Two-Fingered Grasp Planning for Randomized Bin-Picking

Donna C. Dupuis*, Simon Léonard*, Matthew A. Baumann*, Elizabeth A. Croft* and James J. Little†
*Collaborative Advanced Robotics and Intelligent Systems Laboratory
University of British Columbia, Vancouver, BC, Canada
Email: {dupuisd, sleonard, mabauman, ecroft} @mech.ubc.ca
†Department of Computer Science
University of British Columbia, Vancouver, BC, Canada
Email: little@cs.ubc.ca

***Abstract*— This work aims to increase reliability and reduce cycle time in order to realize a commercially-viable vision-guided robotic bin-picking system. We present a novel method for two-fingered grasp generation and target selection for bin-picking of randomized parts. We also propose a definition for grasp robustness, and use this to formulate a new grasp quality measure. A densely-sampled set of grasps is generated and evaluated using our proposed quality measure. The highest-quality grasps are then used to provide more valid picking options in the context of a randomized pile of parts, and to determine the best part to pick up. Our experimental results show a substantial increase in the average number of valid picking options when compared with a typical industrial approach for target selection.**

I. Introduction

In recent years, there has been increasing interest in industry towards developing a commercially-viable vision-guided robotic bin-picking (VGRBP) solution. Such a system would need to be highly reliable; that is, it must be able to continually pick parts out of one or more bins without exceeding an average cycle time per successful pick-and-place operation. To be successful, it is estimated that a bin picking solution should have an average cycle time of 10 seconds or less per operation. Due to the nature of randomly-situated parts within a bin, meeting this requirement is challenging, and is one of the main reasons why randomized bin picking systems have yet to be widely adopted by industry. For example, given a set of pre-defined grasping points on a particular part and a set of candidate parts within the context of a randomized bin, in many cases the pre-defined grasps are obstructed by neighbouring parts or by the walls of the bin. In such cases, the grasps are not feasible since they result in collisions with the gripper. If there is a limited number of pre-defined grasps, it is possible that all grasps for all candidate parts are infeasible, resulting in no viable options for picking.

In some systems, if no valid picking option exists, a second attempt is made at locating a viable candidate; for example, by taking a closer look at the pile, or by mechanically stirring the parts [1] and then re-examining the pile. However, these solutions increase the cycle time.

One can expect that increasing the number of grasping options for a given cycle will reduce the probability of having no feasible picking options, and consequently, reduce overall time spent searching for more candidates. In VGRBP, one way to increase the number of grasping options is to improve the vision recognition system so that more parts are recognized and localized during each cycle. Much research has already been done in the area of computer vision for this purpose (e.g., [2], [3], [4], [5], and [6]). However, once these parts are found, the best candidate must be selected, an issue not widely addressed in the bin-picking literature. A common approach is to select the top-most object in the pile, as is done in [2], [3], and [5]. This can be accomplished using image segmentation methods described in [7] and [8]. Although this would likely produce feasible targets for picking in many cases, it is not clear which part to select when multiple parts are considered to be on top, or when parts are entangled such that no part can be clearly distinguished as being on top. Optimizing the selection of a feasible target is one way to increase system reliability, and is addressed in this work.

Another way to increase the number of picking options is to increase the number of possible high-quality grasps for a given part by sampling the grasp space. Grasp sampling of a particular object has been addressed in [9] and [10]. In [9], objects are represented as a collection of primitive 3-D shapes, each of which is manually assigned a set of grasp starting positions and pre-grasp shapes. In [10], objects are represented by superquadratic decomposition trees in order to reduce the space of possible grasps, and the surface of each superquadratic is sampled at a uniform interval.

Herein, we present a novel method of generating and evaluating a densely-sampled set of grasps, and describe a way to use this set to select the best candidate part to pick up. Grasp generation is tailored for a two-fingered gripper, such as the one shown in Fig. 1, as such grippers are commonly used in industry [11]. Our method is broken down into two stages: (1) offline generation of many high-quality two-fingered grasps for a given part, and (2) online evaluation of candidate parts using these high-quality grasps to select the most desirable target. For evaluating grasp quality offline, we combine the quality measures generated from the simulator GraspIt! [12] with our own measure of grasp robustness, which we define as the insensitivity of the grasp to slight positional errors.

Robustness has been discussed in [13] by examining the effect of rotational variation on grasp quality, although in general this topic is not widely addressed in the grasping literature, and is part of this paper's contribution.

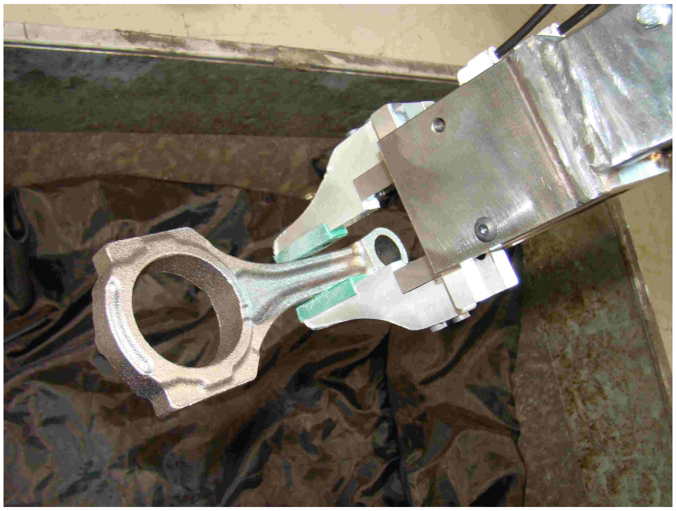

Fig. 1. Standard industrial two-fingered gripper using a nominal grasp to grip a con-rod [14].

For the purposes of illustrating the proposed method in subsequent sections, we have chosen a connecting rod, or con-rod (a common automotive part) as our exemplar part. This part is of a typical size and form of parts that would be suitable for bin-picking applications. Other parts to consider include screws, shafts, and caps, as they are simple in shape and typically delivered to the assembly line jumbled in bins.

This paper is organized as follows: Section II describes the offline grasp generation and evaluation process, Section III describes the online candidate part evaluation process, Section IV describes the experiment and results, and Section V concludes the paper and discusses future work.

II. High-Quality Grasp Generation

Generating an extensive list of grasps for a given part can be computationally expensive, and is very difficult to compute online within the required time constraints. Typically, in the context of industrial bin-picking, a-priori knowledge of the part to be picked is available. This allows for offline generation and evaluation of grasps with minimal concern for computation time.

This section is broken up into two parts. In part A, the approach for generating an extensive list of grasps for a con-rod is detailed. Part B describes how the quality of each grasp is evaluated.

A. Densely Sampling the Grasp Space

For a standard industrial two-fingered gripper, grasps at multiple positions and orientations are generated by intersecting the space between the gripper fingers with the part at uniform intervals. To reduce the complexity of grasp generation, only planar grasps are considered, i.e., the grasp contact points lie in a plane. This choice enables us to model this space as a bounded 2-D (planar) region located at the gripper fingertips (see Fig. 2). To formally describe this intersection process, we present the following definitions:

$\boldsymbol{\Psi_g}$ - the 2-D region of space between the fully-opened gripper fingers, located at the gripper fingertips (see Fig. 2)
$\boldsymbol{S_p}$ - a collection of line segments that comprise a wire-frame approximating the skeleton of the part (see Fig. 3)
$\boldsymbol{N_{sp}}$ - the number of line segments comprising S_p
$\boldsymbol{s_i}$ - a single line segment within S_p
$\boldsymbol{L_i}$ - the length of s_i
$\boldsymbol{d}$ - linear translation parameter along a line segment
$\boldsymbol{\theta}$ - axial rotation parameter about the z-axis of the current part wire-frame line segment
$\boldsymbol{\phi}$ - current-frame rotation parameter about the pinching (or sliding) direction of the gripper, defined in the plane of Ψ_g; the pinching direction is always perpendicular to the line segment
$\boldsymbol{\Delta d}$ - translational step-size for d
$\boldsymbol{\Delta\theta}$ - rotational step-size for θ
$\boldsymbol{\Delta\phi}$ - rotational step-size for ϕ

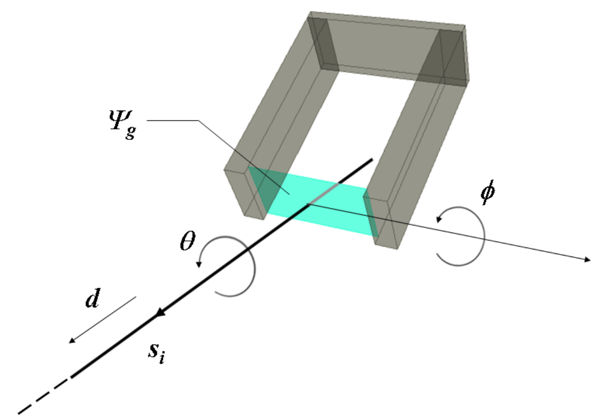

Fig. 2. Illustration of Ψ_g, the gripper, and the sampling directions.

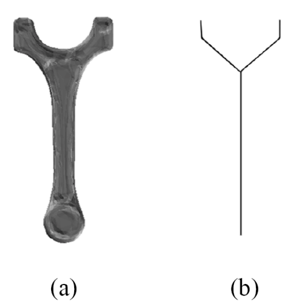

Fig. 3. (a) Part model. (b) Corresponding wire-frame, S_p; $N_{sp} = 5$.

We define the wire-frame S_p manually (see Fig. 3), and restrict the position of Ψ_g to points along S_p. The intersection algorithm is described in Fig. 4; it involves translating the fully-opened gripper (and correspondingly, Ψ_g) in discrete steps along each $s_i \in S_p$, and at each translational step, rotating Ψ_g through a sphere of discrete orientations. At each new position of Ψ_g, the intersection between Ψ_g and the part is computed, resulting in a 2-D cross-section. Grasp points are defined at the extrema of the cross-section along the pinching direction, within a tolerance, ε, to account for soft gripper contacts (see Fig. 5). Only grasps that do not result in a collision between the fully-opened gripper and the part are stored.

Algorithm GRASP_GENERATOR

Let $move(s_i, d, \theta, \phi)$ represent a function that translates and rotates gripper (and correspondingly Ψ_g) to the pose defined by the input parameters

for $i = 1$ to N_{sp}
  for $d = 0$ to L_i; step Δd
    for $\theta = 0$ to $(2\pi - \Delta\theta)$; step $\Delta\theta$
      for $\phi = 0$ to π; step $\Delta\phi$
        $move(s_i, d, \theta, \phi)$
        if fully-opened gripper does not collide with part
          compute intersection between Ψ_g and part
          compute grasp points from this intersection
          store grasp data (contact points + gripper pose)
        end if
      end for
    end for
  end for
end for

Fig. 4. Grasp Generator algorithm, which describes the intersection of Ψ_g with the part.

Parameterizing the rotation with θ and ϕ always ensures that the pinching direction of the gripper is perpendicular to the current wire-frame line segment. The justification for this sampling space is that the generated grasps are (generally) more stable if the forces applied by the gripper fingers are perpendicular to the surfaces they contact, as this minimizes the risk of slippage between the gripper fingers and the part.

The selection of the sampling step size is important to the intersection algorithm. Although a dense sampling is desired, there is a limit on the accuracy of the robot that would be used to grasp the part; it would be superfluous to use a step-size that is smaller than the positional error of the gripper. Thus, we use the robot's positional accuracy as a lower bound on the positional step-size, Δd.

To uniformly sample the grasp space, it is desirable to use a similar step-size in all directions. This is complicated by the fact that one sample direction is translational while the other two are rotational. To address this, we select rotational step-sizes $\Delta\theta$ and $\Delta\phi$ comparable to the translational step size Δd by requiring that the arc length spanned by each rotational step-size at the average radius of the part is equal to Δd.

This algorithm may be used to collect grasps for any object that can be roughly approximated by a wire-frame skeleton of line segments.

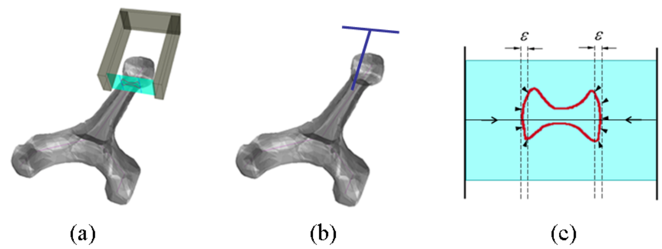

Fig. 5. Illustration of generating a single grasp from a 2-D cross-section. (a) Sample grasp. (b) Minimal representation of grasp using T shape, which depicts approach direction, pinching direction, and position of grasp. (c) Contact points corresponding to sample grasp. Contacts are located at the extrema of the cross-section (indicated by the arrowheads) along the pinching direction of the gripper, within a tolerance, ε.

B. *Grasp Evaluation*

To evaluate grasps, we use the quality measures provided by the simulator GraspIt! [12]. This simulator has been used in [9], [10], [13], and [15] for grasp evaluation. The GraspIt! quality measures are based on the magnitude of the largest disturbance wrench that can be resisted by a unit-strength grasp, as proposed by Ferrari and Canny [16]. Henceforth, for a given grasp, g_i, we will refer to this GraspIt! quality measure as q_i and describe grasps with large values of q as being "highly stable". A grasp is stable if q_i is greater than zero; therefore, we discard any grasps whose quality measure is less than or equal to zero.

Herein, we consider a measure of robustness of a grasp g_i, which we denote as r_i. This is a measure of the insensitivity of q_i to small variations in the position of the grasp. Robustness is important to consider for VGRBP because of the gripper position error as well as the pose estimation error of the target part, i.e., the actual grasp is likely to be offset from the desired grasp. We propose that the robustness, r_i, of a grasp, g_i, is the the inverse of the standard deviation of q within a local region, ρ, centered on g_i. Thus, we present the following definition:

$$r_i = \frac{1}{\sqrt{\frac{\sum^{N_\rho} (q_j - \bar{q})^2}{N_\rho - 1}}} \tag{1}$$

Here, N_ρ is the number of grasps within the local region ρ of the grasp in question, and $\bar{q}$ is the mean within this region. The size of the region, ρ, to consider is an input parameter, and is selected based on the position accuracy of the robotic system.

A feasible, stable grasp is considered to be robust (and is, therefore, accepted) if all neighbouring grasps: (1) exist, (2) are feasible (i.e. they will not result in collisions with the gripper), and (3) are stable ($q_i > 0$).

Finally, we propose the following definition for the overall quality measure, Q_i, of a grasp g_i:

$$Q_i = q_i^{\alpha} \cdot r_i \tag{2}$$

where q_i and r_i have each been normalized between 0 and 1 using the maximum value for each from their respective data sets, and α is a tunable "stability" parameter that is greater than 1 in order emphasize grasp stability over robustness. For the remainder of this paper, when we use the term "quality", we are referring to Q.

Equation (2) ensures that the best grasps are those that both resist large disturbance wrenches and are insensitive to slight position changes. The factors are multiplied rather than weighted and summed, since grasp quality depends on both factors simultaneously rather than either factor independently.

III. Evaluation of Candidate Parts

In VGRBP, a 3-D vision system is typically used to obtain a topographical map of the pile surface, providing information for part localization and obstacle avoidance. In our approach, each localized candidate part is evaluated based on how many pre-generated grasps are collision-free in the context of the pile, using information about neighbouring parts and obstacles obtained from the vision system. A candidate part is considered to be a valid picking option if there exists as least one robust collision-free grasp for picking it up. For the purpose of performing statistical trials to test our approach (detailed in Section IV), we perform our evaluation in simulation, for which we have complete knowledge of all obstacles in the pile.

The process of evaluating candidate parts is performed online, and is described below:

1) Select a set of candidate parts to pick from the pile.
2) For each candidate part:
    a) Obtain the transformation that describes the part's pose in the world co-ordinate frame.
    b) Apply this transformation to each potential grasp (which describes the gripper's pose) and check for collisions between the gripper and all obstacles.
    c) Tally the collision-free grasps. If no collision-free grasps exist, eliminate candidate.
3) Rate the remaining candidates based on the number of available grasps for each, and return this rating, along with each candidate's list of available grasps, to the robot control system.

If only the part, the gripper, and the pile configuration were considered, the best picking option would be the one that provides the most available grasps in the context of the pile. However, some grasps may be impossible due to robot joint limits and workspace constraints. An additional step is then required to process the rated candidate list to check for feasibility with the robot's limits before finally selecting the highest-quality feasible grasp for the highest-rated candidate.

The generated grasp list contains only robust grasps. However, due to the limit on online computation time, we further reduce this list to a set of the highest-quality grasps when evaluating candidates. This results in many good grasping options for the system, and ideally increases system reliability.

IV. Experiment and Results

The parameters that we used for grasp generation for a con-rod are summarized in Table I. Our sampling step-size was chosen to be 3mm. The region size for robustness calculations was restricted to neighbours within one step-size in each of the three directions (d, θ, and ϕ). This can be visualized as a $3 \times 3 \times 3$ array of grasp samples. We selected these values based on what would be typical values for robot accuracy and pose estimation accuracy: $\pm$1mm and $\pm$2mm, respectively. We selected the tunable stability parameter, $\alpha = 2$; future work will investigate the optimal value of α.

Table II summarizes the results of the grasp generation using the parameters shown in Table I. Out of 26650 grasps sampled, 4284 are labeled as robust.

TABLE I
Summary of input parameters used for grasp generation and evaluation.

Grasp generation sampling size (mm)	Soft gripper finger tolerance, ε (mm)	Dimensions of region ρ for robustness calculation	α
3	3	$3 \times 3 \times 3$	2

TABLE II
Grasp generation results. Percentages are in relation to number of grasp samples.

# of grasp samples	Feasible grasps		Stable and feasible grasps		Robust Grasps	
	#	%	#	%	#	%
26650	18041	67.7	15378	57.7	4284	16.1

Fig. 6 visualizes these grasps with respect to the con-rod model from different viewing directions of the model. We have chosen to visually represent each grasp using a T shape, which is to be interpreted as follows:

- The location of the T along the wireframe represents the position of the grasp (as described by d).
- The stem of the T represents the approach direction of the gripper.
- The top bar of the T represents the pinching direction of the gripper.
- The size of the T represents the quality, Q, of the grasp; it has been uniformly scaled according to Q.

The grasp qualities depicted in Fig. 6 are consistent with what one would expect: high quality grasps tend to be those for which (a) there exist many points of contact between the gripper fingers and the part, and (b) forces applied at the gripper finger contacts are generally normal to the part's surface. As expected, the best grasps are clustered near the centre of mass of the part, where disturbance torque is minimized, and few good grasps are found in regions of high surface curvature.

Since grasp stability, q, is dependent on the number of contacts between the gripper fingers and the part, the final quality, Q, is sensitive to the deformability of the gripper fingertips (modeled in our system as ε), as well as imperfections in the surface model of the part. The grasps depicted in Fig. 6 are not perfectly symmetrical because the con-rod model used was obtained from a laser scan of a physical con-rod.

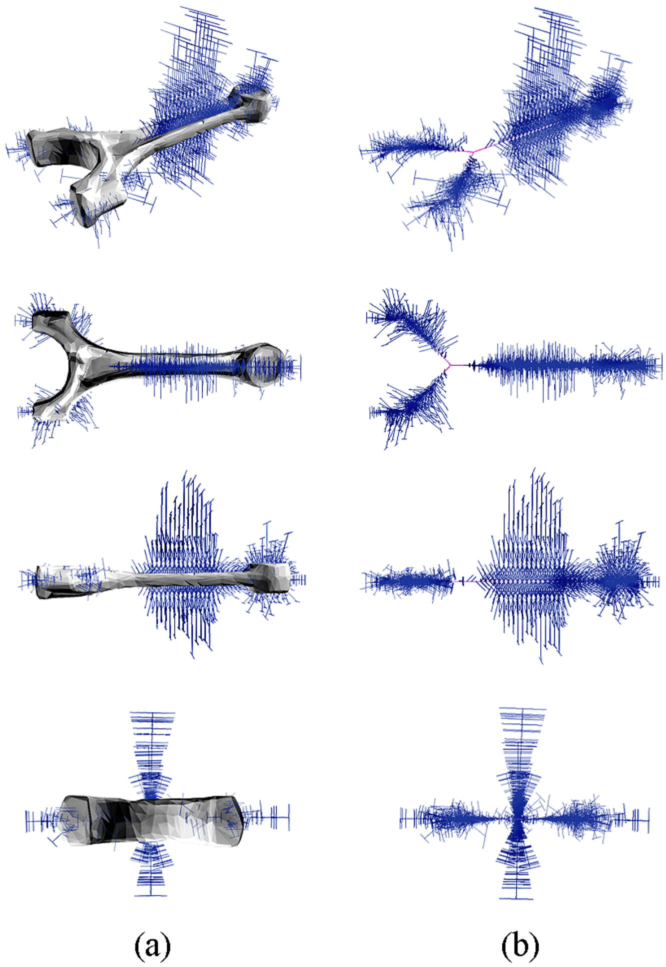

Fig. 6. Visualization from different viewing directions of uniformly-spaced, densely-sampled list of generated grasps with respect to wire-frame, S_p. Each grasp is minimally represented using T shape to indicate both the approach and pinching directions for that grasp. Each T is scaled according to the corresponding grasps computed Q value. Only robust grasps are shown. In (a), part model is overlaid onto wire-frame. In (b), just the wire-frame is shown.

In our experiment, we evaluated candidates within a simulated randomized pile of con-rods using two sets of grasps: (1) a set of top quality grasps from our generated grasp list, $\{G\}$, and (2) a set of 6 nominal "intuitive" grasps, $\{N\}$, that would typically be used in an industrial application. For the first set, grasps were ranked based on Q, and the top $\mu = 10$ were selected, although all grasps could potentially have been included since all are robust. This quantity, μ, is a tunable parameter, and optimizing this value depends on the quality of grasps generated, as well as limits on online computation time. The set of nominal grasps is illustrated in Fig. 7. For each potential grasp, we checked for collisions with the ground plane and all other parts in the pile using the efficient hierarchical Oriented Bounding Box method described in [17].

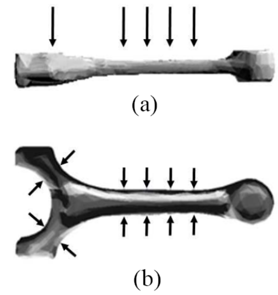

Fig. 7. Visual description of the 6 nominal grasps used for experiments. (a) side-view of part; arrows represent approach directions. (b) top-view of part; arrows represent pinching directions.

We performed this evaluation with 100 different piles of 25 parts; for each trial, we selected the last 15 parts that had been added to the pile as our candidate picks in order to approximate the real-world situation wherein the candidates would be at or near the surface of the pile. Table III summarizes the input parameters for the experiment. Fig. 8(a) shows an example of a pile of parts used in our experiment, with the candidates highlighted in Fig. 8(b). Valid picking options are highlighted and numbered according to their rating in Fig. 8(c)-(d) for the grasp sets $\{G\}$ and $\{N\}$, respectively.

The average number of valid picking options for the set of top grasps, $\{G\}$, and the set of nominal grasps, $\{N\}$, were 8 and 5, respectively. A paired t-test analysis of the null hypothesis that these two methods produce the same distribution of valid parts for picking had a probability of 7.55×10^{-25}, indicating that the distributions are significantly different. These results are summarized in Table IV, and confirm the alternate hypothesis that increasing the number of possible grasps for the part results in an increased number of valid picking options. However, the proposed evaluation does not consider whether or not candidates are pinned down by other parts, and if so, the extent to which they are buried. One would expect that a candidate for which there is an available grasp in the context of the pile, but is deeply embedded in the pile, would be a poor option, and should be eliminated. An example of this situation is illusrated in Fig. 8(c) for the candidates rated 5th and 7th. Future work aims to address this issue.

TABLE III

SUMMARY OF INPUT PARAMETERS USED FOR EXPERIMENT.

# of piles	# of parts per pile	# of candidates per pile	Percentage, μ, of top robust grasps used	Resulting number of grasps used
100	25	15	10%	428

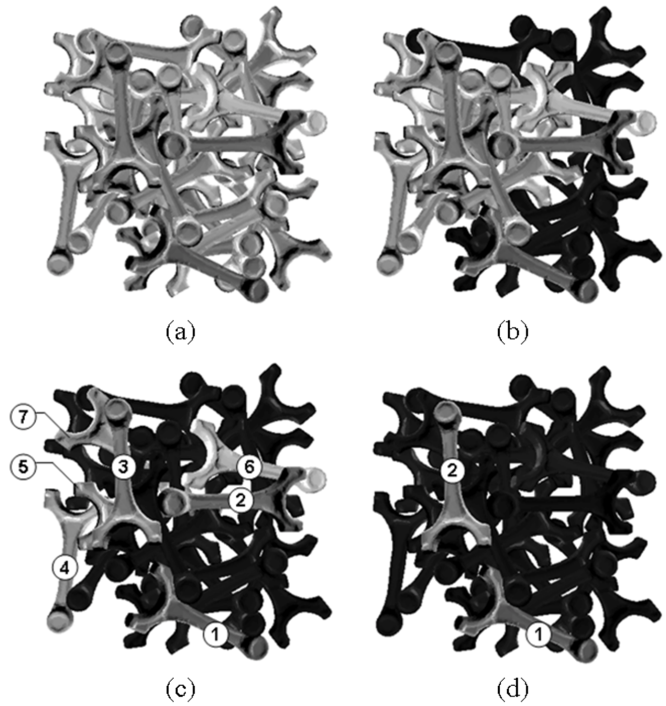

Fig. 8. Comparison of the sets of valid picking options determined for an example pile for grasp sets $\{G\}$ and $\{N\}$. (a) The simulated pile of parts. (b) Highlighted candidates. (c) Highlighted valid picking options found using $\{G\}$, numbered according to their rating. (d) Highlighted valid picking options found using $\{N\}$, numbered according to their rating.

TABLE IV

SUMMARY OF STATISTICAL RESULTS.

Average # of parts with at least one valid grasp		Probability that distributions are the same (paired t-test)
Generated grasps $\{G\}$	Nominal Grasps $\{N\}$	
8	5	7.55×10^{-25}

In addition to using this evaluated densely-sampled set of generated grasps to provide many grasping options online, we can also use this data to establish high quality grasp regions, enabling the selection of nominal grasps offline.

V. CONCLUSION

The main contributions of this paper include a novel method for densely sampling the grasp-space of an object using a two-fingered gripper, a method for evaluating grasp robustness, and a new grasp quality measure. We have presented a way to use the evaluated list of generated grasps in the context of VGRBP to (1) increase the number of pickable candidate parts, and (2) select the best part to pick. Our simulated experimental results show that our approach increases the average number of pickable candidates when compared with a standard industrial approach, and leads to a reliable bin-picking system.

In future work, we will test our method against stereo maps generated by real pile surfaces, and test our evaluation in the context of a physical bin picking experiment. Other future work includes optimization of input parameters, as well as investigating additional factors that affect successful part picking and how to include these factors in the evaluation of candidate parts.

ACKNOWLEDGMENT

The authors gratefully acknowledge the contribution of the National Sciences and Engineering Research Council of Canada, Precarn Inc., Braintech Inc., The Institute for Computing, Information, and Cognitive Systems, and reviewers' comments.

REFERENCES

[1] L. Perreault and P. Olivier, "Bin-picking system for randomly positioned objects," U.S. Patent 7 313 464, Dec. 25, 2007.

[2] E. Al-Hujazi and A. Sood, "Range Image Segmentation with Applications to Robot Bin-Picking Using Vacuum Gripper," *IEEE Trans. SMC*, vol. 20, no. 6, pp. 1,313-1,325, 1990.

[3] O. Ghita and P. F. Whelan, "A bin picking system based on depth from defocus," *Machine Vision and Applications J.*, vol. 13, no. 4, pp. 234-244, 2003.

[4] J. Kirkegaard and T. B. Moeslund, "Bin-Picking based on Harmonic Shape Contexts and Graph-Based Matching," *Intl. Conf. on Pattern Recognition*, vol. 2, pp. 581-584, 2006.

[5] K. Ikeuchi, "Generating an interpretation tree from a CAD model for 3-D object recognition in bin-picking tasks," *Intl. J. Comput. Vision*, vol. 1, no. 2, pp. 145-165, 1987.

[6] A. Zuo, J. Z. Zhang, K. Stanley, and Q. M. J. Wu, "A Hybrid Stereo Feature Matching Algorithm for Stereo Vision-Based Bin Picking," *Intl. J. of Pattern Recognition and Artificial Intelligence*, vol. 18, no. 8, pp. 1407-1422. Dec. 2004

[7] H. S. Yang and A. C. Kak, "Determination of the identity, position and orientation of the topmost object in a pile," *Comput. Vision, Graphics, Image Processing*, vol. 36, pp. 229-255, 1986.

[8] I. Fryndendal and R. Jones, "Segmentation of sugar beets using image and graph processing," *Intl. Conf. on Pattern Recognition*, vol. 2, pp. 1697-1699, 1998.

[9] A. Miller, S. Knoop, P. Allen, H. Christensen, "Automatic Grasp Planning Using Shape Primitives," *IEEE ICRA*, pp. 1824-29, Sep. 2003.

[10] C. Goldfeder, P. Allen, C. Lackner, and R. Pelossof, "Grasp planning via decomposition trees," *IEEE Intl. Conf. on Robotics and Automation*, Rome, Italy, April 2007.

[11] P. J. Sanz, A. Requena, J. M. Inesta, and A. P. Del Pobil, "Grasping the not-so-obvious: vision-based object handling for industrial applications," *IEEE Robotics and Automation Magazine*, vol. 12, no. 3, Sept. 2005, pp. 44-52.

[12] A. Miller, P. Allen, "Graspit!: A Versatile Simulator for Robotic Grasping," *IEEE Robotics and Automation Magazine*, vol. 11, no. 4, Dec. 2004, pp. 110-122.

[13] S. Ekvall and D. Kragic, "Learning and Evaluation of the Approach Vector for Automatic Grasp Generation and Planning," *IEEE Intl. Conf. on Robotics and Automation*, Rome, Italy, April 2007.

[14] J. Beis (Braintech Inc.). Personal communication, May 2008.

[15] A. Morales, T. Asfour, P. Azad, S. Knoop, and R. Dillmann, "Integrated Grasp Planning and Visual Object Localization For a Humanoid Robot with Five-Fingered Hands," *International Conference on Intelligent Robots and Systems (IROS)*, Beijing, China, 2006.

[16] C. Ferrari, J. Canny, "Planning Optimal Grasps," *IEEE ICRA*, pp. 2290-95, 1992.

[17] S. Gottschalk, M.C. Lin and D. Manocha, "OBBTree: A Hierarchical Structure for Rapid Interference Detection," *Computer Graphics (SIGGRAPH '96 Proc.)*, Vol. 30, pp. 171-180, New Orleans, Aug. 1996.