

## Interactive multimedia computing

Meng Wang · Jinhui Tang · Xian-Sheng Hua ·  
Tat-Seng Chua

Published online: 7 December 2010  
© Springer-Verlag 2010

In recent years, we have witnessed the flourish of multimedia data on the Internet. To facilitate humans in accessing and managing the explosively growing multimedia contents, extensive research efforts have been dedicated to automatic multimedia analysis and processing in the past decades, such as categorization, annotation and indexing. However, despite great advances achieved, several key difficulties still exist, such as the well-known semantic gap in multimedia modeling. It is evident from recent results that, without additional information resources, most of the semantic gap problems can hardly be solved automatically within the near future. On the other hand, we have witnessed the power of collective human efforts in the Web 2.0 era in providing high-quality tags and comments to large amounts of images and videos in sites such as Flickr and YouTube. In fact, a lot more can be accomplished through simple online games such as the ESP. Hence, more and more researchers believe that a possible approach to addressing the semantic gap problem is to incorporate the efforts of humans into the computational process, i.e., by combining human intelligence and automated computer processing to jointly tackle the problems in a collaborative manner. The past decade has witnessed the increase of such efforts, such as relevance feedback in content-based image retrieval, active learning in multimedia modeling, the interactive video search evaluation task in TRECVID, new search and browsing interfaces in VideoOlympics to facilitate humans' interaction, and the

recent human computation efforts such as the ESP game on Google image search website.

This special issue is organized with the purpose of introducing novel research work on interactive multimedia computing. Submissions have come from an open call for paper. With the assistance of dedicated referees, five papers have been selected after two rounds of rigorous reviews. These papers cover widely subtopics of interactive multimedia computing, including game-based image annotation, interactive TV, interactive cartoon synthesis, and so on.

In the first paper “Adding Semantics to Image Region Annotations with the Name-It-Game”, Steggink and Snoek introduce a system that accomplishes region-level image annotation with a game. It establishes a set of keywords that describe objects by exploring WordNet, and the keywords are assigned to image regions with a two-player “reveal and guess” game. They also explore WordNet to address the word ambiguity problem. In addition to introducing the system, another contribution of the paper is its review of existing manual image annotation techniques, in particular the comprehensive study of game-based annotation.

In the second paper “Interactive Browsing via Diversified Visual Summarization for Image Search Results”, Wang et al. introduce a scheme for the summarization and browsing of image search results. It adopts a dynamic absorbing random walk approach to summarize the image search results. The summarization is visualized on a 2D panel and users' browsing is facilitated with dynamic scale change and a browsing path tracking tool. Experiments with a set of diverse queries have demonstrated the effectiveness of the approach.

The third paper, “Security and Privacy Requirements in Interactive TV”, discusses the security and privacy issues in the context of interactive TV. It introduces an interactive

---

M. Wang (✉) · J. Tang · T.-S. Chua  
National University of Singapore, Singapore, Singapore  
e-mail: eric.mengwang@gmail.com

X.-S. Hua  
Microsoft Research Asia, Beijing, China

TV architecture for regulating all kinds of interactions and data streams, including all parties with different interests. It densely discusses the privacy concerns, economic interests, legal necessities and approaches, and provides a set of suggestions.

In the fourth paper, “Semi-Automatic Cartoon Generation by Motion Planning”, Yu et al. introduce an interactive cartoon generation approach. Cartoonists manually perform two tasks, frame initialization and path drawing, and then a set of frame insertion points are established. Starting from the initially established frame, the frames at the insertion points are then selected one by one via performing a graph-based learning algorithm, whereby a cartoon clip can be generated.

The final part of the special issue, “Personalized Video Similarity Measure”, introduces a novel video indexing approach with the purpose of facilitating personalized video search. It presents a new distance metric called

personalized video distance. It integrates personalized information and multimodal video properties into a probability based signature, and a new hashing based indexing structure is designed to support fast search. Experiments have demonstrated that the approach can lead to superior search performance and efficiency.

To conclude, the papers in this special issue cover different multimedia computing techniques that jointly explore human and computers. We believe this special issue will benefit researchers and practitioners working in this area.

**Acknowledgments** We would like to thank Prof. Thomas Plagemann for providing us the chance to organize this special issue. We thank the reviewers for their great efforts. Their professional evaluations and constructive comments are vital to securing the high quality of the special issue. Finally, we express our thanks to all the authors who have contributed to this special issue.