Introduction to Special Issue on Social Media

With the surging popularity of the Internet, increasingly rich and massive social media data such as texts, images, audios, videos, and blogs are being posted to the Web. We find social media in all kinds of applications like social networks, photo and video sharing, and photo forums. In such environments, the media content is typically accompanied by rich contextual information. Content descriptions, time and location, social connections, comments, and viewer ratings accompany the media data. This user-generated information can become particularly critical to resolving the long-standing challenge facing multimedia understanding. This new media also introduces many new and challenging research issues, as well as many exciting real-world applications, such as large-scale social image and video analysis and retrieval.

Following an exciting series of ACM International Workshops on Social Media, we initiated this special issue to provide a platform for presenting results from leading edge research in understanding and exploring social media. The goals of this special issue are threefold in presenting the next-generation technologies for

—organizing, indexing, retrieving and mining such rich social media contents;
—employing rich social media contents for multimedia content analysis, computer vision and graphics, and
—enabling and empowering peoples in social networks.

This special issue received 19 high-quality submissions from authors worldwide. The articles underwent a competitive review process with at least three reviews per article. We selected eight articles, which demonstrate the exciting and wide-ranging research area of social media.

Selected Articles

We would like to introduce the articles of this special issue and their perspectives on social media and give you an insight into the great work in this field that is being pursued by the authors.

One means to efficiently and effectively access music from the millions of songs available from music stores is emotion classification. In the article "Exploiting Online Music Tags for Music Emotion Classification" by Yu-Ching Lin, Yi-Hsuan Yang and Homer H. Chen, the authors propose an approach in which they obtain the genre tag of the song by using a Web crawler and then apply the corresponding genre-specific classifier to recognize the emotion of the song.

Social media sites are places where many personal photos are uploaded, but there is no easy way to collect and bundle them into a story. The article "Automatic Creation of Photo Books from Stories in Social Media" by Mohamad Rabbath, Philipp Sandhaus and Susanne Boll proposes an approach to automatically detect media elements that match a query (where, when, what, who) in a user's social network and intelligently arrange and compose them into a printable photo book.

With the explosive growth of the World Wide Web (WWW), efforts are needed to keep children from browsing offensive and obscene content. In the article "Recognition of Adult Images, Videos and Webpage Bags," the authors Weiming Hu, Haiqiang Zuo, Ou Wu, Yunfei Chen, Zhongfei Zhang and David

DOI 10.1145/2037676.2037682 http://doi.acm.org/10.1145/2037676.2037682

Suter develop a system to recognize Web adult images, adult videos, and adult Web page bags. The authors first propose to model the skin patches for adult-image recognition, and then additionally use the information from audio for adult-video recognition. Moreover, they also adopt a few multi-instance learning methods for adult Web page bags recognition.

It is important to effectively detect key changes in the rapidly evolving social networks. The article "SCENT: Scalable Compressed Monitoring of Evolving Multirelational Social Networks," by Yu-Ru Lin, K. Selçuk Candan, Hari Sundaram and Lexing Xie describes a scalable system for extracting useful information from social media data. In SCENT, compressed sensing is extended to tensor streams to encode compact descriptors that allow rapid detection of significant spectral changes in the data with reduced data collection, storage, and processing costs.

Given the overwhelming number of Web videos returned by search engines, it becomes necessary to organize the retrieved videos into semantically meaningful clusters for easy browsing. In the article "Browse by Chunks: Topic Mining and Organizing on Web-Scale Social Media," Jitao Sang and Changsheng Xu propose using a hierarchical topic structure and a semi-supervised hierarchicals topic model to help discover the topic hierarchy. Along with an intuitive visualization user interface, their system enables users to effectively browse the video search results.

Discovering touristic landmarks from geo-tagged photos in photo sharing sites can help us to make better sense of our visual world. The article by Rongrong Ji, Yue Gao, Bineng Zhong, Hongxun Yao, and Qi Tian entitled "Mining Flickr Landmarks by Modeling Reconstruction Sparsity" develops a system for city scene summarization by mining landmarks from geo-tagged Flickr photos. The authors first applied the spectral clustering method to partition each city into geographical regions, and then employed the sparse representation algorithm to mine the representative photos from each landmark region, which is further refined by incorporating the community knowledge.

Social tags and expert-generated tag descriptions can be helpful but their creation is time consuming and expensive. The article "Contextual Tag Inference," by the authors Michael I. Mandel, Razvan Pascanu, Douglas Eck, Yoshua Bengio, Luca M. Aiello, Rossano Schifanella, Filippo Menczer proposes a content-based automatic music tagging scheme by employing the relationships between tags and the clips of songs. The conditional restricted Boltzmann machine models and support vector machines were used to better incorporate context information for improving autotagging.

Video sharing sites host a large number of conversational video blogs in a single talking-head format. In "VlogSense: Conversational Behavior and Social Attention in YouTube," Joan-Isaac Biel and Daniel Gatica-Perez look beyond the concrete topic and content of the video to study the characteristics of vloggers' audiovisual nonverbal behavior by using multimodality analysis techniques. Their analyses of 2,269 vlogs from YouTube discovered interesting insights on common practices of vloggers.

The editors of this special issue hope that you will find these articles interesting!

Susanne Boll, Ramesh Jain, Jiebo Luo, and Dong Xu
Guest Editors