Robust Recovery of Piled Box-Like Objects in Range Images

Dissertation

zur Erlangung des Doktorgrades
der Fakultaet fuer Angewandte Wissenschaften
an der Albert-Ludwigs-Universitaet Freiburg im Breisgau

vorgelegt von

Dimitrios Katsoulas
aus Elateia, Lokridos

2004

Thesis committee:	Prof. Dr. Matthias Teschner (Vorsitz)
	Prof. Dr. Wolfram Burgard (Beisitz)
	Prof. Dr. Hans Burkhardt (Betreuer)
	Prof. Dr. Ales Leonardis (Prüfer)

Date of defense: 04 May 2005

Acknowledgments

I take this opportunity to thank all people involved in a variety of ways in the preparation of this work.

I would firstly like to deeply thank my advisor Prof. Dr. Hans Burkhardt, for strongly believing in my potential to carry out this research, for encouraging me, and for supporting my stay in Freiburg.

In addition, I thank Prof. Dr. Ales Leonardis, by the work of whom a considerable part of this contribution is inspired, for revising the thesis.

Besides, I thank all my colleagues at the department for their constant affection, and encouragement. In particular, I thank Nikos Canterakis for his advice, his promptness to answer all my questions, and for proof reading a part of this thesis.

Furthermore, I thank my colleagues Lothar Bergen, Stefan Rahmann, Marco Reisert, and Sami Romdhani for insightful discussions, Bernard Haasdonk, Klaus Peschke, and Gerd Brunner for proof reading parts of the thesis, my students Christian Cea and Andreas Werber for their efficiency, as well as Stefan Teister for the tireless technical support.

Living in a foreign country for a longer time, may become unbearable without the devotion of good friends. Fortunately enough, I had the opportunity to have such friends. They stood by me in difficult times, which I immeasurably appreciate.

Many thanks go to my parents and my sister, as well as to all my friends in Greece, for their love and understanding. I missed them all very much!

Last but not least, I greatly thank my uncle Demosthenes Kotionis, who illuminated and inspired throughout these years, and to whom this work is dedicated.

Zusamenfassung

Diese Arbeit befasst sich mit dem robotergestützten *bin-picking* Problem, bei dem eine Anzahl Objekte unterschiedlicher Abmessungen, Oberflächenstruktur und Art automatisch erkannt, ergriffen und von einer Palette (rechteckigen Unterlage), auf der sie liegen, zu einem bestimmten benutzerdefinierten Ort transportiert werden müssen. Im Speziellen beschäftigen wir uns mit dem Problem der Entpallettierung, bei dem die Palette deformierbare kistenähnliche Objekte enthält. Der Bedarf für zuverlässige und allgemein einsetzbare automatische Entpalettierungssysteme besteht vor allem in der Automobil- und Nahrungsmittelindustrie. Automatische Entpalettierungssysteme sind von großer Bedeutung, da sie eine Aufgabe ausführen, die für Menschen sehr monoton, anstrengend und manchmal auch gefährlich ist. Zusätzlich kann die zuverlässige und schnelle Umsetzung von der Ausgangspalette zur Zielposition (Palette, Fließband, usw.) die logistischen Abläufe in der Industrie und der Lagerhaltung deutlich beschleunigen, wodurch viel Zeit und Arbeit gespart und die Kosten gesenkt werden.

Für die Handhabung der Objekte benutzen wir einen Industrieroboter mit sechs Freiheitsgraden auf dessen Greifer ein Lasersensor zur bildlichen Erfassung der Stapel angebracht ist. Ein Vakuumgreifer, der ebenfalls am Roboter montiert ist, wird benutzt um die erkannten Objekte an ihren sichtbaren Oberflächen zu greifen. Der Entpalletierungsprozess besteht aus drei Teilen: Erstens, die Oberseite des Stapels wird gescannt, und ein Tiefenbild wird gewonnen. Zweitens, das Bild wird analysiert, und die greifbaren Objekte im Stapel werden lokalisiert. Drittens, der Roboter greift die lokalisierten Objekte von ihrer sichtbaren Oberfläche, und liegt Sie zu einer benutzerdefinierten Position. Dieser Prozess ist iterativ durchgeführt, bis die Palette entleert ist.

Diese Dissertation richtet sich besonderes auf das Objektlokalisierungsproblem. Unsere Strategie für Objektlokalisierung ist modelbasiert. Sie verwendet geometrische parametrische Einheiten zur Objektmodellierung, und besteht aus zwei Aspekten, jede von denen im Zusatz zum Inputtiefenbild ein Kantenbild, das durch einen Kantengewinnungsprozess erhalten ist, benutzt wird: Erstens, global deformierbare Superquadriken werden zur Objektmodellierung verwendet. Objektlokalisierung ist als Optimierungsproblem gestellt, bei dem, ein Tiefenbild gegeben ist, die a-posteriori Wahrscheinlichkeit der Parameter aller greifbaren Objekte maximiert wird. Unser Verfahren erweitert *Recover-and-select*, den weit verbreiteten Ansatz zur Rückgewinnung von Superquadriken in Tiefenbilder, weil es sowohl Tiefeninformation als auch Kanteninformation berücksichtigt. Das ist der Hauptgrund warum unser Verfahren den *Recover-and-select* sowohl an Robustheit als auch an Effizienz übertrifft.

Zweitens, der Rand der sichtbaren Oberflächen der Zielobjekte ist als drei-dimensionales Rechteck modelliert. Die Houghtransformation wird verwendet zur Lokalisierung der Zielobjekte im Kantenbild. Das schierige Problem der Rückgewinnung drei-dimensionaler Rechtecken ist gelöst durch die Zerlegung der Houghtransformation in zwei Probleme von geringer Dimensionalität: Die Lage der Objekte wird zuerst berechnet, und danach ihre Dimensionen. Der größte Vorteil dieses Ansatzes ist seine Effizienz.

Die Entscheidung welcher Ansatz benutzt werden soll, hängt davon ab in wie fern die Objekte starr sind. Die zweite Strategie berücksichtigt keine Objektdeformationen, ist aber schneller

als die erste. Infolgedessen, wenn wir wissen dass der Stapel nur starre Kisten enthält, wird die zweite Strategie verwendet, und sonst die erste. Experimentelle Ergebnisse beweisen dass das resultierende System eine Vielfalt von Vorteilen zeigt, zum Beispiel Flexibilität, Robustheit und Effizienz. Die Kombination dieser Vorteile kann in existierenden Systemen nicht gefunden werden.

Abstract

This work addresses the vision-guided, robotic bin-picking problem, in the context of which a number of piled objects, should be localized, grasped and transferred by a robotic hand from the position they reside, to a specific place defined by the user. We deal in particular with the depalletizing problem according to which deformable box-like objects piled on a rectangle platform, the *pallet*, should be unloaded. The requirement of a robust system for dealing with this problem stems from almost all industrial sectors, and is expected to substantially reduce the costs associated with product handling and distribution.

From the hardware point of view, our system comprises a six degrees-of-freedom industrial robotic arm, on the hand of which a laser sensor is mounted for data acquisition. Besides, a vacuum gripper is mounted on the hand of the robot for object grasping. The object removal process is as follows: Firstly, the top side of the object configuration is scanned by linearly moving the robotic hand along the pallet, and a range image is acquired. Secondly, the image is analyzed and the graspable objects in the pile are localized. Thirdly, the robot grasps the recovered objects from their exposed surfaces and places them at a user defined position. This procedure is executed iteratively, until no objects lie on the pallet.

This thesis mainly focuses on the object localization or recovery process, that is, the way in which given a range image the position and dimensions of the objects is determined. Our strategy for object recovery is model based, uses geometric parametric entities for object modeling, and has two aspects, in both of which, in addition to the input range image a boundary image obtained by the former by means of edge detection is employed. Firstly, globally deformable superquadrics are used for modeling our target objects. The object recovery is posed as an optimization problem, in the context of which given the input range image, the posterior probability of the parameters of all graspable objects in the pile is maximized. Our approach extends the recover-and-select paradigm, the most widespread framework for superquadric recovery from range images, since it incorporates object boundary information into the recovery process. This is the main reason why our approach outperforms the recover-and-select framework in terms of both computational efficiency, and robustness.

Secondly, the boundary of the exposed surfaces of the target objects is modeled as a three dimensional rectangle. The Hough transform is employed to recover the target objects from the boundary image of the object configuration. The seemingly difficult problem of recovering three dimensional rectangles is straightforwardly solved by decomposing the Hough transform into two problems of lower dimensionality: The pose of the objects is recovered, followed by the recovery of their dimensions. This results to a computationally efficient framework.

The decision on which strategy should be adopted for object recovery depends on the rigidity of the objects. The latter strategy does not account for object deformations, but it is faster than the former. Hence, if we know beforehand that the configuration comprises rigid boxes only, the latter strategy is used and in every other case the former. Experimental results demonstrate that the resulting robotic system exhibits a variety of advantages such as robustness, flexibility, accuracy, and computational efficiency, the combination of which

cannot be found in any existing system up to our knowledge.

Contents

List of Figures

List of Tables