

Panel: Cross-media Intelligence

Organizer & Chair:
Yueting Zhuang
Zhejiang University
yzhuang@cs.zju.edu.cn

Panelist:
Ramesh Jain
University of California, Irvine
jain@ics.uci.edu

Panelist:
Wen Gao
Peking University
wgao@pku.edu.cn

Panelist:
Liu Ren
Robert Bosch Research
williamren@gmail.com

Panelist:
Kiyoharu Aizawa
University of Tokyo
aizawa@hal.t.u-tokyo.ac.jp

Panel Summary or Overview

In this panel, we attempt to review and discuss the recent emerging theoretical and technological advances and trends of cross-media. Integrating data-driven machine learning with human knowledge can effectively lead to explainable, robust, and general models. Thus, the effective employment of the interaction between cross-media data during inference and reasoning becomes a challenge to populate the cross-media knowledge graph. Some other fundamental and controversial issues such as leveraging the auxiliary information to boost the cross-media understanding, the existence of unified framework to bridge the gap between multi-modality will also be discussed in this panel.

CCS Concepts/ACM Classifiers

Machine Learning

Keywords: Cross-media Intelligence; knowledge population; AI creativity

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 intelligence 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 to discover knowledge and understand the world.

In this panel, we attempt to discuss some of the emerging trends in terms of cross-media intelligence depicted as follows:

The effective integration of rule-based symbolic reasoning and data-driven learning (i.e., connectionist learning). An appropriate integration is desirable for enhancing the ability to explain intelligent actions, e.g., prediction and classification.

Cross-media knowledge population. Data with multi-modalities from multi-domains can improve the robustness and reliability of inference and reasoning. In general, data with different modalities have different discriminative powers to encode their particular semantics. As a result, how to employ the intrinsic interaction between cross-media during inference and reasoning is a fundamental challenge to populate the cross-media knowledge graph.

Creative ability via artificial intelligence. In recent years, we have witnessed an explosion of AI generated arts such as pop music and painted photos. Whether next generation AI will have creative ability in some domain-specific field is an amazing research direction.

Like other emerging hot topics in multimedia community, cross-media intelligence also has a number of fundamental and controversial issues that must be addressed in order to have a full and complete understanding of the research in this topic. These issues include but are not limited to whether or not there exists a unified framework to bridge the gap between natural language and multimedia, and if there is what such unified framework is; whether or not there exists any computational method that effectively leverages the auxiliary information (i.e., common knowledge and priors) to boost cross-media semantic understanding and if there is what such "method" is and how it is formulated; whether or not there exists a way which can result in breaking the "black box" of deep learning (the dominant method in multimedia community) and the way can perform deep neural reasoning beyond mere classification or detection.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

MM'17, October 23-27, 2017, Mountain View, CA, USA.

© 2017 Copyright is held by the owner/author(s).

ACM ISBN 978-1-4503-4906-2/17/10.

<https://doi.org/10.1145/3123266.3133336>