape retrieval system based on shape alignment and shape orientation analysis. They propose to extract the spatial orientation of the polygon surfaces as the feature of one 3D shape. This information is analyzed by Multi-resolution Wavelet Analysis, and the low frequency components are applied to the feature vector. In the preprocessing stage, they adopt four methods of shape alignment including Principal Component Analysis, Continuous Principal Component Analysis, Normal-based Principal Component Analysis, and Plane Reflection Symmetry Analysis. In the orientation sampling stage, the sampling planes are placed on a cube and a dodecahedron, respectively. Finally, one shape descriptor based on Plane Reflection Symmetry Analysis and Dodecahedron sampling plane (PRSA-DOD) is proposed. The descriptor is selected to construct the search engine of system, and in the system the retrieval interface and search engine are implemented.

The last paper, “AdOn: toward contextual overlay in-video advertising”, by Tao Mei, Jinlian Guo, Xian-Sheng

Hua, and Falin Liu, introduces a contextual video advertising system which supports intelligent overlay in-video advertising. Unlike most current ad-networks that overlay the ads at fixed locations in the videos, the proposed system is able to automatically detect a set of spatio-temporal non-intrusive locations and associate the contextually relevant ads with these locations. The overlay ad locations are obtained on the basis of video structuring, face and text detection, as well as visual saliency analysis, while the ads are selected according to content-based multimodal relevance. The system represents one of the first attempts toward contextual overlay video advertising by leveraging information retrieval and multimedia content analysis techniques.

We believe that the selected papers make a significant contribution to the scientific community as well as practitioners or students working in the areas of multimedia intelligent services and technologies. We would like to thank Prof. Thomas Plagemann, the Editor-in-Chief, for his great support and effort throughout the whole publication process. We are also grateful to the referees for their professional and timely reviews. Finally, we would like to express our sincere appreciation to all the authors for their valuable contributions to this special issue, as without their contributions this special issue would not be possible.