Chapter 8

Conclusions and Future Work

This thesis addressed the depalletizing problem, in the context of which a number of box-like objects piled on a platform, are automatically unloaded by a vision guided robotic manipulator. From the hardware point of view, our approach to the problem comprises a six degree of freedom industrial robot, on the hand of which a time-of-flight laser sensor is mounted. Object grasping is performed by a vacuum gripper, mounted as well on the robot's hand. Given a configuration of a piled objects, our system scans its top side, and a range image of the configuration is acquired. Then the range image is analyzed, and all graspable objects in the pile are recovered. Subsequently, the pose of the center of gravity of the exposed surfaces of the recovered objects are sent to the robot, which grasps the objects and places them at a user-defined position. This operation continues iteratively until no objects lie on the platform.

Usage of a laser sensor for data acquisition results in a *flexible* system: Our system does not depend on the lighting conditions at the installation site on one hand, on the texture of the objects in the configuration on the other. Besides, our system is of *low cost*, and there are two reasons why this is so. Firstly, the sensor has a relatively low price (about 3.000 Euro), and the same holds for the gripper (about 1000 Euro). Secondly, the costs related to system installation are as well low: No tedious sensor calibration procedures are required for rendering our system fully operational. The only thing required for this purpose is just mounting the sensor and the gripper on the robotic arm. This is a process that can be carried out within a couple of minutes.

This thesis primarily focused on the object recovery problem. Our strategy for object recovery has two aspects, both of which are based on modeling the target objects with *geometric parametric models*. According to the first aspect, for the description of which the largest portion of this thesis is assigned, the target objects are modeled via globally deformable Superquadric entities. This approach, performs object recovery by maximizing the posterior probability of the parameters of the objects in the image. The maximization process is performed in two steps: Firstly, a hypothesis about the values of the models' parameters is generated. Then, this hypothesis is locally refined. For both the generation and the refinement of the hypothesis, not only the range image is taken into consideration, but an edge image, obtained from the latter by means of an edge detection operation.

The particular recovery approach exhibits many advantages: The most important is *flexi-*

bility, since it can deal with both neatly placed or jumbled object configurations of objects with *arbitrary* dimensions. Besides, it is able to deal with deformable box-like objects, such as bags or pillows. Furthermore, initial experiments indicate that the approach is able to handle not only box-like objects, but all objects which can be modeled by superquadrics of arbitrary shape. Moreover, this approach is robust: As shown in section 6.7, given 40 configurations in which 174 graspable objects are included in total, our system gave 159 positive responses. Our recovery approach is accurate as well: The average radial euclidean distance of the range points belonging to the objects to the model describing these points was 8.68 mm per graspable object on the average. In addition, the average euclidean distance between the boundary of the region of points belonging to a graspable object, and the boundary of the model describing the object in the image plane was 1.01 pixels per object on the average. Finally, our system proved to be computationally efficient. The average time required for processing a range image of an object configuration was about 92 seconds, and the time required for recovering a single object was about 16 seconds on the average, in a Pentium 4 2.8GHz PC. Note, that our approach has the potential of parallel implementation, and in this event the latter time will be the average time required for processing the entire image.

An alternative method for object recovery discussed in this thesis, is based on the Hough Transform. Object recovery is guided by boundary information within this strategy. More specifically, the rectangular boundary of the exposed surface of the objects is used as modeling entity. Object boundary localization is performed by Hough transforming the input edge image. The seemingly difficult problem of the recovery of the three dimensional rectangle boundary is solved in two stages: Firstly, the pose of the boundary is recovered, via three dimensional vertex detection in the edge image. Secondly, the dimensions of the boundary are extracted. Each of those problems is further decomposed into easier subproblems, each of which is addressed using an one-dimensional Hough transform. This approach is flexible, since it can deal with both neatly placed objects in layers and configurations of jumbled objects, as well as with objects of arbitrary dimensions. However, it cannot deal with deformable objects. When dealing with rigid boxes, the approach exhibits high accuracy and robustness. In terms of computational efficiency, it surpasses the strategy discussed in the preceding paragraph: Processing of an input image lasts about 30 seconds on the average, on a Pentium 3, 600MHz PC. In conclusion, the particular approach should be used, whenever a priori knowledge indicating that all objects in the pile are rigid boxes exists, and the approach discussed in the preceding paragraphs in any other case.

Our robotic system contributes to the state of the art of vision guided systems for depalletizing of box-like objects, since it combines all advantages mentioned in the preceding paragraphs (flexibility, robustness, efficiency, accuracy, low cost), which is not the case for any existing system up to our knowledge. Besides, our image analysis strategies contribute to the state of the art of methods for recovery of multiple parametric geometric models in range images in two ways: The approach using superquadrics as modeling elements extends the recover-and-select paradigm, the most popular strategy for superquadric segmentation, since it employs boundary information in addition to range information for performing the recovery task. The approach using the Hough transform for object recovery, demonstrates that it is feasible to use the Hough transform for efficiently dealing with problems of relatively high dimensionality, the recovery of three dimensional rectangles for example, by decomposing the transform into easier subproblems.

The effectiveness of both approaches for object recovery, depends on the quality of boundary information. A typical example where boundary information is incomplete is when objects of similar dimensions are tightly placed one after another in layers. In this case, the resolution of the laser sensor is not high enough for capturing the edges between the objects, and as a consequence the number of false positive responses our system delivers drastically increases. In the configurations generated for experimenting with the system, this was taken into consideration: When creating configurations of neatly placed objects, we did not place the objects in distance less than 1cm. A solution to this problem is the incorporation of a sensor of higher accuracy, which would inevitably increase the costs of our system. Besides, incorporation of an additional sensor, for example an intensity camera, could provide additional boundary information and alleviate the problem.

Within this thesis, we did not extensively address the problem of recovering the height of the objects. The reason for this is that this problem cannot be solved in its generality, since no adequate range information for determining the height can be acquired by one viewpoint on the top side of the configuration. While the recovery of the height of a box-like object is not important when object grasping is of interest, it is necessary when accurate object placement is required. This is the case when we want the robot to execute the *palletizing* task as well, that is to neatly place the objects on a target platform. It seems that the most efficient way to address this problem is to use an additional sensor, such as a light curtain [9]. Light curtains have a very high repeatability, and they are commonly used in safety engineering to protect people who access a certain area, such as the working area of a hydraulic press. The height recovery scenario using a light curtain could be as follows: After object grasping the robotic hand moves in this way so that the bottom of the grasped object approaches the light curtain. As soon as the bottom side of the parcel touches the light curtain the movement towards the light curtain stops. At that time the height of the object can be recovered, since the current position of the robotic hand is known.

Our primary future target is to complete the design and to continue experimentation of our approach for the recovery of superquadrics of arbitrary shape. As discussed in section 5.4, initial experiments conducted in this direction are very encouraging. In addition, we intend to expand the abilities of the existing system for depalletizing of box-like objects by incorporating additional sensors. The most direct steps towards this direction, is the incorporation of an intensity camera for acquiring more accurate boundary information in cases where the resolution of the existing sensor is not enough on the one hand, and the incorporation of a light curtain for the recovery of the objects' height on the other.