Leveraging Semantic Context: Robot Planning with a Conditional Random Field Model

John G. Rogers III and Henrik I. Christensen

Service robots can assist humans in a variety of domestic tasks such as meal preparation, fetch and carry, and clean-up. Robots can leverage contextual cues within their environments to help them understand how to perform their duties. Context can take many forms such as that which exists between objects (a light switch is close to a door), as well as that which exists between places and objects (certain things are found in certain rooms). By leveraging place and object context, robots can be made to understand domestic environments and perform object search tasks more efficiently.

We have developed a technique described in [3] which uses a probabilistic model to represent the context of objects in places to perform an object search task in a domestic environment. We extend the idea of using *virtual objects and rooms* of [1] through the use of our probabilistic model. We leverage the context of places seen in a topological map when considering which unknown region to explore, to be the one most likely to contain the object for which the robot is searching. Each object found in a room will influence the label on the room and affect the likelihood of finding the target object.

A variety of software and hardware modules are used in our mobile robot system. Our robot is a Segway RMP200 mobile robot base with Hokuyo UTM30 as well as SICK laser scanners. The robot has a Directed Perception PTU 46-70 with an Asus Xtion Pro Live 3D camera, a Nikon D90 DSLR camera. The robot also has a Schunk gripper is also mounted on a linear actuator for picking up objects. The software is based upon the Robot Operating System (ROS) [2]. We use the GTsam nonlinear optimization library with a modular mapping framework that we have developed called OmniMapper. Room regions are segmented using a Gaussian region analysis technique described in previous work. Objects are segmented on tabletop surfaces through 3D point cloud analysis. Object recognition is performed by matching geometrically consistent SURF features against a database of known objects. If the object is not recognized in a database of known objects, then it is classified by a group of relevance vector machines trained on visual word histograms quantized from SURF features.

We introduced a probabilistic cognitive model (PCM) for place and object classification using conditional random fields in [3]. Objects and rooms are each represented as nodes in an undirected graphical model; their labels are given by multinomial distributions. There are two types of edges in this graph: between room nodes which indicate adjacency in a topological map, and between objects and rooms which indicate that object is within that room. The posterior distribution on this graphical model is approximated by loopy belief propagation. Objects which fail to be recognized by the object recognition module, but have been categorized by the object categorization module are given the specific measurement from the object classification module as their prior measurement. If the object recognition module recognizes the object, then it is clamped to the recognized class. Clamped nodes are used to select conditional probability tables in their neighbors and are skipped by the loopy belief propagation algorithm. If a new measurement is made of a previously mapped object, then the distribution with the lowest entropy is used as the measurement in the graphical model.

We have developed a planning module which uses the probabilistic model to select robot actions to perform an object search task. The PCM Planner selects actions in a Markov Decision Process framework. At each state, the robot chooses an action from the set *Move*, *Search*, *Examine*, *Fetch*. The *Move* action can be performed to transition to another room which is topologically adjacent to the one in which the robot currently inhabits. The PCM hypothesized by the planner for this action is the same as the current PCM, but with the robot in the desired adjacent room. The *Search* action is used to try to find objects in the current room using the segmentation module. The result of the *search* action in the hypothesized next state is that a new object is found by the robot. A uniform prior is placed on this new object; after the posterior is computed with LBP, the context of the current room will affect the label of the hypothesized object. The *Examine* action is selected by the planner to look at a previously segmented or categorized object and try to identify it directly with the recognition module. The hypothesized next PCM state upgrades the object categorized label distribution to a clamped, known recognition. The final action, *Fetch*, is terminal and represents the robot making its final choice; choosing this object as the solution to the object search task.

The PCM planner computes a sequence of actions leading it to the terminal *Fetch* action; however, since this relies heavily upon hypothesized results, we only execute the first action of the sequence and then re-plan based upon its result. The *Move* action sends coordinates of the center of the target room Gaussian as the destination for the motion controller. The *Search* action instructs the base controller to move to waypoints aligned with the major and minor axes of ellipsoidal regions described by the covariance matrix of the room Gaussian. While the robot is moving to these positions, objects on table surfaces are likely to be seen as they pass nearby. The *Examine* action moves the base to within 2 meters of the target object, saccades the PTU to aim the DSLR camera, and takes a high resolution image which is processed by the recognition and classification modules.

We have presented a technique for leveraging contextual cues in the form of room adjacency and object in room affinity in a Markov decision planner framework. The problem addressed is to leverage the context offered from room and object identification to find an element of an object class in an unknown indoor environment.

References

[1] A. Aydemir, M. Göbelbecker, A. Pronobis, K. Sjöö, and P. Jensfelt. Plan-based object search and exploration using semantic spatial knowledge in the real world. In *Proc. of the European Conference on Mobile Robotics (ECMR'11), Orebro, Sweden*, 2011.

[2] M. Quigley, B. Gerkey, K. Conley, J. Faust, T. Foote, J. Leibs, E. Berger, R. Wheeler, and A. Ng. ROS: an open-source robot operating system. In *ICRA Workshop on Open Source Software*, 2009.

[3] J. G. Rogers III and H. I. Christensen. A conditional random field model for place and object classification. In *IEEE International Conference on Robotics and Automation (ICRA)*, 2012.