

## Editorial of the special issue on cross-media analysis

Zhongfei Zhang · Yueting Zhuang · Ramesh Jain ·  
Jia-Yu Pan

Published online: 30 June 2014  
© Springer-Verlag London 2014

Today, there are lots of heterogeneous and homogeneous media data from multiple sources, such as news media websites, microblog, mobile phone, social networking websites, and photo/video sharing websites. Integrated together, these media data represent different aspects of the real-world and help document the evolution of the world. Consequently, it is impossible to correctly conceive and to appropriately understand the world without exploiting the data available on these different sources of rich multimedia content simultaneously and synergistically. Cross-media analysis is a research area in the general field of multimedia content analysis that focuses on the exploitation of the data with different modalities from multiple sources simultaneously and synergistically

---

This work is supported in part by US NSF through grants (IIS-0812114 and CCF-1017828). ZZ is also supported by the National Basic Research Program of China (2012CB316400), Zhejiang University—Alibaba Financial Joint Lab, and Zhejiang Provincial Engineering Center on Media Data Cloud Processing and Analysis.

---

Z. Zhang (✉)  
Computer Science Department, Watson School of Engineering  
and Applied Science, State University of New York,  
Binghamton, NY 13902-6000, USA  
e-mail: zhongfei@cs.binghamton.edu

Y. Zhuang  
College of Computer Science, Zhejiang University,  
Hangzhou 310027, P.R. China  
e-mail: yzhuang@cs.zju.edu.cn

R. Jain  
Department of Computer Science, University of California,  
Irvine, CA 92697-3425, USA  
e-mail: jain@ics.uci.edu

J.-Y. Pan  
Google, Mountain View, CA 94041, USA  
e-mail: jypan@google.com

to discover knowledge and understand the world. Specifically, we emphasize two essential elements in the study of cross-media analysis that help differentiate cross-media analysis from the rest of the research in multimedia content analysis or machine learning. The first is the simultaneous co-existence of data from two or more different data sources. This element indicates the concept of "cross", e.g., cross-modality, cross-source, and cross-cyberspace to reality. Cross-modality means that heterogeneous features are obtained from the data in different modalities; cross-source means that the data may be obtained across multiple sources (domains or collections); cross-space means that the virtual world (i.e., cyberspace) and the real world (i.e., reality) complement each other. The second is the leverage of different types of data across multiple sources for strengthening the knowledge discovery, for example, discovering the (latent) correlation or synergy between the data with different modalities across multiple sources, transferring the knowledge learned from one domain (e.g., a modality or a space) to generate knowledge in another related domain, and generating a summary with the data from multiple sources. These two essential elements help promote cross-media analysis as a new, emerging, and important research area in today's multimedia research. With the emphasis on knowledge discovery, cross-media analysis is different from the traditional research areas, such as cross-lingual translation. On the other hand, with the general scenarios of the leverage of different types of data across multiple sources for strengthening the knowledge discovery, cross-media analysis addresses a broader series of problems than the traditional research areas, such as transfer learning. Overall, cross-media analysis is beneficial for many applications in data mining, causal inference, machine learning, multimedia, and public security.

After a rigorous, two rounds of peer review, we have accepted five papers into this special issue, representing five different research aspects or different applications of cross-media analysis.

The paper titled *Interactive Cross and Multimodal Biomedical Image Retrieval Based on Automatic Region-Of-Interest (ROI) Identification and Classification* by Rahman, You, Simpson, Antani, Fushman, and Thoma studies the problem of cross-modality biomedical image retrieval from articles that contain text and images. Specifically, visual features of the detected ROIs (regions of interest) of a biomedical image are mapped to text concepts for image retrieval. The retrieval can be facilitated from a pure perceptual search to a conceptual search based on the cross-modality analysis. Evaluations are reported using a biomedical article dataset of thoracic CT scans.

The paper titled *Optimization of Information Retrieval for Cross Media contents in a Best Practice Network* by Cenni, Nesi, and Bellini addresses the cross-media search problem in the application of performing arts. Specifically, the paper proposes a search prototype system for European Collected Library of Artistic Performance (ECLAP) performing art social network data. Technical issues, such as scalability, robustness, and error resistance due to the presence of the large-scale heterogeneous data across different modalities are appropriately addressed in the prototype system.

The paper titled *Person Instance Graphs for Mono-, Cross-, and Multi-Modal Person Recognition in Multimedia Data* by Bredin, Roy, Le, and Barras focuses on the application of speaker identification in TV broadcast. An effective solution is developed using the person instance graph, which reduces the person identification problem to the graph mining

problem to find the best mapping between person instance vertices and identify vertices, resulting in the development of a unified framework for mono-, cross-, and multi-model person identification in multimedia data.

The paper titled *MET: Media Embedded Target for Connecting Paper to Digital Media* by Liu, Girgensohn, Wilcox, Shipman, and Dunnigan addresses a specific industrial application of cross-media analysis of recognizing a media embedded target (MET) and then navigating through the link of the MET. A MET is an iconic mark printed in a blank margin of a page with a media link to a nearby region of the page. Though this is a relatively specific application of cross-media analysis, the paper demonstrates a nice case study of the wide spectrum of the techniques available for cross-media analysis.

The paper titled *Topic Detection in Cross-Media: A Semi-Supervised Co-clustering Approach* by Xue, Li, Zhang, Pang, and Huang studies the problem of social media topic detection. With the presence of multiple modalities of data in typical social media, this work showcases the methodology research in cross-media analysis. Specifically, the paper proposes an effective solution to the topic detection problem in the context of cross-media analysis as a semi-supervised co-clustering approach using constrained non-negative matrix factorization.

We hope that this special issue shall help bring the attention of the researchers and practitioners in the related communities to the emerging area of cross-media analysis and shall lead to the further development of the research as well as the applications of the technologies from the research in this area to benefit the whole society.