

Feature Learning from RGB-D Data for Multimedia Applications

Background

Recent advances in 3D depth cameras using either structured light sensor or time-of-flight sensor have made it possible to capture depth images in real time. Thanks to these innovations, nowadays even a consumer-level device equipped with a pair of spatially aligned RGB camera and depth camera can already simultaneously provide high-quality visual (RGB) and depth information of a scene. This is a new type of data to interpret the human-centric environments, thus facilitating many existing multimedia-related applications such as immersive media, 3D SLAM, augmented reality, and smart assistive living. Among its follow-up research efforts, one important topic is to investigate how to efficiently represent a scene by taking advantage of the characteristics of both RGB and depth signals.

Feature learning is a technique that learns a transformation from a raw input to a compact representation that can describe the scene. Compared to the traditional feature representation methods, it allows for learning expressive features directly from the raw data without manual annotations or hand-crafted heuristic rules. The set of used features could be continuously modified during the learning procedure, which is able to adapt to a changing and more realistic environment. In the recent a few years, feature learning techniques (e.g., dictionary learning and deep learning) have been widely exploited to many applications, such as classification, segmentation, and scene understanding, etc. As the RGB-D data becomes available, researchers are keen to develop feature learning algorithms in this new area.

Objectives

The aim of this special issue is to solicit the state-of-the-art feature learning approaches and technical solutions in the area of multimedia technology and applications using RGB-D cameras. This special issue will bring together leading researchers in the related fields of 3D data capture and representation, RGB-D based human motion modeling, and human computer interaction via contactless sensor (e.g. Kinect), in order to advocate and promote research in vision and multimedia applications. The issue will provide a convincing forum for researchers and practitioners to disseminate their latest research results on feature learning for RGB-D or other multiple-channel data. The submission focusing on multiple-channel images, unmanned aerial vehicle (UAV), or satellite are also encouraged.

This special issue covers all aspects of vision and multimedia technology and applications using RGB-D cameras. Topics include but are not limited to:

- Feature learning for multiple-channel features
- Fast feature learning for complex features
- Augmented virtual reality
- Depth fusion, 3D reconstruction and modeling
- Human detection, tracking, re-identification, and activity understanding
- Human-machine interactions
- Scene understanding and segmentation
- Navigation, localization, SLAM and semantic mapping

- Object description, detection, tracking and recognition
- Annotation and retrieval of RGB-D data
- 3D multimedia data coding
- Multimedia content analysis and event detection
- Synthesis and animation of 3D human motions
- Immersive media applications: culture, entertainment, education, art, health, personal media, etc.
- Multimodal approaches to media analysis, indexing, search, and retrieval
- Multimodal approach on multiple-channel images, such as unmanned aerial vehicle (UAV), and satellite images.

Submission Details

Format: All the papers should be full journal length versions and follow the guidelines set out by Multimedia Tools and Applications:

<http://www.springer.com/computer/information+systems/journal/11042>.

Submission and Review: Manuscripts should be submitted online at <https://www.editorialmanager.com/mtap/> choosing "Feature Learning from RGB-D Data for Multimedia Applications" as article type. When uploading your paper, please ensure that your manuscript is marked as being for this special issue. All the papers will be peer-reviewed following the MTAP reviewing procedures.

Important Dates

Paper submission: October 31, 2015

Review results to authors: December 15, 2015

Revised paper submission: January 15, 2016

Final acceptance notification: March 1, 2016

Guest Editors

Prof. Baochang Zhang(Lead editor), Beihang University, Beijing, China (bczhang@buaa.edu.cn; bczhang@139.com)

Dr. Jungong Han, Civolution Technology, Eindhoven, the Netherlands (jungonghan77@gmail.com)

Prof. Ling Shao, Northumbria University, Newcastle upon Tyne, UK (ling.shao@northumbria.ac.uk)