Semantic object search using semantic categories and spatial relations between objects

Patricio Loncomilla, Marcelo Saavedra and Javier Ruiz-del-Solar
Advanced Mining Technology Center & Dept. of Elect. Eng., Universidad de Chile
{ ploncomi, jruizd}@ing.uchile.cl

Abstract. In this work, a novel methodology for robots executing informed object search is proposed. It uses basic spatial relations, which are represented by simple-shaped probability distributions describing the spatial relations between objects in space. Complex spatial relations can be defined as weighted sums of basic spatial relations using co-occurrence matrices as weights. Spatial relation masks are an alternative representation defined by sampling spatial relation distributions over a grid. A Bayesian framework for informed object search using convolutions between observation likelihoods and spatial relation masks is also provided. A set of spatial relation masks for the objects "monitor", "keyboard", "system unit" and "router" were estimated by using images from Label-Me and Flickr. A total of 4,320 experiments comparing six object search algorithms were realized by using the simulator *Player/Stage*. Results show that the use of the proposed methodology has a detection rate of 73.9% that is more than the double of the detection rate of previous informed object search methods.

Keywords: Semantic search, Informed search, Co-occurrence matrix

1 Introduction

Object search is an important ability for a mobile robot. Some previous work on object search are based on the use of spatial object-place relations, assuming that in a scene the searched object will be readily available to the robot's field of vision. However, there are sets of objects inside a setting that tend to appear near of each other as they have a particular spatial relation, making it possible to infer the existence of an object A, given an object B.

For example, when looking inside offices, it can be noted that the object "keyboard" is often very near the object "monitor". For a human, the ability to deduce that object A tends to be "near", "very near" or "far" from object B is a trivial task, since human beings can learn spatial relations between objects by observing a large number of similar settings. In this paper, we focus on giving robots the ability to find objects using existing semantic relations with other objects such as "near" or "far," in a given environment.

We conducted an exhaustive analysis of spatial relations between real-world objects that share a common use. For this objective, we created a database of spatial co-occurrences where distance values between objects are represented by the linguistic variables "very near," "near," "far," and "very far". We say that two objects "co-occur" when they appear together in several images showing a particular spatial relation.