

**CALL FOR PAPERS**  
**ACM Transactions on Multimedia Computing, Communications, and Applications**

**Special Issue on Multimodal machine learning for human behavior analysis**

Analyzing human behaviors in multimedia data has become one of the most interesting topics in intelligent multimedia perception. Recently, with the widespread availability of advanced visual and non-visual sensors and a growing need of user-friendly interface, integrating multi-modality data for human behavior analysis, has received a great deal of research interests from the community of multimedia analysis. Compared with the traditional single-modality human behavior analysis, multi-modality human behavior analysis provides deeper level understanding to human identification and event detection, and a more comprehensive perspective for understanding the intrinsic interaction and connections of humans.

Although the studies of human behavior analysis in multimodality data are invaluable for both academia and industry, there are many fundamental problems unsolved so far, such as learning representation of human appearance and behaviors from multiple modalities, mapping data from one modality to another to achieve cross-modality human behavior analysis, identifying and utilizing relations between elements from two or more different modalities for comprehensive behavior analysis, fusing information from two or more modalities to perform a more accurate prediction, transferring knowledge between modalities and their representations, and recovering missing modality data given the observed ones. In the past decade, several multimodal machine learning models have been developed and shown promising results in some real-world examples such as multimedia descriptions and retrieval, which prepares us to exploit and develop effective multimodal machine learning algorithms for addressing fundamental issues in human behavior analysis.

This special issue aims at providing a forum for researchers from natural language processing, multimedia, computer vision, speech processing and machine learning to present recent progress in machine learning research with applications to multimodal multimedia data. The list of possible topics includes, but not limited to:

**Theories**

- Multimodal representation learning
- Multimodal translation and mapping
- Multimodal aligning
- Multimodal fusing and co-learning

**Applications**

- Multimodal affect recognition including emotion, persuasion and personality traits
- Multimodal media description including image captioning, video captioning and visual question answering
- Multimodal action recognition
- Cross-media information retrieval
- Large-scale multimodal datasets

Tutorial or overview papers, creative papers outside the areas listed above but related to the overall scope of the special issue are also welcome. Prospective authors can contact the Guest Editors to ascertain interest on such topics. Submission of a paper to ACM TOMM is permitted only if the paper has not been submitted, accepted, published, or copyrighted in another journal. Papers that have been published in conference and workshop proceedings may be submitted for consideration to ACM TOMM provided that (i) the authors cite their earlier work; (ii) the papers are not identical; and (iii) the journal publication includes novel elements (e.g., more comprehensive experiments). For submission information, please refer to the ACM TOMM journal guidelines (see <https://tomm.acm.org/authors.cfm>). Manuscripts should be submitted through the online system (<https://mc.manuscriptcentral.com/tomm>).

### **Important dates**

Submission deadline: April 15, 2019

First notification: June 15, 2019

Revision submission: August 15, 2019

Notification of acceptance: September 31, 2019

Online publication: March 2020

### **Guest editors**

Prof. Shengping Zhang,  
Harbin Institute of Technology, China, [s.zhang@hit.edu.cn](mailto:s.zhang@hit.edu.cn)

Dr. Huiyu Zhou,  
University of Leicester, United Kingdom, [hz143@leicester.ac.uk](mailto:hz143@leicester.ac.uk)

Prof. Dong Xu,  
University of Sydney, Australia, [dong.xu@sydney.edu.au](mailto:dong.xu@sydney.edu.au)

Prof. M. Emre Celebi,  
University of Central Arkansas, USA, [ecelebi@uca.edu](mailto:ecelebi@uca.edu)

Prof. Thierry Bouwmans,  
University of La Rochelle, France, [thierry.bouwmans@univ-lr.fr](mailto:thierry.bouwmans@univ-lr.fr)