Signal Processing 120 (2016) 721–730

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

Error-tolerant manipulation by caging

Weiwei Wan, Feng Lu*, Rui Fukui

The University of Tokyo, Tokyo, Japan

ARTICLE INFO

Article history:
Received 4 September 2014
Received in revised form
19 November 2014
Accepted 6 December 2014
Available online 15 December 2014

Keywords:
Caging
Grasping
Robot finger

ABSTRACT

This paper delivers a preliminary attempt to find the optimized caging positions for a three-finger robotic hand designed for home-use logistical environments. The main idea behind optimizing caging positions falls in that optimal caging can afford largest margins to stop target objects from escaping into infinity. By employing the advantages of largest margins, optimal caging grasp can be robust enough to endure dramatic perception noises or errors and low sensing resolutions. This paper optimizes object grasping towards caging. Specifically, our algorithm utilizes Genetic Algorithm (GA) to accelerate the searching procedure and evaluate a fitness of the GA population by examining a combination of max–min, which corresponds to intersections of neighbour fingers' $\mathcal{CC}$ space margins, and least inter-finger distance for optimization. Simulation results show that the manipulation strategy proposed in this paper could in the worse case coordinate with sensors whose resolution are less than one pixel per centimeter.

1. Introduction

Many work has been devoted to robot manipulation in the past decades, including finger synthesizing [1], multiple robot cooperation [2], etc. Despite the various projects and publications, the research in manipulation can be divided into two groups where one group studies the manipulation problem from the viewpoint of actuation while the other group studies the problem from the viewpoint of perception.

In actuation, contemporary interests mainly lie in closures like form/force closure [3] and object closure/caging [4]. There also exists many practical applications of nonprehensile manipulation. Previous researches in closure tend to optimize force synthesizing for external force sets [5] or specific tasks [6], sometimes taking into kinematic constraints of hardware structures [7]. Theoretically, these optimization are complete and interesting. However, seldom some pragmatic systems consider closures due to toughness of perception. Traditional closure research [8] assumes exact positions, orientations and shapes of target objects. They suffer from crashing in real applications where perfect perception information cannot be obtained. Caging [9], as an extension to force closures, offers a passive and force-less way of manipulation. Although caging does not require exact force, it is far from error-tolerant synthesizing. Even the preliminary step, caging test, has been demonstrated challenging.

In perception, researchers try to build target information into boundary/surface clouds, curves, polygons or polyhedrons by using image segmentation [10,11]. Some works, such as curve fitting [12,13], have achieved certain success in cooperation with actuation. The cooperation involves two procedures. A geometric shape is modeled in the first procedure and acts as the medium of synthesizing in the second procedure. However, it is not always satisfying and may collapse with novel or unusual target shapes [14]. Some other works try to match target objects with database models, or namely, to recognize target objects. Modeling and retrieving indicate an excellent solution to

* Corresponding author.
E-mail addresses: wanweiwei07@gmail.com (W. Wan), lufeng@ut-vision.org (F. Lu).

http://dx.doi.org/10.1016/j.sigpro.2014.12.006

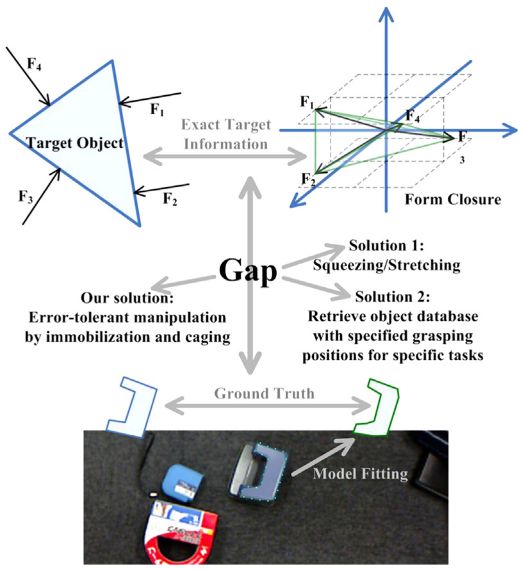

Fig. 1. The gap between perception and actuation.

offer precise shape information, even in the presence of overlaps. Nevertheless, errors arise from scale, translation and orientation [15]. Ambiguous configurations usually lead to major failure of actuation. Recently, researchers begin to seek for remedies of configuration errors by matching more precisely or modeling with larger compliance. But it remains hard to evaluate the flexibility.

Both research in actuation and perception have made conspicuous achievements. However, there is a gap between them. Researches in actuation take little consideration of perception information while the synthesizing approaches employed in perception researches guarantee no robust manipulation. The gap motivates us to explore synthesizing approaches that best tolerate perception errors. Fig. 1 demonstrates the gap and some popular solutions towards it.

In order to fill up the gap, we propose a robust finger synthesizing approach in this paper by considering model errors caused from the perception phase. Intrinsically, the approach is immobilization synthesizing that ensures (1) in the best case when perception is perfect, a target object is immobilized by fingers (2) even though certain perception errors occur, the target object is still constrained in a compact configuration region. If each finger is taken as a single-parameter link, our proposal finally converges into another immobilization by shrinking individual parameters.

Our approach is based on the relationship between finger objects in $\mathcal{C}$ space and $\mathcal{CC}$ space. It seeks maximum error-tolerance by simultaneously maximizing the breaking margins formed by intersections of adjacent finger neighbours and minimizing rotation moments caused by finger scatter. Major contributions of our proposal are as follows: (1) The proposal in this paper not only works with 2D planar objects but also works with 3D targets. (2) It can choose finger numbers automatically according to actual requirements or limitations of robustness. (3) A redundant solution is proposed to deal with concave objects, rather than 2-finger squeezing or stretching.

Our work is like [16,17] where correlative models or hierarchical models were developed for multimedia recognition. To our best knowledge, this paper is the initial work which applies such ideas to robotic manipulation. Using a max–min constraint plus a least inter-finger distance constraint for optimization essentially shares the same idea with the multi-modal based technique in signal processing. We use the optimization to process the vision information collected from KINECT sensors. We believe our work is of great interest to both researchers in robotics and signal processing.

The organization of this paper is as follows: background works like immobilization, $\mathcal{C}$ space, $\mathcal{CC}$ space, caging test, etc., are presented in Section 2. Section 3 discusses our consideration in perception errors and the routine of robust synthesizing in planar environment, followed by an extension to 3D environment in Section 4. These two sections compose major contributions of our work. The strategy to deal with concave objects is presented in Section 5 to ensure completeness of our proposal. Section 6 shows some experiments and analysis with both real sensor data and virtual 3D models. Finally, conclusions are drawn in the last section.

2. Background works

Form closure theoretically presents how to synthesize contacts that can resist certain external forces [18]. It requires at least $n_{dof}+1$ force contacts to grasp a target object with n_{dof} degree of freedom. Formally, form closure can be achieved by fulfilling $0 \in \text{int}(\text{conv}(\{\mathbf{f_i}\}))$. Although form closure indicates an important conclusion, it is not as practical with real robotic hands where friction and finger constraints always exist and finger number seldom satisfies requirements.

In actual application, it is pragmatic to simply take into account immobilization grasp [19]. Fingers in immobilization grasp can be considered as fixed obstacles once they are placed to certain contact configurations. In that case, fingers do not exert forces explicitly but bear wrenches caused by target weights and external forces passively. Immobilization grasp owns reasonable benefits compared with the loose equilibrium grasp which depends too much on target properties and the strict form closure grasp. An immobilization grasp requires at least 2 or $n_{dim}+1$ [1] to $2 \times n_{dim}$ finger contacts to immobilize an object in n_{dim} dimensional space.

Given a finger $\mathcal{A}_i$, we employ $\mathcal{C}_{oi}$ to indicate its correspondent configuration space ($\mathcal{C}$ space) presence. Immobilization grasp means that a target object, denoted by a configuration point q in $\mathcal{C}$ space, is isolated as a single point. Consequently, $\mathcal{C}$ space can be divided into three parts when immobilization grasp is formed, namely q, $\cup_{i=0}^{n}\mathcal{C}_{oi}$ and $\mathcal{C}_f$ ($\mathcal{C}_{free}$). According to this definition, a circle can never be immobilized since its 2 order derivative [20] is equal to 0 and no independent q can be isolated. At least four fingers are required to immobilize a semi-circle.

[1] Two fingers are sometimes enough for certain concave targets. However, we prefer $n_{dim}+1$ in this paper for safety and convenience.