Active modelling of an unknown object

Frederik Hagelskjær, Rudi Hansen, Kent Stark Olsen and Leon Bonde Larsen
Group number: 4. Supervisor: Unknown.

Introduction

Models are necessary in most applications of robotics. Often the models are build in advance and used by the robot as a priori knowledge, but this approach limits the application of robots in dynamic environments.

The main purpose of a model is to plan collision free trajectories for the robot and therefore the precision of the model is dependent on the task or sub task performed by the robot. In situations with high clearance to obstacles the planning can be based on a rather coarse model, but in cases where the robot needs to move close to obstacles, for example when grasping an object, a much finer resolution is required. The computational cost of collision checking is proportional to the level of detail in the model and therefore having a fixed resolution of the model will cause either waste of resources in high clearance situations or lack of precision in low clearance situations. It is therefore desired to use a coarse model in general, while being able to refine it whenever needed. This causes a need for active refinement of the model.

Actively building a model requires actuated sensory input in 3D as well as exploratory decision making to acquire and select the input maximising the information gain.

It is hypothesised, that a coarse a priori model of a robot environment can be actively refined, based on input from a camera controlled by the robot.

The idea presented in this work is to mount a camera on a robot arm and place an unknown object in the work cell. The robot will then autonomously move around the static object and build a 3D model of it. Using the estimated pose of the robot and images from the camera, a point cloud is generated and is used to gradually refine the model. The robot motion is based on desired poses maximising gain of information.

The problem of selecting the camera view that gains most new information is known as the Next Best View problem and many algorithms have been proposed, often with good results, although many have been tested in simulation only. Several different sensors are available ranging from mono- and stereo cameras over RGB-D and Time-of-Flight cameras to laser range scanners and some proposed hybrids.

For knowledge representation many approaches use evidence grids (Martin *et al.* 1996) such as the occupancy grid (Torabi *et al.* 2010, Perrollaz *et al.* 2012) representing the probability of a cell being occupied, either seeking to formulate and optimise an information gain function or to maximise configuration space entropy (Torabi *et al.* 2007). The advantage of the occupancy grid is that it has a compact and comprehensible format that can benefit from the octree data structure. Other uses of evidence grids are in the form of voxels labelled with type, surface normal and quality (Vasquez-Gomez 2009) or a 3x3 covariance matrix representing uncertainties in 3D (Foix *et al.* 2012). These methods are computationally more expensive, but holds much more information. Later methods adopt the surfel map representation of viewer facing discs known from medical scanner data representation instead of conventional polygonal modelling (Stückler *et al.* 2014).

There are many approaches to select the next best view. Some uses occlusion prediction (Ding *et al.* 2010), limit visual surface (Zhou *et al.* 2009a, Zhou *et al.* 2009b) or trend surface (Li *et al.* 2009) to focus the next view on the occluding lines of the current image. One approach (Khalfaoui *et al.* 2012) exploits the Mass Vector Sum that has been proven to be the null vector for closed models (Yuan *et al.* 1995), as a measure for the incompleteness of the model. This approach makes it possible to construct the surface without having a volumetric representation.

Materials and methods

To test the hypothesis a closed loop system based on the GOFAI environment interaction model (Pfeifer *et al.* 2007) was developed (Figure 1). Notice that there is no communication between the robot planner and the model, meaning that collision detection is only based on the a priori model of the work cell. The currently expected solutions for the individual parts of the system are described in the following.

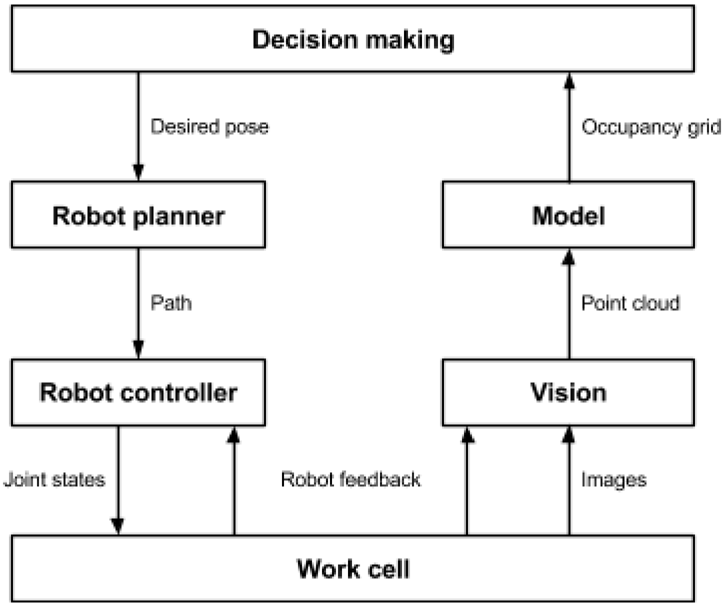

Figure 1 - Block diagram of the robot system

Preliminary test

As this project is still in its earliest stages, the exact solutions implemented are not yet known. The approach will be to first run a preliminary test, in which the basic parts of the system will be developed and evaluated. The purpose of the preliminary test is to have all parts of the system work together and to collect data that are as close to real as possible, thus forming a baseline for later tests. The evaluation of the test will be based on requirements specifications for the final system with regards to tolerances and constraints.

The test setup is based on a simplified block diagram (Figure 2) where the decision making consists of a preset trajectory and where the modelling component has been removed.

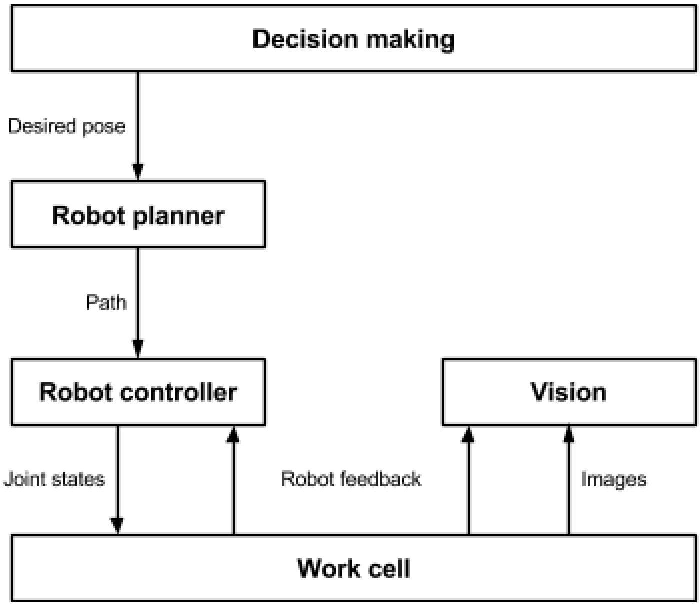

Figure 2 - Simplified block diagram

Robot planner

The robot planner executes the desired pose based on the RRT-connect algorithm (Kuffner *et al.* 2000) subject to the constraint that the camera always points to the center of the object being modelled. The path is post-processed using path pruning and path clearance (Geraerts *et al.* 2007). The robot planner will be based on the moveIt stack in ROS and collision checking will be based on the a priori model of the work cell.

Robot controller

The robot controller further processes the path adding kinematics and time tessellation to meet velocity and acceleration constraints. The controller interpolates the joint states and handles closed loop control with feedback from the workcell. The robot controller is also based on the moveIt stack in ROS.

Work cell

The workcell contains a 6dof RX60 robot arm with a stereo camera mounted on the end effector. The work cell interface is based on a ROS node communicating with the robot and a node broadcasting data from the camera.

Vision

The vision module takes input from the stereo camera mounted on the end effector of the robot. A disparity image is generated based on epipolar geometry and the result is passed on as a point cloud.

Model and decision making

Although not implemented in the preliminary test, different methods for surface reconstruction and volume representation are being evaluated based on the output of an RGB-D camera. The interface between model and decision making is also being evaluated and as part of this the representation of knowledge and the filtering of point clouds. The decision making will most likely be based on the formulation of information gain for next best view selection although other methods are also being evaluated.

The computer vision focus of this work is the active view planning while the robotics focus is to run the system independent of RobWork, meaning that for example path optimization and constrained planning has to be implemented.

References

Ding, H., He, B., & Zhou, X. (2010). A Novel Occlusion Planning Method for Unknown 3D Objects Automatic Reconstruction. *2010 International Conference on Artificial Intelligence and Computational Intelligence*, 463–467. doi:10.1109/AICI.2010.103

Foix, S., Kriegel, S., & Fuchs, S. (2012). Information-gain view planning for free-form object reconstruction with a 3d tof camera. *Advanced Concepts for …*. Retrieved from http://link.springer.com/chapter/10.1007/978-3-642-33140-4_4

Geraerts, R., & Overmars, M. H. (2007). Creating High-quality Paths for Motion Planning. *The International Journal of Robotics Research*, *26*(8), 845–863. doi:10.1177/0278364907079280

Khalfaoui, S., Seulin, R., Fougerolle, Y., & Fofi, D. (2012). View planning approach for automatic 3d digitization of unknown objects. *Computer Vision–ECCV 2012 …*, 496–505.

Kuffner, J. J., & LaValle, S. M. (2000). RRT-connect: An efficient approach to single-query path planning. *Proceedings 2000 ICRA. Millennium Conference. IEEE International Conference on Robotics and Automation. Symposia Proceedings (Cat. No.00CH37065)*, *2*(Icra), 995–1001. doi:10.1109/ROBOT.2000.844730

Li, Y. F. (2009). A view planning method for automatic 3D modeling based on the trend surface and limit region. *2009 IEEE Intrumentation and Measurement Technology Conference*, (May), 803–808. doi:10.1109/IMTC.2009.5168560

Martin, M., & Moravec, H. (1996). Robot Evidence Grids., (March). Retrieved from http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=ADA310957

Perrollaz, M., Yoder, J.-D., Negre, A., Spalanzani, A., & Laugier, C. (2012). A Visibility-Based Approach for Occupancy Grid Computation in Disparity Space. *IEEE Transactions on Intelligent Transportation Systems*, *13*(3), 1383–1393. doi:10.1109/TITS.2012.2188393

Pfeifer, R., Lungarella, M., & Iida, F. (2007). Self-organization, embodiment, and biologically inspired robotics. *Science (New York, N.Y.)*, *318*(5853), 1088–93. doi:10.1126/science.1145803

Stückler, J., & Behnke, S. (2014). Multi-resolution surfel maps for efficient dense 3D modeling and tracking. *Journal of Visual Communication and Image Representation*, *25*(1), 137–147. doi:10.1016/j.jvcir.2013.02.008

Torabi, L., Kazemi, M., & Gupta, K. (2007). Configuration space based efficient view planning and exploration with occupancy grids. *2007 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2827–2832. doi:10.1109/IROS.2007.4399451

Torabi, L., & Gupta, K. (2010). Integrated view and path planning for an autonomous six-DOF eye-in-hand object modeling system. *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 4516–4521. doi:10.1109/IROS.2010.5651203

Vasquez-Gomez, J. I., Lopez-Damian, E., & Sucar, L. E. (2009). View planning for 3D object reconstruction. *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 4015–4020. doi:10.1109/IROS.2009.5354383

Yuan, X. (1995). A mechanism of automatic 3D object modeling. *Pattern Analysis and Machine Intelligence, IEEE …*, *17*(3), 307–311. Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=368196

Zhou, X., He, B., & Li, Y. F. (2009). Research on new view planning method for automatic reconstruction of unknown 3D objects. *2009 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, 965–970. doi:10.1109/ROBIO.2009.5420729

Zhou, X., He, B., & Li, Y. F. (2009). View Planning Method Based on the Visual Region of the Vision System. *2009 International Conference on Artificial Intelligence and Computational Intelligence*, 196–200. doi:10.1109/AICI.2009.62