conferenceseries.com SciTechnol

J Comput Eng Inf Technol 2017, 6:4(Suppl)
DOI: 10.4172/2324-9307-C1-012

Multimedia knowledge extraction: Get things right about complex information

Heng Ji
Rensselaer Polytechnic Institute, USA

Knowledge extraction and representation have been the common goals for both the text domain and the visual domain. A few significant benchmarking efforts, such as TREC and TRECVID, have also demonstrated important progress in information extraction from data of different modalities. However, none of the media modality research is complete and fully reliable. Systems using text Knowledge Base Population (KBP) tools cover important high-level events, entities, and relations, but they often do not provide the complete details depicting the physical scenes, objects, or activities. Visual recognition systems, despite the recent progress, still suffer from inadequate abilities in extracting high-level semantics comparable to the counterparts from the text part. In this talk, we will present our recent efforts at developing a scalable, portable, and adaptive multi-media knowledge construction framework which can exploit cross-media knowledge, resource transfer and bootstrapping to dramatically scale up cross-media knowledge extraction processes. We have developed novel cross-media methods (including a cross-media deep learning model and "Liberal" KBP) to automatically construct multimodal semantic schema for event, improve extraction through inference and conditional detection, and enrich knowledge through cross-media cross-lingual event co-reference and linking.

jih@rpi.edu

Collaborative active learning from crowds for visual recognition

Chengjiang Long
Kitware Inc., USA

Active learning is an effective way to relieve the tedious work of manual annotation in many applications of visual recognition. The vast majority of previous works, if not all of them, focused on active learning with a single human oracle. The problem of active learning with multiple oracles in a collaborative setting has not been well explored. Moreover, most of the previous works assume that the labels provided by the human oracles are noise free, which may often be violated in reality. To solve the above-mentioned issues, we proposed two models (i.e., a distributed multi-labeler active learning model and a centralized multi-labeler active learning model) for collaborative active visual recognition from the crowds, where we explore how we can effectively model the labelers' expertise in a crowdsourcing labeling system to build better visual recognition models. Both two models are not only robust to label noises, but also a principled label quality measure to online detect irresponsible labelers. We also extended the centralized multi-labeler active learning model from binary cases to multi-class cases and also incorporate the idea of reinforcement learning to actively select both the informative samples and the high-quality annotators, which better explores the trade-off between exploitation and exploration. Our collaborative active learning models have been validated in the real-world visual recognition benchmark datasets. The experimental results strongly show the validity and efficiency of the two proposed models.

clong@stevens.edu