

# Abstract

---

This thesis addresses the problem of learning object classification using multi-view range data. Class membership is determined by shared characteristics, which can be visual, structural or functional. The major steps in object recognition and object classification are: segmentation, feature extraction, object representation and learning.

This research introduces segmentation methods to decompose a scene into shape primitives. The first segmentation method is a new approach for producing high-quality planar segments, while the second method employs a commonly used, standard library for creating planar, cylindrical and spherical regions. A set of higher-level, relational features is extracted from the segmented regions. Thus, features are presented in three different levels: single region features, pair-region relationships and features of all regions forming an object instance.

The extracted features are represented as predicates in Horn clause logic. Positive and negative examples are produced for learning by the labelling and training facilities developed in this thesis. Inductive Logic Programming (ILP) is used to learn relational concepts from instances taken by a depth camera. As a result, a human-readable representation for each object class is created.

The methods developed in this research have been evaluated in experiments on data captured from a real robot designed for urban search and rescue, as well as on standard datasets. RoboCup Rescue competition arenas and other natural indoor scenes were the source of much of the data. There are also several published standard sets of range data that allow comparison with other 3D object classification methods. The results show that ILP is successful in recognising objects encountered by a robot and are competitive with the other state-of-the-art methods.

The main contribution of this thesis is in developing an object classification system that integrates data gathering, segmentation, relational feature representation, and relational learning that is capable of performing well in complex unstructured scenes.