Feature fusion for object recognition in RGB-D images

PAL Robotics, S.L.

Extended Abstract

Commodity RGB-D image sensors are revolutionizing entire sectors of computer vision. The introduction of the Kinect sensor from Microsoft has enabled a new generation of applications capable of reasoning directly about the 3D geometry of viewed scenes, persons and objects. It has had an especially acute impact in robotics, where autonomous systems can now recover accurate 3D information about scenes. In this Masters project we will examine the use of RGB-D sensors for object recognition and how to best fuse depth information with color and texture.

Traditional state-of-the-art approaches to object recognition are based on pooling of local features into global image descriptors that are informative for recognizing objects of predetermined classes [Chatfield 2011]. Recently, significant advancement in the state-of-the-art has been achieved through the use of Deep Convolutional Neural Networks (CNNs) to simultaneously learn local feature representations, pooling and classification in a single optimization framework [Krizhevsky 2012]. For both the classical and the CNN approaches, the fusion of multiple feature modalities (e.g. shape, texture, color and depth) into a single representation remains an active area of research [Gehler 2009, Gupta 2014].

In this project we will look at feature modality fusion techniques for object class recognition in RGB-D image streams. Specifically, we will investigate the use of techniques to fuse local RGB information with local depth information in order to improve object class recognition. The primary objective of the project will be to derive an effective fused representation (in terms of recognition performance) that is computationally efficient. The student will be expected to:

1. conduct a critical survey of the state-of-the-art in RGB-D object recognition (classical and CNN approaches);
2. investigate the potential of fusion approaches such as Portmanteau Vocabularies [Khan 2011] or KCCA [Lisanti 2014] for RGB-D feature fusion;
3. experiment with identified candidate fusion approaches on standard benchmark datasets [Browatzki 2011, Lai 2011] (or on data provided by PAL Robotics); and
4. conduct a thorough comparative analysis of all approaches.

Further Information
Andrew Bagdanov (CVC) -- bagdanov@cvc.uab.es
Bibliography


