Proceedings of the 38th Chinese Control Conference
July 27-30, 2019, Guangzhou, China

An Unsupervised Grasp Detection for Water-surface Object Collection

Shihan Kong[1,2], Xingyu Chen[1,2], Zhengxing Wu[2], Junzhi Yu[2,3]

1. University of Chinese Academy of Sciences, Beijing 100049, China

2. State Key Laboratory of Management and Control for Complex Systems,
Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China

3. State Key Laboratory for Turbulence and Complex System, Department of Mechanics and Engineering Science, BIC-ESAT,
College of Engineering, Peking University, Beijing 100871, China

E-mail: {kongshihan2016, chenxingyu2015, zhengxing.wu, junzhi.yu}@ia.ac.cn

Abstract: Aquatic environment has been damaged by human activities for years. Trash on the water surface is the prominent problem, e.g., dumped plastic bottles are ubiquitous. To clean water surface by aquatic robots, grasp detection is crucial task in the process of collection. Owing to lack of labeled data for training and the limited computational capacity of aquatic robots, supervised grasp detection method is intractable. In this context, an unsupervised grasp detection strategy is proposed in this paper for bottle grasping. Firstly, the bottle is approximately abstracted as an elongated cylinder, then a grasping model is constructed, in which a graspable position is described by the central point, long axis direction angle, and the maximum width. Secondly, to extract location points of bottles on the water surface, a high speed background elimination method based on RGB histogram is utilized in the gridding bounding box. Next, principal component analysis (PCA) with an outliers removing algorithm is employed to estimate parameters of the graspable position. Finally, experimental results of grasp detection show that estimated position is significant to grasp control. Especially, the accuracy of long axis direction angle is up to 3°, and computation speed of the background elimination method is 1.47 times the speed of Canny.

Key Words: Grasp detection, grasping model, background elimination, principal component analysis

1 Introduction

The oceans and rivers cover almost 71% of the earth's surface and provide a suitable home for billions of aquatic organism. Particularly, marine plants produce about 70% of the oxygen in the earth's atmosphere [1], and the abundant resources in oceans have driven both human civilization progress and social economic development throughout history. Despite the aquatic environmental necessity to both life and society on the earth, humanity does not treat the aquatic environment in a friendly manner.

The water pollution resulted from human negligence has been accumulating for decades. The waste in water consists of dredge, industrial garbage, sewage, and radiation [2], in which plastic trash is the prominent problem. The United Nations Joint Group of Experts on the Scientific Aspects of Marine Pollution estimates that $60-80\%$ of the waste in the oceans is made up of plastic debris [3, 4]. The dumped plastic bottles, as the most common aquatic trash, are here and there in rivers, lakes, and even oceans. For protecting aquatic environment, an aquatic robot equipped with mechanical arm is designed to autonomously collect the plastic bottles on the water surface.

There are three key technical aspects of collecting plastic bottles by grasping: 1) plastic bottles detection for locating objects; 2) grasp detection for estimating the best graspable position; and 3) grasp control for control the motion of mechanical arm based on grasp position information. The grasp detection plays an important role in this task since it bridges the preceding detection and the following grasp control. Therefore, it is imperative to improve the grasp detection performance for real-world grasping.

Grasp detection is a visual task to detect graspable regions in images [5]. In recent years, supervised learning based on Convolutional Neural Network (CNN) is prevalently adopted in the research community, because of its efficient end-to-end training and inferring manners [6, 7]. In this context, Redmon *et al.* and Asif *et al.* focused on directly predicting grasp rectangles (defined by its position, orientation, width and height) in images, both of which obtained state-of-the-art grasp detection accuracies on the Cornell grasp dataset [8, 9]. One great challenge for these methodology is the feasibility of gathering a large amount of labeled training data, when the application scenarios are different from existing datasets. Because it is very expensive and time consuming to generate an available dataset in the real-world. To overcome this problem, a physics simulation engine was utilized to generate training data [10]. However, the training process need several weeks on multiple GPUs in parallel working [11]. The good performance of deep learning on grasp detection was demonstrated according to various attempts, but the data-driven mode is not the most suitable method of grasp detection for plastic bottles on the water surface. There are two primary reasons for it. On the one hand, there lacks labeled training data for a special scene. On the other hand, the aquatic robot with a poor computation power cannot fit within the considerable amount of computational cost. In this situation, an unsupervised method of grasp detection with high speed is necessary.

As a commonsense, the shape of plastic bottle can be abstracted as an elongated cylinder. Therefore, the graspable position of a bottle can be modeled by three parameters, i.e., the central point, the long axis direction angle, and the maximum width. After the object detection, bounding boxes of plastic bottles are obtained. Each bounding box consists of

This work is supported by the National Natural Science Foundation of China under Grant 61725305, Grant 61633004, Grant 61836015, and Grant 61633020.

the water surface background and the plastic bottle. To extract the locations of plastic bottle, the region of bottle should be separated from the bounding box. The water surface in the neighborhood of bottles has the characters of single color and balanced illumination. The RGB histogram of water surface is in accordance with normal distribution. Based on this character, a high speed background elimination method is proposed. After background elimination, the region of bottle is reserved and its pixel-level locations are extracted. Finally, the principal component analysis (PCA) is employed in the location points to estimate the aforesaid three parameters of graspable position.

The primary contributions of this paper are concluded into three parts. Firstly, the graspable position of plastic bottles on the water is modeled by the central point, the long axis direction angle, and the maximum width. Secondly, a high speed background elimination method based on RGB histogram is proposed to obtain pixel-level locations of plastic bottles. Thirdly, PCA with an outliers removing algorithm is employed to estimate the graspable position parameters, and experimental results demonstrate the feasibility of the proposed grasp detection method.

The rest of this paper is organized as follows. In Section 2, the grasping model is presented to introduce the significance of three graspable parameters. Section 3 deals with a high speed background elimination method based on RGB histogram. The process of position estimation based on PCA is detailed in Section 4. Experimental results and discussion are shown in Section 5. Finally, the conclusion and future work are given in Section 6.

2 Modeling Robotic Grasp

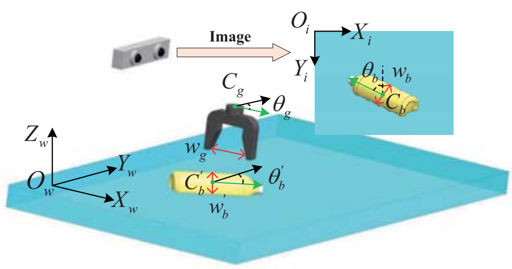


Fig. 1: Illustration of the grasping model.

The shape of plastic bottles can be abstracted as an elongated cylinder, and the grasping model of plastic bottles is illustrated in Fig. 1. For an akin-cylinder plastic bottle, its position in the world coordinate system can be described as

$$P_b{}' = \{C_b{}', \theta_b{}', w_b{}'\} \tag{1}$$

where $C_b{}'$ denotes the central point, $\theta_b{}'$ is the long axis direction, and $w_b{}'$ represents the maximum width of a plastic bottle.

In the process of controlling a two-finger gripper to grasp a plastic bottle on the water, the pose of the griper and the distance between two fingers w_g serve as the main control variables, the former of which consists of the coordinates of the gripper joint C_g and the direction between two fingers θ_g. In practice, the direction between two fingers is parallel to the water surface. The objective of grasp control can be described as

$$\begin{cases} C_g = C_b{}' + \vec{t} \\ \theta_g = \theta_b{}' + \frac{\pi}{2} \\ w_g = w_b{}' + \Delta w \end{cases} \tag{2}$$

where $\vec{t}$ denotes a translation transformation, and Δw is a redundant parameter to guarantee grasp success.

Therefore it is feasible to describe the plastic bottle grasp by the three parameters in the world coordinate system. Furthermore the plastic bottle in the image coordinate system can be presented as

$$P_b = \{C_b, \theta_b, w_b\} \tag{3}$$

where C_b denotes pixel coordinate of the central point, θ_b is the long axis direction angle, and w_b represents the maximum pixel width of a plastic bottle in an image.

A RGB-D image can be obtained by the 3D-reconstruction with a binocular camera. Then the mapping relationship between $P_b{}'$ and P_b can be determined. In the context, the accurate estimation of graspable position P_b in image is a crucial task.

3 High Speed Background Elimination Method

After the object detection, a bounding box of the plastic bottle is located on the image. Nevertheless, each bounding box consists of the water surface background and the plastic bottle. To extract the pixel-level locations of the plastic bottle accurately, the water background should be eliminated and the segment of the bottle should be reserved. To this end, an high speed methodology for water background elimination based on RGB histogram is proposed to extract the region of bottles.

3.1 Normal Distribution of Water Background

In the same water area, the color of water is similar, and the illumination is balanced and consistent. The RGB histograms of water surface images in different water areas are shown in Fig. 2. It is apparent that the distribution of RGB histogram is accordant to normal distribution as

$$P_c(x) = \frac{1}{\sqrt{2\pi}\sigma_c} e^{-\frac{(x-\mu_c)^2}{2{\sigma_c}^2}} \tag{4}$$

where $P_c(x)$ denotes the possibility at intensity x in channel c, μ_c and σ_c denote mean value and standard deviation in channel c respectively.

In this situation, μ_c and σ_c can describe the characteristic of a water surface at a specific area. The different mean value and standard deviation between water surface and plastic bottle are used as a basis to segment images.

3.2 Process of Background Elimination Method

To eliminate the background, there are four steps, i.e., background learning, image gridding, background elimination, and pixel-level object location. The process is illustrated in Fig. 3.

Background learning. At the beginning of grasping plastic bottles, images of water surface should be collected to obtain the μ_r and σ_r for reference. μ_r and σ_r are arithmetically

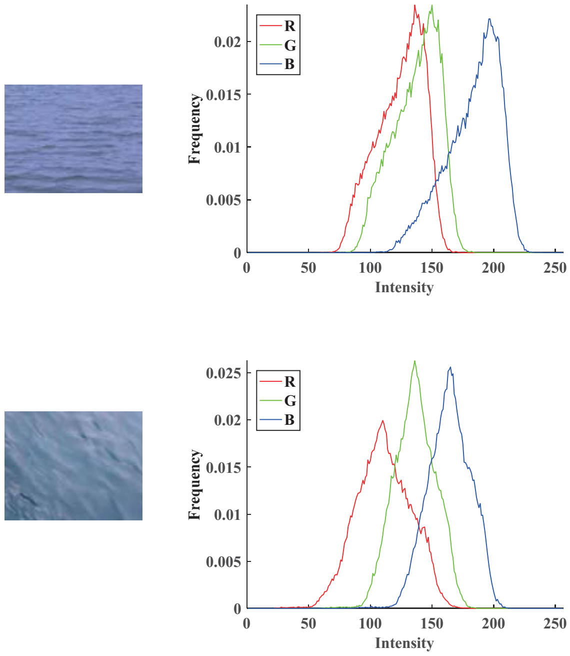


Fig. 2: RGB histograms of water surface images.

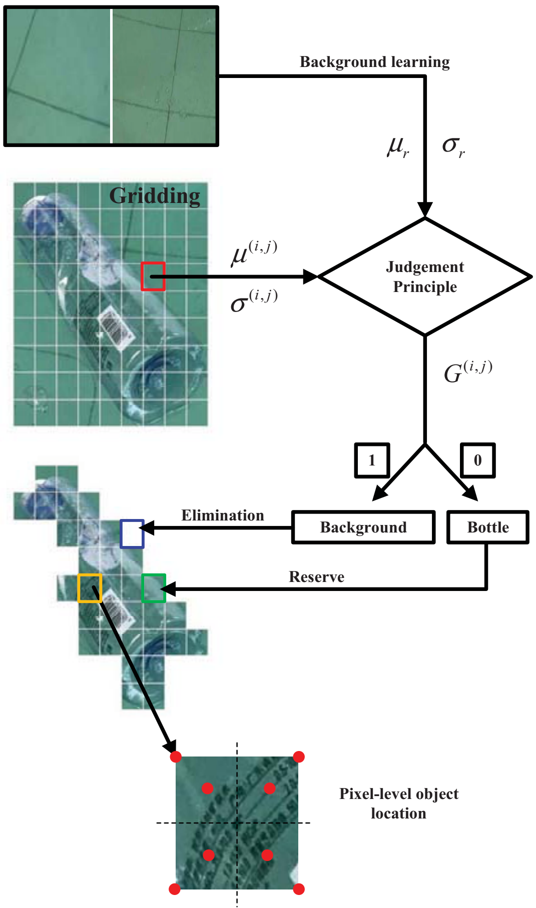


Fig. 3: Process of background elimination method.

derived as

$$\mu_r = \frac{1}{3} \sum_{c \in \{R,G,B\}} \mu_c \tag{5}$$

$$\sigma_r = \frac{1}{3} \sum_{c \in \{R,G,B\}} \sigma_c. \tag{6}$$

Image gridding. The bounding box is gridded into $N \times N$ grids. Then mean value and standard deviation are computed in each grid as

$$\mu^{(i,j)} = \frac{1}{3} \sum_{c \in \{R,G,B\}} \mu_c^{(i,j)} \tag{7}$$

$$\sigma^{(i,j)} = \frac{1}{3} \sum_{c \in \{R,G,B\}} \sigma_c^{(i,j)} \tag{8}$$

where $i, j \in [1, N]$. This step decreases the computational cost from by converting pixel-level computation to grid-level estimation.

Background elimination. In this step, a binary state $G^{(i,j)}$ of each grid is provided to label whether a grid belongs to background or not. $G^{(i,j)} = 1$ means the grid belongs to background, and $G^{(i,j)} = 0$ represents that the grid belongs to region of bottles. The principle of judgement is described as

$$G^{(i,j)} = \begin{cases} 1 & |\mu^{(i,j)} - \mu_r| < \lambda\sigma_r \quad \text{or} \\ & |\sigma^{(i,j)} - \sigma_r| < \beta\mu_r \\ 0 & else \end{cases} \tag{9}$$

where the factors of λ and β are determined empirically. After the judgement, grids with $G^{(i,j)} = 1$ are deleted to achieve the background elimination. Then, a set of grids belonging to region of bottles Q_{bottle} is obtained as

$$Q_{bottle} = \{(i,j) | G^{(i,j)} = 0, \forall i, j \in [1, N]\} \tag{10}$$

where (i, j) denotes the index of grids.

Pixel-level object location. For each grid in Q_{bottle}, eight location points are determined by a uniform pattern as

$$Point^{(i,j)} = \{(x_c^{(i,j)} + pw^{(i,j)}, y_c^{(i,j)} + qh^{(i,j)}) \\ |p, q \in \{\pm\frac{1}{4}, \pm\frac{1}{2}\}\} \tag{11}$$

where $x_c^{(i,j)}$ and $y_c^{(i,j)}$ denote the central point coordinates of the (i, j)th grid, $w^{(i,j)}$ and $h^{(i,j)}$ denote the width and height of the (i, j)th grid respectively.

After aforesaid steps, a complete set of location points for plastic bottles has been constructed. The set can be described as

$$\boldsymbol{I} = \{(x, y) | (x, y) \in Point^k, \forall k \in Q_{bottle}\}. \tag{12}$$

4 PCA-based Grasp Position Estimation

PCA has been widely used in dimension reduction. By maximizing the variance in data, it captures the dominant features in an N-dimensional data in descending order through an orthogonal transformation [12, 13]. The central point, the long axis direction, and the maximum width are dominant one-dimension features which can be captured from two-dimensional coordinates of location points. Therefore, PCA is a feasible method to estimate the graspable position.

4.1 Theory of PCA

A dataset is given as $\boldsymbol{X} = [\boldsymbol{x_1}, \boldsymbol{x_2}, ..., \boldsymbol{x_n}]$, where $\boldsymbol{x_n}$ is a N-dimensional vector. Then the mean value $\boldsymbol{\mu}$ and the covariance matrix $\boldsymbol{S}$ of this dataset is derived as

$$\boldsymbol{\mu} = \frac{1}{n}\sum_{i=1}^{n}\boldsymbol{x_i} \tag{13}$$

$$\boldsymbol{S} = \sum_{i=1}^{n}(\boldsymbol{x_i} - \boldsymbol{\mu})(\boldsymbol{x_i} - \boldsymbol{\mu})^T. \tag{14}$$

With the purpose of reducing dimension from N to m, a projection matrix $\boldsymbol{P}$ transforms each element in the $\boldsymbol{X}$ as

$$\boldsymbol{y_i} = \boldsymbol{P}^T(\boldsymbol{x_i} - \boldsymbol{\mu}). \tag{15}$$

To keep the information of original data as much as possible, a least squares reconstruction error is in need. The reconstruction data is derived by

$$\hat{\boldsymbol{x}}_{\boldsymbol{i}} = \boldsymbol{P}\boldsymbol{y_i}. \tag{16}$$

Thus the squares reconstruction error is described as:

$$e = \sum_{i=1}^{n}||\boldsymbol{x_i} - \boldsymbol{\mu} - \hat{\boldsymbol{x}}_{\boldsymbol{i}}||^2. \tag{17}$$

In this context, to minimize the squares reconstruction error is equal to the optimization as follows:

$$\boldsymbol{P} = \arg\max_{\boldsymbol{P}} |\boldsymbol{P}^T\boldsymbol{S}\boldsymbol{P}|. \tag{18}$$

To solve the optimization, the covariance matrix $\boldsymbol{S}$ is decomposed as:

$$\boldsymbol{S} = \boldsymbol{Q}\boldsymbol{\Sigma}\boldsymbol{Q}^{-1}. \tag{19}$$

After rearrangement of eigenvalues in a descending order, the top m eigenvalues are determined, and then the corresponding eigenvectors compose $\boldsymbol{P}$.

4.2 PCA for Grasp Position Estimation

Location point set $\boldsymbol{I}$ of plastic bottle obtained in Section 3 serves as the feature in PCA process. The central point C_b, the direction of long axis $\boldsymbol{D_l}$, and the direction of short axis $\boldsymbol{D_s}$ can be computed through PCA as shown in Fig. 4.

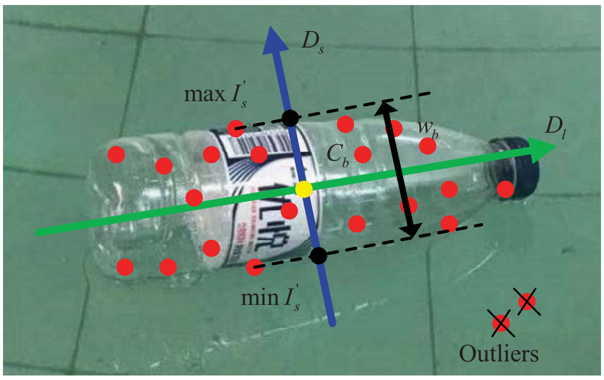


Fig. 4: Parameters of grasp position.

Firstly, the central point of bottle is determined by

$$C_b = \frac{1}{M}\sum_{p_i \in \boldsymbol{I}}^{M} p_i \tag{20}$$

where M denotes the number of points in $\boldsymbol{I}$. The the covariance matrix $\boldsymbol{S'}$ is shown as

$$\boldsymbol{S'} = \sum_{p_i \in \boldsymbol{I}}(p_i - C_b)(p_i - C_b)^T. \tag{21}$$

Next, $\boldsymbol{S'}$ is decomposed by

$$\boldsymbol{S'} = [\boldsymbol{D_l}\ \boldsymbol{D_s}]^T \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} [\boldsymbol{D_l}\ \boldsymbol{D_s}] \tag{22}$$

where λ_1 and λ_2 ($\lambda_1 > \lambda_2$) are the eigenvalue of $\boldsymbol{S'}$.

To estimate the maximum width of plastic bottle, location points will be projected on the short axis as

$$\boldsymbol{I'_s} = \{\boldsymbol{D_s}(p_i - C_b)|\forall p_i \in \boldsymbol{I}\} \tag{23}$$

where $\boldsymbol{I'_s}$ denotes the set of points projected on the short axis. Then the maximum width w_b is derived as:

$$w_b = \max \boldsymbol{I'_s} - \min \boldsymbol{I'_s}. \tag{24}$$

So far the grasp position parameters are determined, i.e., the central point C_b, the long axis direction $\boldsymbol{D_l}$, and the maximum width w_b. Note that the angle of long axis direction θ_b indicated in Section 2 can be computed by $\boldsymbol{D_l}$.

4.3 Outliers Removing Algorithm

As shown in Fig. 4, there are some outliers in location point set, which affect the accuracy of w_b tremendously. According to observation distribution of location points on bottle, most of points assemble around the long axis. Inspired by this characteristic of distribution, the set of points projected on short axis $\boldsymbol{I'_s}$ is divided into a valuable point subset $\boldsymbol{V_s}$ and outlier subset $\boldsymbol{O_s}$. It is apparent that the distance of two adjacent points in the same subset is shorter than two points in different subsets. Based on this, the outliers removing algorithm is proposed as shown in **Algorithm 1**, which compensates for the error resulted by background elimination method.

5 Experiments and Discussion

To demonstrate the feasibility of the proposed strategy, evaluation criterion, performance of the proposed strategy, and contrast experiments with other methods are specified in this section. Note that images of various plastic bottles in different orientations on the water surface are collected in laboratory environment, and each of these images is labeled with groundtruth grasp position parameters manually.

5.1 Evaluation Criterion

The groundtruth position parameters, i.e., the central point, the long axis direction angle, and the maximum width, are denoted with C_b^t, θ_b^t, and w_b^t, respectively. Relatively, the estimated value of aforesaid three parameters are expressed as $\hat{C}_b$, $\hat{\theta}_b$, and $\hat{w}_b$ separately. Evaluation criterions of the grasp position estimation performance include the relative offset ratio of central point Δp, the absolute error angle of long axis direction ε_{angle}, and the relative error ratio of maximum width d_w. The definition of relative offset ratio of central point Δp is shown as

$$\Delta p = \frac{|\hat{C}_b - C_b^t|}{w_b^t}. \tag{25}$$

Algorithm 1 Outliers removing algorithm

$\boldsymbol{I}_s' = Sort(\boldsymbol{I}_s')$;
  // Sort $\boldsymbol{I}_s'$ in the ascending order
$n = Size(\boldsymbol{I}_s')$;
  // Initialize the number of points in $\boldsymbol{I}_s'$
α;
$m = 0$;
$\boldsymbol{V_s} = \emptyset$;
$\boldsymbol{O_s} = \emptyset$;
for $i = 1; i < n; i++$ **do**
  $\boldsymbol{Dis}(i) = \boldsymbol{I}_s'(i+1) - \boldsymbol{I}_s'(i)$;
    // Compute the distance of adjacent points of points
end for
$dis_{mean} = Mean(\boldsymbol{Dis})$;
for $i = 1; i < n; i++$ **do**
  $\boldsymbol{Set}(m) \cup \boldsymbol{I}_s'(i)$;
  if $(\boldsymbol{Dis}(i) >= \alpha dis_{mean})$ **then**
    $m = m + 1$;
  end if
end for
$Index = \arg\max_m Size(\boldsymbol{Set}(m))$;
$\boldsymbol{V_s} = \boldsymbol{Set}(Index)$;
for $i = 1; i <= m; i++$ **do**
  if $(i \neq Index)$ **then**
    $\boldsymbol{O_s} = \boldsymbol{O_s} \cup \boldsymbol{Set}(i)$;
  end if
end for
$Delete(\boldsymbol{O_s})$;

The absolute error angle of long axis direction ε_{angle} is denoted by

$$\varepsilon_{angle} = |\hat{\theta}_b - \theta_b^t|. \quad (26)$$

The relative error ratio of maximum width d_w is expressed with

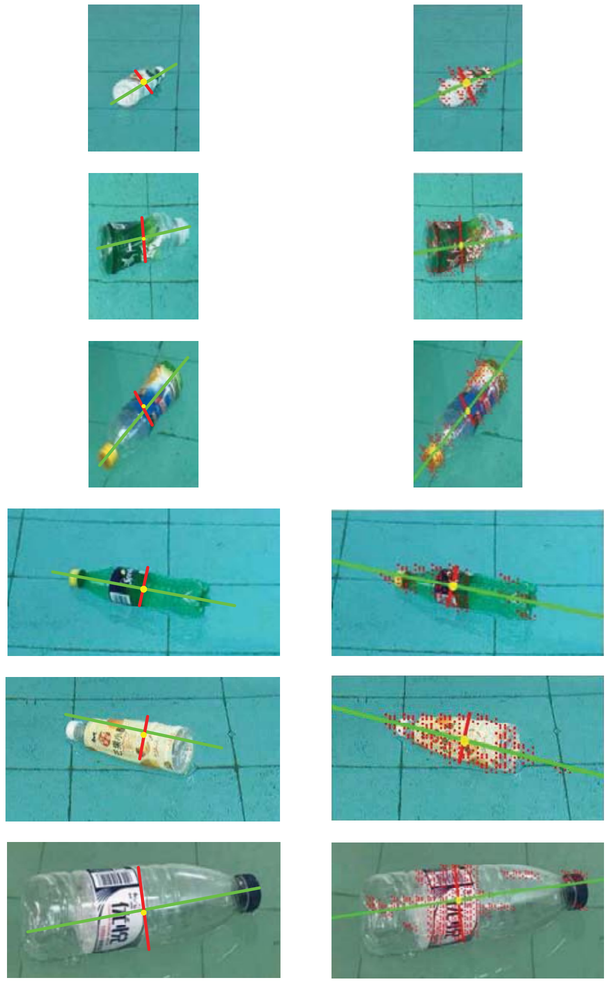
$$d_w = \frac{|\hat{w}_b - w_b^t|}{w_b^t}. \quad (27)$$

5.2 Performance of Grasp Position Estimation

After object detection, bounding boxes of bottles are obtained. Note that bounding boxes are with different sizes. For each bounding box, the groundtruth is labeled as shown in Fig. 5: (a). The grasp position estimation strategy is executed in Matlab, and the results are drawn on each bounding box as shown in Fig. 5: (b). Note that the factor in background elimination λ and β are set as 2 and 3, respectively. And the parameter α in outliers removing algorithm is set as 5. Qualitatively, the results of grasp position estimation are satisfactory.

For quantitative analysis of the performance, the three evaluation criterions are computed in Table. 1. Because of the redundant values in the practical grasp control process, if the estimated results meet $\Delta p < 0.15$, $\varepsilon_{angle} < 5°$, and $d_w < 0.15$ simultaneously, the estimation results are regarded as feasible values.

Furthermore, in the process of practical grasp control, the angle of long axis direction is more important than others, for two fingers of the griper can splay in a wider manner to compensate for errors of the estimated central point and the maximum width. Appropriately, the error of estimated long axis direction with the proposed strategy is pretty slight, even less than 3°. Therefore, the proposed strategy is feasible to estimate grasp position and guide the grasp process.



(a) (b)

Fig. 5: Results of grasp position estimation. (a) Groundtruth. (b) Estimated results. The yellow point denotes the central point, the green straight line denotes the long axis direction, the red line segment expresses the maximum width, and the red scatter points in right figure are location points.

Table 1: Evaluation criterion of each grasp position estimation

Bottle order	Δp	ε_{angle}	d_w	Feasibility
1	0.0928	0.9776°	0.0912	✓
2	0.0581	2.7644°	0.0405	✓
3	0.1309	2.9771°	0.1191	✓
4	0.1220	1.2825°	0.0772	✓
5	0.1423	2.8607°	0.0827	✓
6	0.0232	1.4890°	0.1066	✓

5.3 Contrast Experiments on Computation Speed

In this paper, a high speed background elimination method is proposed to obtain location points of plastic bottles on the water surface. On account of the gridding step, the computational complexity is simplified. To extract the feature points and obtain the pixel-level locations, Canny and Prewitt are available methodologies [14, 15]. The proposed method and the two aforesaid methods are employed to extract the feature points of plastic bottles, contrastively. Then respective location points are used to estimate the grasp position by the PCA-based strategy. With the Computation speed S_p of proposed method as reference, results are shown in Table. 2. It is noticeable that three methods with PCA obtain feasible estimated parameters, but the proposed background elimination method has a higher speed.

Table 2: Comparison on computation speed

Method	Computation Speed	Feasibility
Proposed + PCA	S_p	✓
Canny + PCA	$0.6779S_p$	✓
Prewitt + PCA	$0.9125S_p$	✓

5.4 Discussion

The experimental results indicate that the PCA-based unsupervised grasp detection framework is feasible and efficient, and that background elimination method has higher Computation speed. Therefore, the proposed grasp detection strategy is an appropriate choice for the real-time plastic bottle collection on aquatic robots. The reasons for the satisfactory performance on grasp detection accuracy and computation speed lie in two aspects. The first reason is that the majority of location points of bottle distribute along with an elongated cylinder and the dominant information of 2-dimensional data derived by PCA is the long axis direction. The second reason is that the computation of extracting location points is simplified by converting pixel-level computation to grid-level estimation. In practice, errors resulted by the high speed background elimination method might cause few outliers of location points. But the proposed outliers removing algorithm compensates for the defect efficiently. Furthermore, the unsupervised grasp detection strategy does not requires training process, which avoids consuming much time and labor for building a labeled dataset.

The major limitation of the proposed strategy is that pixel-level location points are sparse to some extent. On the one hand, the bounding box is divided into grids, whose size restricts the number of location points. On the other hand, some regions of plastic bottles could be semitransparent so that the color could be similar to water background. In this situation, the principle of judgment based on mean value and standard deviation could be hard to describe the water surface entirely. Therefore, the principle of judgment will be improved in the future.

6 Conclusion and Future Work

As a crucial part of collecting plastic bottles for the aquatic robot, an unsupervised grasp detection strategy is detailed in this paper. Based on the approximation that the plastic bottle can be regarded as an elongated cylinder, a grasping model is constructed. To extract pixel-level locations of bottles on the water surface, a high speed background elimination method is utilized in the gridding bounding box. PCA with an outliers removing algorithm is employed to estimate graspable position. Finally, experimental results of grasp detection demonstrate the feasibility and the high-efficiency of the proposed strategy. Especially, the accuracy of long axis direction angle is up to $3°$, and computation speed of the background elimination method is 1.47 times the speed of Canny.

In the future, the background elimination method will be improved to compensate for the limitation of sparse location points. Furthermore, the grasp detection strategy will be utilized on the aquatic robot to collect the plastic bottles and protect the aquatic environment.

References

[1] J. R. Jambeck, R. Geyer, C. Wilcox, *et al.*, "Plastic waste inputs from land into the ocean," *Science*, vol. 347, no. 6223, pp. 768–771, 2015.

[2] L. C. M. Lebreton, J. van der Zwet, *et al.*, "River plastic emissions to the worlds oceans," *Nat. Commun.*, vol. 8, no. 15611, 2017.

[3] Plastics Europe, "Plastics-the facts 2016: an analysis of European plastics production, demand and waste data," `http://www.plasticseurope.org`, 2016.

[4] L. Guern, "Plastic pollution", `http://coastalcare.org/2009/11/plasticpollution/`, 2009.

[5] I. Lenz, H. Lee, and A. Saxena, "Deep learning for detecting robotic grasps," *Int. J. Robot. Res.*, vol. 34, no. 4–5, pp. 705–724, 2015.

[6] S. Levine, P. Pastor, A. Krizhevsky, *et al.*, "Learning hand-eye coordination for robotic grasping with deep learning and large-scale data collection," *Int. J. Robot. Res.*, vol. 37, no. 4–5, pp. 421–436, 2018.

[7] S. Kumra and C. Kannan, "Robotic grasp detection using deep convolutional neural networks," in *Proc. IROS*, Vancouver, Canada, Sep. 2017, pp. 769–776.

[8] J. Redmon and A. Angelove, "Real-time grasp detection using convolutional neural networks," in *Proc. ICRA*, Seattle, Washington, USA, May. 2015, pp. 1316–1322.

[9] U. Asif, M. Bennamoun, and F. A. Sohel, "RGB-D object recognition and grasp detection using hierarchical cascaded forests," *IEEE Trans. Robot.*, vol. 33, no. 3, pp. 547–564, 2017.

[10] E. Johns, S. Leutenegger, and A. J. Davison, "Deep learning a grasp function for grasping under gripper pose uncertainty," in *Proc. IROS*, Daejeon, Korea, Oct. 2016, pp. 4461–4468.

[11] U. Asif, J. Tang, and S. Harrer, "GraspNet: an efficient convolutional neural network for real-time grasp detection for low-powered devices," in *Proc. IJCAI*, Stocholm, Sweden, Jul. 2018, pp. 4875–4882.

[12] M. D. Farrell and M. M. Russell, "On the impact of PCA dimension reduction for hyperspectral detection of difficult targets," *IEEE Geosci. Remote Sens. Lett.*, vol. 2, no. 2, pp. 192–195, 2005.

[13] A. Malhi and R. X. Gao, "PCA-based feature selection scheme for machine defect classification," *IEEE Trans. Instrum. Meas.*, vol. 53, no. 6, pp. 1517–1525, 2004.

[14] L. Ding and A. Goshtasby, "On the Canny edge detector," *Pattern Recognition*, vol. 34, no. 3, pp. 721–725, 2001.

[15] G. T. Shrivakshan and C. A. Chandrasekar, "A comparison of various edge detection techniques used in image processing," *Int. J. Comput. Sci. Issues*, vol. 9, no. 5, pp. 269–276, 2012.