CrossMark

EDITORIAL

Recent advances in social multimedia big data mining and applications

Jitao Sang[1] · Yue Gao[2] · Bing-kun Bao[1] · Cees Snoek[3] · Qionghai Dai[4]

Abstract In the past decade, social media contributes significantly to the arrival of the Big Data era. Big Data has not only provided new solutions for social media mining and applications, but brought about a paradigm shift to many fields of data analytics. This special issue solicits recent related attempts in the multimedia community. We believe that the enclosed papers in this special issue provide a unique opportunity for multidisciplinary works connecting both the social media and big data contexts to multimedia computing.

1 Introduction

Recent years have witnessed the popularity of social media websites, such as Flickr, YouTube, Twitter and Facebook. Social media and Web2.0 applications have transformed the web into an interactive sharing platform where huge amounts of unstructured data are generated every minute. Nowadays, social media and big data converge to describe what happens in the world on content production, collection and analysis. Conducting big data analysis has a significant role in maximizing the utility of social media. The key issues in exploiting and fostering the marriage between "social media" and "big data" are the process is large-scale, weakly labeled, heterogeneous, unstructured, multi-source and redundant. In this special issue, we are delighted to select twelve papers which attempt to address the above-mentioned issues and develop solutions for multimedia-related problems.

The first paper "Chat with illustration" by Jiang et al. addresses a novel multimedia application by exploiting the huge available user-generated social media data. The authors propose to enrich the traditional text-based instant messaging service with visual content. A novel visual-assisted scheme called Chat With Illustration (CWI) is introduced, with goal to present users a new chatting experience augmented with relevant images. With combination of multi-modal messages, users are validated to obtain a more interesting and vivid communication experience. Another advantage of CWI is its language-free characteristic, which means for different native language speakers, it can help them cross language barrier to some degree.

In the paper titled "SocialRobot: a big data-driven humanoid intelligent system in social media services", Liu et al. proposed a humanoid intelligent system named SocialRobot in Sina Weibo, which is the most famous social media platform in China. SocialRobot is the first robot "living" in the social media services. It can exchange, recommend and gather domain specific information for its followers as well as chatting with them. Hence, it can be seen as a quasi-Turing test, which is an ultimate goal of the Artificial Intelligence.

✉ Jitao Sang
  jtsang@nlpr.ia.ac.cn

  Yue Gao
  kevin.gaoy@gmail.com

  Bing-kun Bao
  bingkunbao@gmail.com

  Cees Snoek
  cgmsnoek@uva.nl

  Qionghai Dai
  qhdai@tsinghua.edu.cn

[1] Chinese Academy of Sciences, Beijing, China

[2] University of North Carolina, Chapel Hill, USA

[3] University of Amsterdam, Amsterdam, Netherland

[4] Tsinghua University, Beijing, China

With the aim of handling large-scale video data, the third paper titled "Efficient video copy detection using multimodality and dynamic path search" by Li et al. proposes an efficient and robust video copy detection method jointly utilizing the characteristics of temporal continuity and multimodality of video. The video is converted to a continuous sequence of states, and both the visual and auditory features are extracted for temporal frames. To facilitate tolerance of the length variations caused during video re-targeting, an efficient dynamic path search method is proposed to detect the target video clips, and highly compact audio fingerprint and visual ordinal features are jointly utilized in a flexible frame. The proposed scheme not only achieves high computational efficiency but also guarantees effectiveness in real applications.

The next paper, titled "Semi-supervised feature selection via hierarchical regression for web image classification" by Song et al., proposes a semi-supervised feature selection algorithm for large-scale image data analysis, called Feature Selection via Hierarchical Regression (FSHR). Different from most of the existing semi-supervised feature selection algorithms, the proposed method preserves the manifold structure of each feature type during the training phase. Local classifiers and global classifiers are learnt simultaneously. Once the optimal global classifiers are obtained, the features are sort accordingly in descending order and the top ranked ones are selected. The selected discriminative features are proved to be effective to classify web images, when combined with a multi-class SVM classifier in the experiments. This method takes advantages of both semi-supervised learning and $l_{2,1}$-norm-based sparsity.

To utilize the weakly labeled user-generated social media data, the paper titled "Enriching one-class collaborative filtering with content information from social media" by Yuan et al. investigates into One-Class Collaborative Filtering (OCCF) problem in recommendation applications. Different from existing attempts which discover the negative examples solely based on the statistical properties of users' historical behavior, the authors propose to incorporate rich content information as additional evidence for modeling. For methodology, conventional Matrix Factorization model is modified with probabilistic topic modeling to obtain content topic feature for both users and items.

With more and more textual and image information available on social networks, logo information retrieval from these platforms become a practically important research topic. In the paper titled "Logo information recognition in large-scale social media data", Wang et al. proposed a new logo detection method. A new dense histogram-based visual feature was proposed to detect logo in images, with textual information helping filter out irrelevant candidate, it is demonstrated with promising performance on real microblog data.

The seventh paper titled "Cross-domain semantic transfer from large-scale social media" by Nie et al. addresses the heterogeneous issue in exploiting social media big data. In the context of Location-based Social Networks (LBSN), the authors introduce an efficient cross-domain image annotation scheme to discover the semantic correlation between texts and images. The results lead to a set of specific textual description for the user-generated images. The discovered multimodal correlation also enables the management of textual and visual information around specific venues, and improves the user experience in location-based services.

Also addressing the weakly labeled characteristic of social media data, the paper titled "Characterizing users' check-in activities using their scores in a location-based social network" by Jin et al. focuses on users' score in the check-in activities which shows the accumulative points the user is awarded from all of his/her mobility activities in Foursquare. The authors model a user's consecutive scores and extract properties to characterize the mobility patterns using Non-negative Matrix Factorization. The experimental results demonstrate how to uncover interesting spatiotemporal patterns by utilizing the aggregate measures released by an LBSN service.

To facilitate the automatic organization and management of the increasingly unstructured web video comment, the ninth paper, titled "Accurate online video tagging via probabilistic hybrid modeling" by Shen et al., proposes a novel video tagging framework called comprehensive video tagger (CVTagger). A hierarchical structure with multi-layer concept model is embedded in the framework to classify the web videos considering both multimodal information and temporal dynamics.

To manage the unstructured social multimedia data arriving in stream, the paper titled "Online web video topic detection and tracking with semi-supervised learning" by Li et al. investigates the problem of web video Topic Detection and Tracking (TDT) in an online fashion. The historical data are leveraged to learn offline topic detector which is then used to classify the online new video. The authors claim that the proposed solution is capable of tracking the topics of interest, predicting the topic evolution and discovering the emerging trends in a timely way.

The next paper, titled "BHoG: binary descriptor for sketch-based image retrieval" by Fu et al., addresses the application of sketch-based image retrieval on mobile devices. Computation complexity and memory space are critical issues in mobile-based applications. To alleviate the computation and memory costs in processing multimedia resources, the authors introduce a discriminative and efficient BHoG feature as signatures for the hand-drawn sketches. The proposed solution framework is validated on a large-scale social media dataset with promising

experimental results not only in flexibility and efficiency, but with less memory occupation.

In the last paper titled "Decomposed human localization from social photo album", Li et al. propose a decomposition-based human localization model dealing with localizing highly deformable persons, as commonly exhibited in personal photo albums, in which the existing methods cannot do well. The model first detect upper body, and adopt an appearance model-based approach to score expanding instances. Then to achieve similar pose retrieval in album browsing application, Similar Pose Feature (SPF) is proposed to achieve this task, which records the relative offset between geometric center of upper body and whole person.