Special issue on Benchmark Evaluation of RGB-D based Visual Recognition Algorithms

Machine Vision and Applications

Visual recognition is a critical component of machine intelligence. For a robot to behave autonomously, it must have the ability to recognize its surroundings (I am in the office; I am in the kitchen; On my right is a refrigerator). Natural human computer interaction requires the computer to have the ability to recognize human's gestures, body languages, and intentions. Recently, the availability of cheap 3D sensors such as Microsoft Kinect has made it possible to easily capture depth maps in real time, and therefore use them for various visual recognition tasks including indoor place recognition, object recognition, and human gesture and action recognition. This in turn poses interesting technical questions such as:

1. What are the most discriminative visual features from 3D depth maps? Even though one could treat depth maps as gray images, depth maps consist of strong 3D shape information. How to encode the 3D shape information is an important issue for any visual recognition tasks.
2. How to combine depth maps and RGB images? An RGB-D sensor such as Microsoft Kinect provides a depth channel as well as a color channel. The depth map contains shape information while the color channel contains texture information. The two channels complement each other, and how to combine them in an effective way is an interesting problem.
3. What are the most suitable paradigms for recognition with RGB-D data? With depth maps, foreground background separations are easier, and in general, better object segmentations can be obtained than with conventional RGB images. Therefore the conventional bag of feature approaches may not be the most effective approaches. New recognition paradigms that leverage depth information are worth exploring.

Scope

This special issue covers all aspects of RGB-D based visual recognition. It emphasizes on the evaluation on two benchmark tasks: ImageCLEF Robotic Vision Challenge (http://www.imageclef.org/2012/robot) and CHALEARN Gesture Challenge (http://gesture.chalearn.org/). The special issue is also open to researchers that did not submit runs to either of the two challenges, provided they will test their methods on at least one of the two datasets. In addition to the two benchmark tasks, researchers are welcome to report experiments on other datasets to further validate their techniques.

Topics include but are not limited to:

o   new machine learning techniques that are successfully applied to either of the two benchmark tasks
o   novel visual representations that leverage the depth data
o   novel recognition paradigms
o   techniques that effectively combine RGB features and depth features
o   analysis of the results of the evaluation on either of the two benchmark tasks

- theoretical and/or practical insights into the problems for the semantic spatial modeling task, and/or for the robot kidnapping task in ImageCLEF Robotic Vision Challenge
- theoretical and/or practical insights into the one-shot recognition problem in the CHALEARN Gesture Challenge
- computational constraints of methods in realistic settings
- new metrics for performance evaluations

Information for Authors:

Authors should prepare their manuscripts according to the author guideline from the online submission page of Machine Vision and Applications (http://www.editorialmanager.com/mvap/).

Important Dates (tentative):

- Manuscript submission deadline: January 30, 2013
- First round review decision: May, 2013
- Second round review decision: September, 2013
- Final manuscript due: November, 2013
- Expected publication date: January, 2014

Guest Editors:

- Barbara Caputo, Idiap Research Institute, Switzerland
- Markus Vincze, The Institute of Automation and Control Engineering, Austria
- Vittorio Murino, Istituto Italiano di Tecnologia, Italy
- Zicheng Liu, Microsoft Research, United States