ACL 2018

First Grand Challenge and Workshop on Human Multimodal Language (Challenge-HML)

Proceedings of the Workshop

July 20, 2018
Melbourne, Australia

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)
209 N. Eighth Street
Stroudsburg, PA 18360
USA
Tel: +1-570-476-8006
Fax: +1-570-476-0860
acl@aclweb.org

ISBN 978-1-948087-46-9

Introduction

Welcome to the First Grand Challenge and Workshop on Human Multimodal Language (Challenge-HML). This grand challenge is co-located with ACL 2018 in Melbourne, Australia. During this grand challenge, we aim to gauge the performance of current natural language processing models in understanding the complete form of human language: from language, vision and acoustic modalities all used in a coordinated manner to convey intentions.

Computational analysis of human multimodal language is an emerging research area in Natural Language Processing (NLP). It expands the horizons of NLP to study language used in face to face communication and in online multimedia. This form of language contains modalities of language (in terms of spoken text), visual (in terms of gestures and facial expressions) and acoustic (in terms of changes in the voice tone). At its core, this research area is focused on modeling the three modalities and their complex interactions. The first Grand Challenge and Workshop on Human Multimodal Language aims to facilitate the growth of this new research direction in NLP community. The grand challenge is focused on multimodal sentiment analysis and emotion recognition on the recently introduced CMU Multimodal Opinion Sentiment and Emotion Intensity (CMU-MOSEI) dataset. The grand-challenge will be held in conjunction with the 56th Annual Meeting of the Association for Computational Linguistics 2018.

Communicating using multimodal language (verbal and nonverbal) shares a significant portion of our communication including face-to-face communication, video chatting, and social multimedia opinion sharing. Hence, it's computational analysis is centric to NLP research. The challenges of modeling human multimodal language can be split into two major categories: 1) studying each modality individually and modeling each in a manner that can be linked to other modalities (also known as intramodal dynamics) 2) linking the modalities by modeling the interactions between them (also known as intermodal dynamics). Common forms of these interactions include complementary or correlated information across modes. Intrinsic to each modality, modeling human multimodal language is complex due to factors such as idiosyncrasy in communicative styles, non-trivial alignment between modalities and unreliable or contradictory information across modalities. Therefore computational analysis becomes a challenging research area.

Organizers:

Amir Zadeh Language Technologies Institute, Carnegie Mellon University
Louis-Philippe Morency Language Technologies Institute, Carnegie Mellon University
Paul Pu Liang Machine Learning Department, Carnegie Mellon University
Soujanya Poria Temasek Laboratories, Nanyang Technological University
Erik Cambria Temasek Laboratories, Nanyang Technological University
Stefan Scherer Institute for Creative Technologies, University of Southern California

Invited Speakers:

Bing Liu, University of Illinois at Chicago (UIC)
Sharon Oviatt, Monash University
Roland Goecke, University of Canberra

Table of Contents

Grand Challenge and Workshop Program

July 20th 2018

9:00–10:30 Session 1

9:00–9:10 ***Opening Remarks***

9:10–10:00 *Keynote*
Bing Liu

10:00–10:10 *Getting the subtext without the text: Scalable multimodal sentiment classification from visual and acoustic modalities*
Nathaniel Blanchard, Daniel Moreira, Aparna Bharati and Walter Scheirer

10:10–10:20 *Recognizing Emotions in Video Using Multimodal DNN Feature Fusion*
Jennifer Williams, Steven Kleinegesse, Ramona Comanescu and Oana Radu

10:20–10:30 *Multimodal Relational Tensor Network for Sentiment and Emotion Classification*
Saurav Sahay, Shachi H Kumar, Rui Xia, Jonathan Huang and Lama Nachman

10:30–11:00 ***Coffee Break***

11:00–12:30 Session 2

11:00–11:50 *Keynote*
Sharon Oviatt

11:50–12:00 *Advances in Multimodal Datasets*
Paul Pu Liang

12:00–12:10 *Convolutional Attention Networks for Multimodal Emotion Recognition from Speech and Text Data*
Woo Yong Choi, Kyu Ye Song and Chan Woo Lee

12:10–12:20 *Sentiment Analysis using Imperfect Views from Spoken Language and Acoustic Modalities*
Imran Sheikh, Sri Harsha Dumpala, Rupayan Chakraborty and Sunil Kumar Kopparapu

July 20th 2018 (continued)

12:20–12:30 *Polarity and Intensity: the Two Aspects of Sentiment Analysis*
Leimin Tian, Catherine Lai and Johanna Moore

12:30–13:30 ***Lunch Break***

13:30–15:00 **Session 3**

13:30–14:20 *Keynote*
Roland Goecke

14:20–14:30 *ASR-based Features for Emotion Recognition: A Transfer Learning Approach*
Noé Tits, Kevin El Haddad and Thierry Dutoit

14:30–14:40 *Seq2Seq2Sentiment: Multimodal Sequence to Sequence Models for Sentiment Analysis*
Hai Pham, Thomas Manzini, Paul Pu Liang and Barnabas Poczos

14:40–14:50 *DNN Multimodal Fusion Techniques for Predicting Video Sentiment*
Jennifer Williams, Ramona Comanescu, Oana Radu and Leimin Tian

14:50–15:00 ***Grand Challenge Results***

15:00 ***Workshop End***