A Practical and Intuitive Quality Measure for Evaluating Antipodal Grasp Poses in The Image Space

Tian Tan, Redwan Alqasemi, Rajiv Dubey, Sudeep Sarkar

***Abstract*— Practical robot applications that require grasping often fail due to failed grasp planning, and a good grasp quality measure is the key to successful grasp planning. In this paper, we developed a novel grasp quality measure that quantifies and evaluates grasp quality in real-time. To quantify the grasp quality, we compute a set of object movement features from analyzing the interaction between the gripper and the object's projections in the image space. The normalizations and weights of the features are tuned to make practical and intuitive grasp quality predictions. To evaluate our grasp quality measure, we conducted a real robot grasping experiment with 1000 robot grasp trials on ten household objects to examine the relationship between our grasp scores and the actual robot grasping results. The results show that the average grasp success rate increases, and the average amount of undesired object movement decrease as the calculated grasp score increases, which validates our quality measure. We achieved a 100% grasp success rate from 100 grasps of the ten objects when using our grasp quality measure in planning top quality grasps.**

I. INTRODUCTION

Generating grasp pose candidates and evaluating their qualities are core problems in grasp planning. In this work, we focus on addressing the grasp pose evaluation problem in the image space. Existing approaches to this problem can be divided into two categories, analytical [1]-[3] and data-driven [4][5]. The analytical approaches extract various hand-crafted features from analyzing the input sensory data and calculating an overall quality score. On the other hand, data-driven approaches use auto-generated features. The grasp quality can be found by either feature combination or comparison with known-quality grasps in the feature space. This type of approach requires extensive training data or a large reference database.

Our grasp evaluation method design is application-driven. We consider assistive robots in grasping household items with parallel jaw grippers. Assistive robots designed for the old and individuals with disabilities must be reliable. Due to the limitations of robots on perceiving and decision making, the most practical and reliable control strategy is a semi-autonomous or shared control, which combines the superior human perceptive and cognitive abilities with the motor functions and basic sensing and reasoning capabilities of robots. We designed a two-stage semi-autonomous system to simplify the six-dimensional grasp planning problem and make the task allocation between the human user and the autonomous agent more natural. In the first stage of this system, the object position is found using a depth camera, and the gripper is positioned at a certain distance from the object. With the gripper's z-axis always pointing towards the object, the gripper pitch and yaw are controlled by an autonomous grasp planner and the human user in a shared control mode. The goal of controlling the gripper pitch and yaw is to set the gripper (with a hand-camera) in a proper grasp direction. So, in the second stage, we can take advantage of the parallel jaw gripper's grasping mechanism, which allows us to reduce the 3D grasp evaluation problem to 2D. The object's boundary in the 2D image can represent the object's potential contact region when it is grasped from the direction where the image is taken. Therefore, we can perform grasp-planning on a 2D silhouette image of the object taken from a proper grasp direction. When projecting the 3D object to a 2D silhouette, the change of the object's cross-sections in the projection direction should be considered as in [[6], Fig.6].

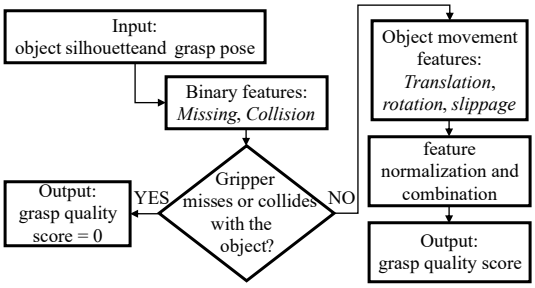

Fig. 1. Grasp quality evaluation flowchart.

This paper presents our analytical grasp evaluation method for evaluating grasp poses in the image space. We designed eight grasp features, consisting of two low-level visual features that capture the grasp region's geometry properties and six high-level quality features that measure the object movement during gripper closing. The overall quality score is calculated as the weighted sum of the normalized feature values. Fig. 1 shows the flowchart of the evaluation process.

II. RELATED WORK AND BACKGROUND

The existing approaches for grasp evaluation include data-driven and analytical approaches. The most successful data-driven approaches are deep learning approaches. The most popular deep learning approaches [7]-[10] can achieve 80-100% grasp success rate in grasping common household items. However, the grasp success rate alone cannot validate the grasp network's auto-generated grasp quality measure since there can be a significant quality difference between successful grasp poses. For example, grasps A and B both picked up a toothbrush successfully, but grasp A did not

*This work was supported by NSF

move the toothbrush during gripper closing, while grasp B rotated the toothbrush 60 degrees. If using only success rate to judge the grasp quality, then grasp A and B are the same quality, but grasp A is apparently better than grasp B. The main problem of the grasp networks is that even they can achieve high grasp success rates, the grasp poses' actual qualities are unpredictable. Thus, grasp networks are suitable for such tasks as bin picking and storing/retrieving non-fragile goods, but not reliable enough to be used alone on assistive robots.

There are two sub-categories of analytical approaches, the contact geometry-based, and the contact wrench-based approaches. The contact geometry-based approaches typically evaluate grasp quality by examining if the contact region's geometric properties/features are consistent with the designed criteria. Davidson [11] and Vahedi [12] presented geometric caging methods for finding immobilizing grasps. Calli [13] used the object contour curvature derived from its contour function to help determine graspable regions on the object contour. Blake [14] exploited the properties of object contour local reflectional and rotational symmetry. Due to the lack of good feature designs, the contact geometry-based approaches often have many limitations.

The contact wrench-based approaches judge the grasp quality by analyzing the contact wrenches (forces and torques) acting on the object [2]. Most of the contact wrench based methods are built upon the grasp force closure property [15]. Nguyen [16] presented a method to construct force closure grasps, which is one of the representative works in the field. Ferrari and Canny [17] developed the most popular grasp quality measure(often referred to as Q measure) for force closure grasps. The Q measure is a measure of how well a grasp can resist external disturbances. It is currently the most versatile and reliable analytical grasp evaluation method. Many variants of the Q measure were developed [18]-[21] to improve the method in different aspects. Other than the Q measure, other contact wrench-based approaches [22] often have many restrictions similar to the geometry-based approaches.

Compared to deep learning-based grasp quality measures, our quality measure is more practical and reliable. Comparing to the Q measure, the advantages of our quality measure are: (1) our quality score is normalized and can be tuned to approximate the grasp success rate, while the quality score of the Q measure is a distance measure in the grasp wrench space, which makes our quality scoring more intuitive than the Q measure; (2) our method can evaluate grasps with all types of contact conditions while the Q measure only works for grasps with non-slippery contact (the grasp force has to be inside the friction cone); (3) our method is less sensitive to imperfect object modeling than the Q measure, making our method work better with real camera data.

III. Grasp Quality Evaluation

A. Grasp Representation

To evaluate the gripper-object interaction, we established mathematical models for both the gripper and the object projections in the image. An object's projection is a silhouette image of the object, which is mathematically a binary matrix. As for the gripper projection, we designed a different grasp representation in the image space. Instead of describing an antipodal grasp as a rectangle, we treat it as a set of line segments, $\{gl_1, gl_2, \ldots, gl_n\}$, as shown in Fig. 2. Therefore, the gripper projection can be discretely modeled as a set of line segment functions. Let point P_{ij} be the j^{th} point on grasp line *i* (gl_i), then the function of gl_i can be expressed as the coordinates of $P_{ij}(P_{ij}^x, P_{ij}^y)$ in the image frame {I}:

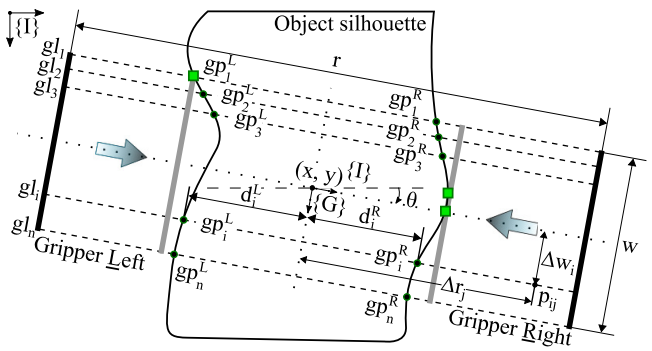

Fig. 2. Illustration of grasp lines(gl_i), grasp points(gp_i^e), grasp distances(d_i^e), contact points and grasp pose parameters(x, y, θ, r, w, n). The thicker lines labeled as L and R represent the left and right gripper fingers' contact lines. The arrow indicates the gripper closing direction.

$$P_{ij}^x = x - \Delta w_i \sin\theta + \Delta r_j \cos\theta \tag{1}$$

$$P_{ij}^y = y + \Delta w_i \cos\theta + \Delta r_j \sin\theta \tag{2}$$

$$\Delta w_i = \frac{w(2i - n - 1)}{2(n-1)} \tag{3}$$

Where $i = 1, 2, \ldots, n$ is the index of a grasp line. Δw_i is the distance from the center of the gripper projection to the i^{th} grasp line. Δr_j is the distance from a grasp line's center to its point j. Also, (Δr_j, Δw_i) is the coordinate of P_{ij} in the gripper projection frame {G}. Δr_j is positive if P_{ij} is on the right side of the grasp line center, and negative if otherwise ($\Delta r_j \in (-\frac{r}{2}, \frac{r}{2})$). The parameters of this grasp representation, (x, y, θ, w, r, n), are composed of the planar center location (x, y), planar orientation (θ), grasp width(w), grasp range(r), and the number of grasp lines ($n \in [1, w]$) of the gripper projection in the image coordinates.

B. Terminologies

Here we introduce some terminologies and annotations to help describe the gripper-object interaction and better explain our algorithm. We use superscripts L and R to differentiate the gripper's left and right sides in our equations. We use superscript e to indicate that the term is side specific, and it should be replaced by either L or R. The intersections of the grasp lines and the object outline are referred to as *grasp points* of the object, such as gp_i^e ($i = 1, 2, ..., n$) in Fig. 2. The distances between grasp points and the center points of the corresponding grasp line segments are referred to as the *grasp distances* (d_i^e) of those grasp points. When the gripper closes along the grasp line, some of the grasp

points will contact the gripper, and these grasp points become *contact points*. The green square shape grasp points in Fig. 2 are the contact points corresponding to that grasp pose. The collection of grasp points is the *grasp region*, and the collection of contact points is the *contact region*.

C. Low-level Visual Features

The visual features are the grasp region profile vector (GRPV) and the contact region feature vector (CRFV), both of which carry information about the geometry and the contact properties of the object contour inside the grasp region. The GRPV is a vector of grasp distances of all grasp lines.

$$GRPV = \begin{bmatrix} GRPV^L \\ GRPV^R \end{bmatrix} = \begin{bmatrix} d_1^L & d_2^L & \dots & d_n^L \\ d_1^R & d_2^R & \dots & d_n^R \end{bmatrix} \quad (4)$$

The grasp distances can be found from searching the intersections of the object outline and the grasp lines. If there are no intersections between the grasp line and the object silhouette, then the left and right grasp distances are set to $-r$ and r, respectively. If the endpoint of one side of the grasp line segment is inside the object silhouette, then the grasp distance of that side is $-2r$ (left) or $3r$ (right). These abnormal grasp distance values are used to indicate gripper misses or collides with the object.

After calculating the GRPV, the left and right *primary contact points* can be defined as the grasp points with maximum absolute grasp distances (excluding abnormal grasp distances) in the left and right grasp regions, respectively. If the grasp distances' difference between a grasp point and the primary contact point of the same side is less than a threshold, this grasp point is a *secondary contact point* on that side. Adding secondary contact points into consideration decreases the effect of false contact prediction caused by imperfect object silhouette detection. Once the contact points are found, we calculate their normals and center offsets and organize them in the CRFV:

$$CRFV = \begin{bmatrix} CRFV^L \\ CRFV^R \end{bmatrix} = \begin{bmatrix} CPN^L \\ CPD^L \\ CPN^R \\ CPD^R \end{bmatrix} \quad (5)$$

CPN^e and CPD^e are the vectors of contact point normals and contact point center offsets, respectively. The contact point center offset equals to the distance from the gripper center to the grasp line that the contact point is on, which is the Δw_i in equation (3).

D. High-level Quality Features

Based on the two low-level features, we derived a set of more in-depth features that can be used to quantify the quality of the grasp. These high-level features are the missing detector, the collision detector, the object translation predictor, the object rotation predictor, the type 1 and type 2 object slippage predictors. *The missing detector (D_m)* and *the collision detector (D_c)* are binary terms that indicate if the gripper will miss the object and if the gripper will collide with the object, respectively. As shown in Fig. 3,

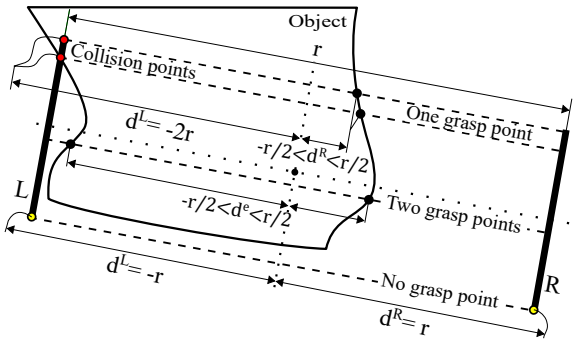

Fig. 3. Features that relate to *missing* and *collision*.

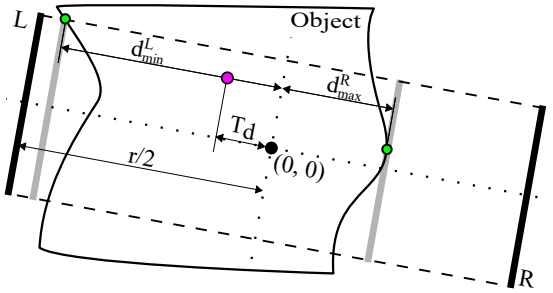

Fig. 4. Feature that relates to object *translation* in the gripper closing direction.

when a grasp line misses the object, there is no grasp point correspond to the grasp line, and the grasp distances are abnormal values. Similarly, when the gripper collides with the object, the colliding side's grasp distance is also abnormal. Therefore, using these abnormal grasp distances, we can calculate D_m and D_c from the GRPV:

$$D_m = \begin{cases} 1 & \Delta d_{max} = 0 \; and \; d^L_{max} = -r \\ 0 & Otherwise \end{cases} \quad (6)$$

$$D_c = \begin{cases} 1 & \Delta d_{max} \geq \frac{r}{2} \\ 0 & 0 \leq \Delta d_{max} < \frac{r}{2} \end{cases} \quad (7)$$

$$\Delta d_{max} = max(\big|\,|\,GRPV^L\,| - |\,GRPV^R\,|\,\big|) \quad (8)$$

$$d^L_{max} = max(GRPV^L) \quad (9)$$

When D_m is 1, the gripper misses the object, and when D_c is 1, there is a collision.

The object translation predictor (D_t) estimates the ratio of the object travel distance during gripper closing to half of the grasp range ($\frac{r}{2}$). As shown in Fig. 4, T_d is the object translate distance in the gripper closing direction. It is measured from the center of the two primary contact points to the grasp rectangle's center. Using GRPV, the ratio of the object translate distance to the gripper half grasp range is calculated as:

$$D_t = \left| \frac{min(GRPV^L) + max(GRPV^R)}{r} \right| \quad (10)$$

The minimum and maximum operations here only apply to normal grasp distance values. We use the distance ratio

instead of the distance as the grasp quality measure because using ratio allows this feature to adapt to objects and grippers of different sizes.

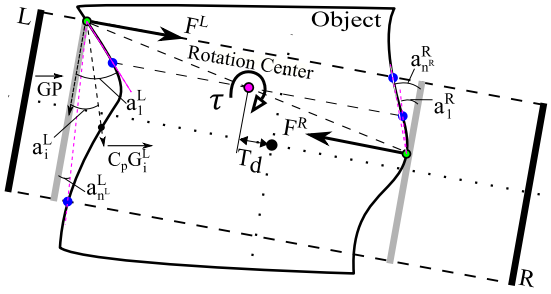

Fig. 5. Feature that relates to object planar *rotation* during grasping.

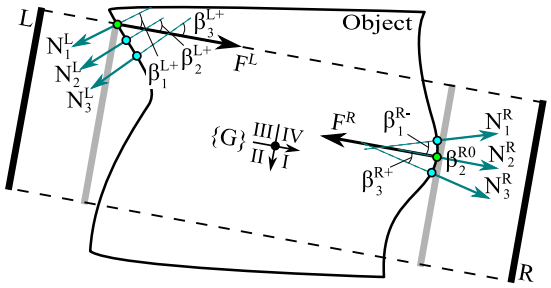

Fig. 6. Feature that relates to contact angle and indicates the likelihood of slippage.

The object rotation predictor (R_r) predicts the angle of object rotation during grasping. The object rotates when there is torque, and the torque occurs when the two primary contact points are not on the same grasp line. With the two primary contact points, the rotation center and the rotation direction can be determined. The *potential after-rotation contact region (PARCR)* is defined using the center of rotation and the rotation direction.

As shown in Fig. 5, the PARCR on each side is the object outline segment between the two grasp points highlighted in blue. One grasp point is on the same grasp line as the rotation center, and the other is the last grasp point of the region that is approaching the gripper in the rotation process. Once the PARCRs are determined, the angle of rotation can be found by selecting the minimum required rotation angle for a grasp point in the PARCR to be the new contact point after rotation. Geometrically, the rotation angle (a_i) is defined as the angle between the gripper plate and the line that connects the primary contact point (approximate pivot point) and the new after-rotation contact point. This can mathematically be expressed as:

$$R_r = min(a_1^L, a_2^L, \ldots, a_{n^L}^L, a_1^R, a_2^R, \ldots, a_{n^R}^R) \tag{11}$$

Where a_i^e's are the rotation angles. Each rotation angle is calculated as the angle between the gripper plate vector $\overrightarrow{GP}$ and $\overrightarrow{C_pG_i^e}$ the vector from the primary contact point to the i^{th} grasp point in the PARCR.

The type 1 slippage predictor (S_a) is a feature that measures how likely the object will slip during grasping due to the slope of the contact surfaces. This feature is defined as the average *contact angle* of the most slippery contact region.

As shown in Fig. 6, the contact angle ($\beta \in [-90, 90]$) is defined as the angle between the contact force ($\mathbf{F}$) and the contact point's inward normal ($-\mathbf{N}_i$). If the contact point normal ($\mathbf{N}_i$) points towards the 1^{st} and 2^{nd} quadrant of the gripper frame {G}, then the contact angle is positive; and it is negative otherwise. If one contact region's contact angles have the same polarity, then the contact region is slippery, and the slipperiness is indicated by the average of all the contact angles. The larger the average contact angle, the more likely the object will slip. The following equations describe this slippage prediction:

$$\beta_i^e = \begin{cases} \angle(\mathbf{F}^e, -\mathbf{N}_i^e) & \mathbf{N}_{iy}^e > 0 \\ -\angle(\mathbf{F}^e, -\mathbf{N}_i^e) & \mathbf{N}_{iy}^e < 0 \end{cases} \tag{12}$$

$$SI^e = \sum_{i=1}^{n^e} \left|\beta_i^e\right| - \left|\sum_{i=1}^{n^e} \beta_i^e\right| \tag{13}$$

$$S_a = \begin{cases} \frac{1}{2}(\left|\frac{\sum_{i=1}^{n^L} \beta_i^L}{n^L}\right| + \left|\frac{\sum_{i=1}^{n^R} \beta_i^R}{n^R}\right|) & SI^e = 0 \\ 0 & SI^e \neq 0 \end{cases} \tag{14}$$

Where N_{iy}^e is the y coordinate of the normal N_i^e in {G}, and n^e is the number of contact points of the e side of the gripper. SI^e is the slippage indicator. When SI^e is 0, slippage is likely to happen because all contact angles of the same side have the same polarity.

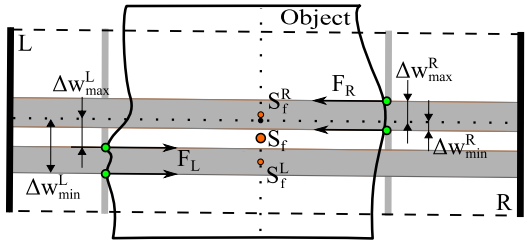

Fig. 7. Feature relates to grasp force location and indicates the likelihood of slippage. The shaded areas are the force zones.

The type 2 slippage predictor (S_f) is a feature that predicts slippage through the grasp force placement. This feature favors grasp forces that are balanced and centered on the gripper contact surface. It is calculated as the average of the left and right force zone center offsets.

As shown in Fig. 7, the force zone of one side of the gripper is the continuous region that contains all contact points of that side. The center offset is the distance from the center of the force zone to the gripper's center. Using CPD^e, this feature can be calculated as:

$$S_f^e = \frac{min(CPD^e) + max(CPD^e)}{2} \tag{15}$$

$$S_f = \frac{S_f^L + S_f^R}{2} \tag{16}$$

Where $min(CPD^e)$ and $max(CPD^e)$ are the Δw^e_{min} and Δw^e_{max} respectively. Note that we do not directly consider friction force in predicting slippage. However, our slippage predictors are related to how the gripper forces are applied to the object. Minimizing the slippage predictors' values maximizes the friction forces applied to the object.

E. Grasp Quality Scoring

The quality features are in different units and scales. Before combining them, we first normalize them with linear functions defined by two endpoints (0, 1) and (τ_i, 0), where x is the feature value and y is the normalized feature score. τ_i is the threshold feature value for the 0 feature score. Since D_t is a ratio, its threshold value is 1. R_r and S_a are measures of angles, and their thresholds are set to 60 degrees based on our experience. S_f is the force placement feature, its value range from 0 to the gripper projection's half-width ($\frac{w}{2}$). We used two feature thresholds $\frac{w}{4}$ and $\frac{w}{2}$ for feature scores 0.7 and 0, respectively, which makes the function has two different slopes when normalizing low and high feature values. We found this yields more practical quality scores than a single slope function. After feature normalization, we calculate the grasp quality measure S as the weighted sum of the feature scores:

$$S = ks_{min} + (1-k)s_{o^-} \tag{17}$$

Where s_{min} is the minimum feature score, and s_{o^-} is the average of other feature scores. Since the grasp pose's quality mainly depends on its worst quality feature, the weight k should be assigned in a minimum-dominant fashion ($k >> 0.5$), and it can be empirically determined by specifying the grasp quality to the desired value when $s_{min} = 0$ and $s_{o^-} = 1$ (in this paper we used $k = 0.9$).

IV. REAL ROBOT GRASPING EXPERIMENT

A. Experiment Setup

We used the Baxter robot [24] from Rethink Robotics as our test platform. The computations were performed on a laptop PC running Ubuntu 18.04 with a 2.2 GHz Intel Core i7-8750H CPU, 8 GB of RAM, and an NVIDIA GeForce GTX 1060 (laptop) graphics card. The graphics card is only used for computing the objects' silhouettes, and other computations are all performed on the CPU. We used a parallel jaw gripper with a pair of narrow fingers(1.3 cm wide). The fingers' rubber contact surfaces are masked by scotch tape to reduce friction so that slipping is more likely to happen when the grasp pose is prone to slippage. Fig. 8 shows the hardware and the 10 objects used in the grasping experiment.

This experiment examines how our calculated grasp quality score relates to the actual robot grasping performance. In the experiment, an eye-to-hand camera (Intel Realsense L515) was used to locate the object. Because controlling the gripper pitch and yaw is outside the scope of this work, we set them as -90 and 0 degrees, respectively, to form top-down grasps. Then from the top of the object, an eye-in-hand camera (Baxter hand camera) was used to take a closer shot at the object. Facebook Detectron2 [23] was used to extract the object's silhouette from the close shot object image. Assuming the gripper is at the object's location, we calculate the gripper's projection in the object silhouette image using calibrated camera intrinsics and the object's distance to the camera. With the object and gripper projections, we randomly generate grasp candidates within the object bounding box and evaluate them with our quality measure to select the one with the desired quality score for robot execution. Randomly generates grasp poses help prove our quality measure's robustness, since to get the desired pose, we need to evaluate hundreds and thousands of candidates.

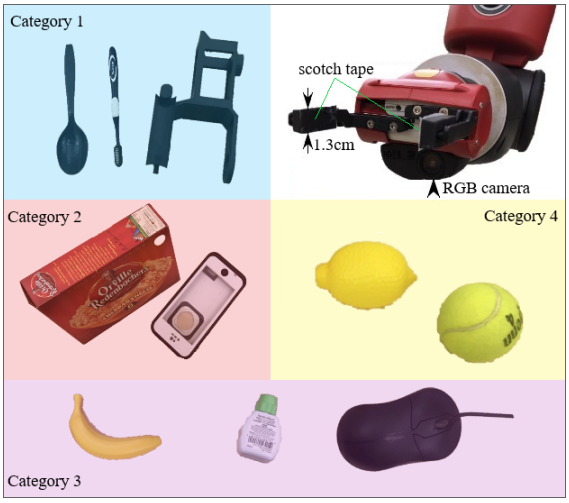

Fig. 8. The robot gripper and the objects used in the robot grasping experiment.Scotch tape was used to reduce the friction of the gripper contact surfaces, so that grasping is prone to failure when the grasp pose is not good enough.

To test the whole grasp score value domain, we categorized the grasp scores into ten score levels, 0-0.1, 0.1-0.2, . . . , 0.9-1. We performed ten grasp trials within each score level for each object, which is 1000 total grasps. For each grasp trial, we record the number of evaluated grasp candidates, the time used in grasp evaluation, and the robot grasp outcome. The robot grasp outcome includes: (1) a binary term that indicates if the grasp is successful, and (2) two hand camera images (BA images) were taken right before and after the gripper closes, which capture the object's movement during gripper closing. If the robot can lift the object for 2 seconds and put it back without visible slippage, then the grasp is recorded as successful. The object's rotation and its translation in the gripper closing direction are measured as the key lines' orientation change and the key points' position change, respectively, in the BA images. The key lines are drawn using the object's appearance features that are visible in both BA images. The position key point is drawn as the object's center on the gripper's center grasp line.

B. Results

In terms of the computational cost, the recorded average grasp generation time for the 1000 grasps is 0.534 s. 261868

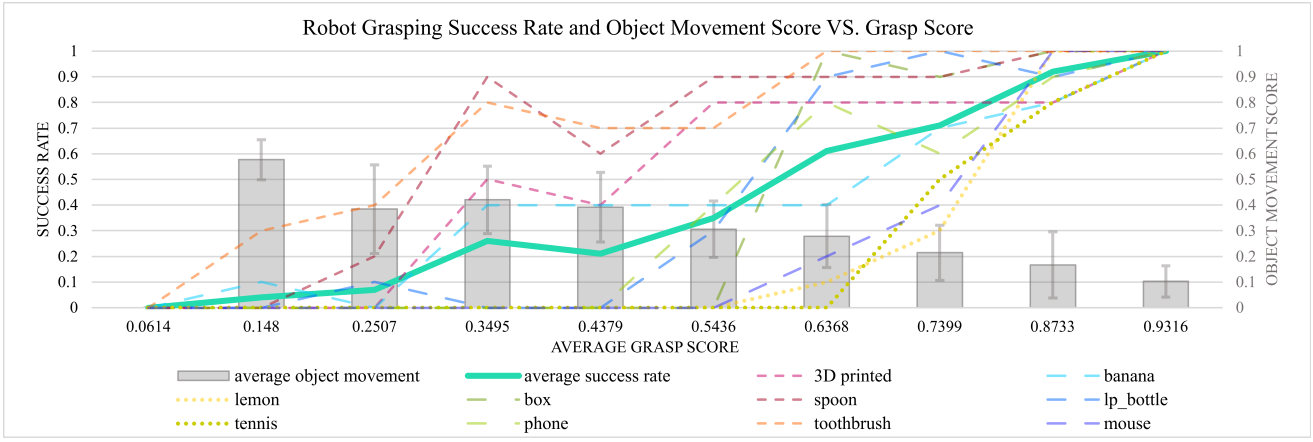

Fig. 9. The success rate and the object movement score of the robot grasping experiment vs. the grasp score.

grasp candidates were evaluated to generate those grasps; the average grasp evaluation time is 2.04 ms, with a standard deviation of 0.264 ms.

To analyze the accuracy of our quality measure, the results of our robot grasping experiment are visualized in Fig. 9. The left y-axis is the success rate; the right y-axis is the normalized object movement score. Because the objects' movements in failed grasp trails are unpredictable and not useful for evaluating our method, we ONLY measure the objects' movements in succeeded grasp trials to reveal the quality difference among those successful grasp poses. The object movement score is calculated as the average of the normalized object rotation and translation (in the gripper closing direction). The thresholds used for normalizing the object translation and rotation are 100 pixels and 60 degrees, respectively. The x-axis is the average grasp score. The level-wise average grasp score of each object is very close to the level-wise average of all objects; the maximum standard deviation of the all objects average is 0.0087 at the 0.6-0.7 score level. The thin non-solid lines show each object's average success rate at each score level. The thick solid line shows the average success rate, and the bars show the average object movement of all objects at each score level. To help interpret the results, we classified our test objects into four categories based on their geometrical properties. As shown in Fig. 8, objects in categories 1 and 2 both have flat contact surfaces. However, the graspable parts of category-1 objects are much thinner than category-2 objects. Objects in category 4 have curved contact surfaces, and category-3 objects have both flat and curved contact surfaces.

From the object-wise success rate, we can see that objects in the same category tend to have similar success rates under the same score level. Also, category-1 objects and category-4 objects have the highest and lowest overall grasp success rate, respectively. The category-1 objects are long and thin and have flat contact surfaces in the projection, making them the easiest for parallel-jaw grippers to grasp, while category-4 objects are the hardest to grasp because they are round and our gripper fingers are narrow. The category-1 objects and the banana have relatively high success rates even when the grasp qualities are low. This result does not disprove our quality measure because the object movements in those grasps are very high. Therefore, even though the grasps were successful, they were evaluated as low quality because they were expected to move the objects a lot.

Despite the differences between different objects' success rates in the medium quality score levels, at the highest and lowest score level, all objects' success rates merge to the same points. This observation indicates that our quality measure can distinguish good and bad grasps regardless of the object type. We achieved a 100% success rate from grasping the ten objects 100 times with grasps of an average score of 0.93 and an average movement score of 0.1 with a standard deviation of 0.0618. Overall, the average grasp success rate increases, and the average object movement score decreases as the quality score increases. All the results show that our grasp quality measure is accurate, practical, and intuitive.

V. CONCLUSIONS

This paper presents the detailed design and experimental results of our novel grasp evaluation method, which calculates grasp poses' quality by analyzing the interactions between the gripper and the object through their projections in the image space. Two low-level geometrical features were directly extracted from the image, and six higher-level grasp quality features were derived from these geometrical features. The overall grasp quality is calculated as the weighted sum of the normalized features. The real robot grasping results show that our grasp quality measure is practical and intuitive. The main limitation of our method is that it only works for parallel jaw grippers. We will develop a versatile object silhouette extraction method and a semi-autonomous grasp direction control system in future works. So that our quality measure based grasp planning can generate grasps in all directions and work on objects of complex shapes.

REFERENCES

[1] K. B. Shimoga, "Robot Grasp Synthesis Algorithms: A Survey," The International Journal of Robotics Research, vol. 15, no. 3, pp. 230–266, Jun. 1996.

[2] A. Bicchi and V. Kumar, “Robotic grasping and contact: a review,” in Proceedings 2000 ICRA. Millennium Conference. IEEE International Conference on Robotics and Automation. Symposia Proceedings, 2000, vol. 1, pp. 348–353.
[3] M. A. Roa and R. Suárez, “Grasp quality measures: review and performance,” Auton Robot, vol. 38, no. 1, pp. 65–88, Jan. 2015.
[4] L. E. Zhang, M. Ciocarlie, and K. Hsiao, “Grasp evaluation with graspable feature matching,” in RSS Workshop on Mobile Manipulation: Learning to Manipulate, 2011.
[5] S. Caldera, A. Rassau, and D. Chai, “Review of Deep Learning Methods in Robotic Grasp Detection,” Multimodal Technologies and Interaction, vol. 2, no. 3, p. 57, Sep. 2018.
[6] Y. Domae, H. Okuda, Y. Taguchi, K. Sumi, and T. Hirai, “Fast graspability evaluation on single depth maps for bin picking with general grippers,” in 2014 IEEE International Conference on Robotics and Automation (ICRA), Hong Kong, China, May 2014, pp. 1997–2004.
[7] J. Mahler et al., “Dex-net 2.0: Deep learning to plan robust grasps with synthetic point clouds and analytic grasp metrics,” in Proc. Robot., Sci., 2017.
[8] J. Mahler, M. Matl, V. Satish, M. Danielczuk, B. DeRose, S. McKinley, and K. Goldberg, "Learning ambidextrous robot grasping policies." Science Robotics 4, no. 26 (2019).
[9] U. Viereck, A. Pas, K. Saenko, and R. Platt, “Learning a visuomotor controller for real world robotic grasping using simulated depth images,” in Conference on Robot Learning, 2017, pp. 291–300.
[10] D. Morrison, P. Corke, and J. Leitner, “Closing the loop for robotic grasping: A real-time, generative grasp synthesis approach”, Robotics: Science and Systems XIV, 2018, pp.1-10.
[11] C. Davidson and A. Blake, “Error-tolerant visual planning of planar grasp,” in Sixth International Conference on Computer Vision (IEEE Cat. No.98CH36271), Bombay, India, 1998, pp. 911–916.
[12] M. Vahedi and A. F. van der Stappen, “Caging Polygons with Two and Three Fingers,” The International Journal of Robotics Research, vol. 27, no. 11–12, pp. 1308–1324, Nov. 2008.
[13] B. Calli, M. Wisse, and P. Jonker, “Grasping of unknown objects via curvature maximization using active vision,” in 2011 IEEE/RSJ International Conference on Intelligent Robots and Systems, 2011, pp. 995–1001.
[14] A. Blake, M. Taylor, and A. Cox, “Grasping visual symmetry,” in 1993 (4th) International Conference on Computer Vision, Berlin, Germany, 1993, pp. 724–733.
[15] X. Markenscoff, L. Ni, and C. H. Papadimitriou, “The Geometry of Grasping,” The International Journal of Robotics Research, vol. 9, no. 1, pp. 61–74, Feb. 1990.
[16] V. Nguyen,“Constructing Force- Closure Grasps” The International Journal of Robotics Research, 1988, 7(3), pp.3-16.
[17] C. Ferrari and J. F. Canny, “Planning optimal grasps.,” in ICRA, 1992, vol. 3, pp. 2290–2295.
[18] F. T. Pokorny and D. Kragic, “Classical grasp quality evaluation: New algorithms and theory,” in 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems, 2013, pp. 3493–3500.
[19] S. Liu and S. Carpin, “A fast algorithm for grasp quality evaluation using the object wrench space,” in 2015 IEEE International Conference on Automation Science and Engineering (CASE), 2015, pp. 558–563.
[20] Y. Zheng, “Computing the best grasp in a discrete point set with wrench-oriented grasp quality measures,” Auton Robot, vol. 43, no. 4, pp. 1041–1062, Apr. 2019.
[21] M. Pozzi, A. M. Sundaram, M. Malvezzi, D. Prattichizzo, and M. A. Roa, “Grasp quality evaluation in underactuated robotic hands,” in 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Daejeon, South Korea, Oct. 2016, pp. 1946–1953.
[22] S. Liu et al., “Grasp quality evaluation and planning for objects with negative curvature,” in 2017 IEEE International Conference on Robotics and Automation (ICRA), 2017, pp. 2223–2229.
[23] Y. Wu and A. Kirillov and F. Massa and W. Lo and R. Girshick, “Detectron2,” https://github.com/facebookresearch/detectron2, 2019
[24] RethinkRobotics Baxter. https://github.com/RethinkRobotics/sdk-docs/wiki/Baxter-Overview