Multimedia Data Management in Mobile Computing

Sergio Ilarri
University of Zaragoza, Spain

Florence Sèdes
Université Paul Sabatier, France

Francesco De Natale
University of Trento, Italy

Alan Hanjalic
Delft University of Technology, The Netherlands

The widespread availability of network-enabled portable multimedia devices able to capture, share, and access vast amounts of multimedia contents has led to an ever-increasing interest in new technologies providing improved multimedia data management in mobile environments. Some of the open challenges in this field include dealing with computational and storage limitations of mobile terminals, overcoming the significant limitations of batteries, achieving better bandwidth and latency in network access, enabling mobile access to large multimedia repositories, and providing more effective user interaction. At the same time, modeling metadata and semantics become more and more challenging. Most of these problems can be attacked from different viewpoints, for instance, by increasing the capacity of devices versus reducing the consumption of resources and energy, by increasing the efficiency of software tools versus strengthening hardware architectures, or by maximizing the network throughput versus representing the information more efficiently. Although the technology underlying these challenges has been studied for years within the broader multimedia community, mobility introduces some peculiar features that make the problem particularly challenging and still attractive for research.

This special issue focuses on the management and exploitation of multimedia data in mobile computing environments. The main goal was to collect articles reporting the latest advances in the technologies, algorithms, models, standards, and applications for managing and exploiting multimedia data in mobile computing. Specifically, we selected four representative articles that encompass topics such as image databases for mobile devices, mobile photo recommendations, logbook generation from context-tagged images, keyword-based multimedia retrieval, and querying multimedia data in vehicular networks.

Special Issue Articles

In David M. Chen and Bernd Girod's article "Memory-Efficient Image Databases for Mobile Visual Search," the focus is on mobile visual search (MVS) systems, which can identify objects in a user's surroundings from the images captured by the camera on the user's device, retrieve interesting information related to those objects, and show that information overlaid in the camera's viewfinder. In particular, the authors focus on detecting the relevant nearby objects by comparing the images obtained by the camera with a multimedia database that stores labeled images. An important novelty of their work is that, instead of relying on a remote server where the image database is stored, the mobile device itself hosts a compact database of image signatures that summarize statistics of the image features. This approach improves the user experience by minimizing wireless communications, which leads to a low latency. To realize this vision, suitable methods to build the compact image database and an efficient database update mechanism from a remote server are necessary. In particular, Chen and Girod present four methods for building the compact databases: tree histogram coding (THC), inverted index coding (IIC), residual enhanced visual vector (REVV), and scalable compressed fisher vector (SCFV). In addition to an experimental evaluation of these methods, they also present the results of a REVV-based

1070-986X/14/$31.00 © 2014 IEEE

Published by the IEEE Computer Society

implementation of an MVS system on an Android smartphone that recognizes landmarks and media covers.

"Mobile Photo Recommendation and Logbook Generation Using Context-Tagged Images" by Windson Viana, Reinaldo Braga, Fabrício D.A. Lemos, João M.O. de Souza, Rafael A.F. Carmo, Rossana M.C. Andrade, and Hervé Martin focuses on context-tagged images. Specifically, the article presents two mobile applications (MMedia2U and CAPTAIN) based on the Context-Aware Mobile and Multimedia Architecture (CoMMediA) framework, which advocates the joint use of context-awareness and semantic technologies to improve the management of mobile multimedia. CAPTAIN (a context-aware system based on personal tracking) is a tool evaluated in the context of a sea expedition (ZeroCO2). It includes both a mobile iOS-based application and a desktop application. The mobile application captures the information about the user's context (such as location and speed), which is used to automatically tag multimedia content. The desktop application stores the information provided by the mobile device until there is an Internet connection available that allows the data to be transferred to a Web server. Moreover, it can enhance the context information by querying remote servers if there is a connection available. Finally, the annotated multimedia resources can be used to build structured multimedia documents, such as Web logbooks. MMedia2U (Mobile Multimedia to You), on the other hand, is a photo recommender system that exploits a collection of context-labeled images to provide the user with photos created in similar contexts.

The article "A Multimedia Semantic Retrieval Mobile System Based on HCFGs" by Yimin Yang, Hsin-Yu Ha, Fausto C. Fleites, and Shu-Ching Chen tackles the problem of keyword-based multimedia retrieval in mobile applications. In particular, it presents a multimedia semantic retrieval system based on the use of hidden coherent feature groups (HCFGs). The authors advocate exploiting the correlation between features, integrating multiple decision models by using a suitable scheme to fuse the selected scores from multiple models, and incorporating a user feedback mechanism to enhance the semantic retrieval performance. The authors have developed an iPad application and provide an experimental evaluation based on a dataset of disaster images to show the system's effectiveness.

This special issue inspires a number of possible directions for future work in this challenging and cross-disciplinary domain.

Ouri Wolfson and Bo Xu close this special issue with their article "Querying Blobs in Vehicular Networks," which focuses on query processing strategies for multimedia data in vehicular networks. Vehicular ad hoc networks (Vanets)[1] represent a special mobile computing environment where vehicles communicate with each other using short-range wireless communication technologies, such as WiFi or dedicated short-range communications (DSRC). In this context, the exchange of multimedia information (voice, images, or video clips) among vehicles can enable situation awareness or urban monitoring. The idea is to enable the processing of queries about multimedia resources received from other vehicles. For example, drivers may wish to obtain multimedia information about potential road hazards ahead or the upcoming traffic conditions. The authors define different query processing strategies based on three dimensions: push (where blobs are proactively disseminated) versus pull (where queries are disseminated and blobs are disseminated as an answer to received queries), WiFi-based ad hoc communications versus the use of a cellular communication infrastructure along with WiFi, and metadata dissemination separated from blob dissemination versus metadata and blob dissemination (a multimedia report can be matched with a query based only on its metadata). An experimental evaluation of strategies using the SWANS++ simulator is provided for the case of range queries that retrieve multimedia reports captured within a certain spatial area.

Future Directions

Although the present issue provides only a glimpse of this huge research area, it inspires a number of possible directions for future work in this challenging and cross-disciplinary domain. A noncomprehensive list of open problems in

Progress in this area could enable a range of innovative applications and services in a number of different user scenarios.

the area of multimedia data management in mobile computing, characterized by structure heterogeneity and semantic richness, would include the following:

- fitting a rich multimedia application into a small-screen device with limited interactivity;
- dealing with a user with partial attention or interaction (such as walking or driving a car);
- adapting the huge amount of data available on the Internet to the severely limited capabilities of a mobile device;
- managing the variability in location or proximity to places or other users;
- seamlessly handling the unpredictable variation of network coverage;
- enhancing the management of multimedia, user profiling, and personalization by embedding semantic technologies (such as ontologies[2] and reasoners[3]) on mobile devices with constrained capabilities; and
- smartly connecting the portable device to different sensors and actuators to create a real multimodal portal to the surrounding world.

Furthermore, progress in this area could enable a range of innovative applications and services in a number of different user scenarios, such as

- *event capturing and sharing* (such as in scenarios of participatory sensing), where a user can exploit his/her handheld to spot an interesting situation and distribute it in real time (such as for urban monitoring, surveillance, or news production);
- *vehicular networks*, where a car equipped with cameras may transmit multimedia data to other drivers or to a remote center, sharing traffic information, driving alerts, or other interesting information;
- *TV broadcasting/multicasting*, where mobile cameras could be managed and any mobile user with a smartphone could potentially become a source of multimedia data for a newspaper or news agency if he or she is located in an area where something interesting is happening;
- *emergency management*, where multimedia data can be captured and transmitted by different types of devices to help the emergency services prepare to optimally assist people;
- *mobile social networks*, where users can exchange multimedia information among friends or people located in a certain area, possibly georeferenced or enriched with geospatial data related to maps and/or points of interest (POIs);
- *augmented or enriched reality*, where users can benefit from external information sources to add semantics to the world around them, for instance, by improving their experience in a museum or learning how to operate a device; and
- *elderly or disabled support*, where a mobile portable device can become an enabling technology to improve the safety of its user as a guide, to provide constant health monitoring, and to connect users to medical assistance.

Technological breakthroughs are fostered in all such fields. Such research still requires a large, collaborative effort in different scientific and technologic disciplines. We hope that this special issue will help to encourage further joint research in these areas. **MM**

References

1. H. Hartenstein and K. Laberteaux eds., *Vehicular Applications and Inter-Networking Technologies,* John Wiley & Sons, 2010.
2. M. Uschold and M. Gruninger, "Ontologies and Semantics for Seamless Connectivity," *SIGMOD Record,* vol. 33, no. 4, 2004, pp. 58–64.
3. R.B. Mishra and S. Kumar, "Semantic Web Reasoners and Languages," *Artificial Intelligence Rev.,* vol. 35, no. 4, 2011, pp. 339–368.

Sergio Ilarri is an associate professor in the Department of Computer Science and Systems Engineering at the University of Zaragoza, Spain. His research

interests include data management issues for mobile computing, vehicular networks, mobile agents, information systems, and the Semantic Web. Sergio has a PhD in computer science engineering from the University of Zaragoza, Spain. Contact him at silarri@unizar.es.

Florence Sèdes is a professor at the University of Toulouse 3, Paul Sabatier University, France, deputy director of IRIT (Institut de Recherche en Informatique de Toulouse), and head of the national research network GDR i3 (Information Interacting Intelligence) of the French CNRS. Her research interests include information systems, multimedia data management, and pervasive computing. Florence has a PhD in computer science from the University of Toulouse, France. Contact her at sedes@irit.fr.

Francesco De Natale is a professor at the University of Trento, Italy, a member of the Management Board of CNIT, and the representative of UniTN within the EIT ICT Labs. His research interests include image and signal processing, with particular attention to multimedia data processing, analysis, and retrieval. Francesco has a PhD in telecommunication engineering from the University of Genova, Italy. Contact him at denatale@ing.unitn.it.

Alan Hanjalic is a professor at the Delft University of Technology, The Netherlands, and the head of the Multimedia Signal Processing Group. His research interests include multimedia signal processing, with focus on multimedia information retrieval and recommender systems. Hanjalic has a PhD in computer science from the Delft University of Technology, The Netherlands. Contact him at A.Hanjalic@tudelft.nl.

cn *Selected CS articles and columns are also available for free at http://ComputingNow.computer.org.*