IWSCR: An Intelligent Water Surface Cleaner Robot for Collecting Floating Garbage

Shihan Kong, Manjun Tian, Changlin Qiu, Zhengxing Wu, and Junzhi Yu, *Senior Member, IEEE*

***Abstract*—In this article, a robot system for intelligent water surface cleaner named IWSCR is developed to collect floating plastic garbage. It is able to accomplish three major tasks autonomously, i.e., cruise and detection, tracking and steering, and grasping and collection. The challenges behind these tasks involve how to realize the accurate and real-time garbage detection, how to resist the disturbances while IWSCR conducts vision-based steering, and how to grasp the floating garbage reliably despite the turbulent conditions on the surface of the water. To overcome these difficulties, three key techniques are proposed for IWSCR. First, the YOLOv3 network, which is widely applied in the high speed and accuracy object detection field, is trained on the proposed floating garbage dataset to realize accurate and real-time garbage detection. Next, to improve the ability of resisting disturbances, a control law based on the sliding-mode controller is proposed for vision-based steering. Furthermore, inspired by the stability of floating bottles in fluid, a feasible grasping strategy is utilized for IWSCR. Finally, the experimental results demonstrate that IWSCR is competent to carry out the task of water surface cleaning.**

***Index Terms*—Deep learning, grasping strategy, intelligent water surface cleaner, robot system, sliding-mode controller (SMC).**

I. Introduction

WATER is the source of life. The oceans and rivers cover almost 71% of the earth's surface and provide a suitable home for billions of aquatic organisms [1]. However, humanity does not treat the aquatic environment in a friendly manner. The water pollution resulting from human negligence has been accumulating for decades. The waste in water consists of dredge, industrial garbage, sewage, radioactive materials, and plastic trash [2]. With expanding applications for robots, many authors have reported their attempts to develop robots designed to clean local environments [3]–[10]. Naturally, robots designed to clean bodies of water merit development and study.

To collect plastic pollution, the semimanual refuse-removal vessels are widely applied recently. However, the refuse-removal vessel is large in size which is only appropriate to rivers with huge acreage or much accumulated waste. In this context, it is impracticable to clean small, low-density waste in small waters by the refuse-removal vessel. Notably, this method lacks the capacity to determine which floating objects merit removal and which ones do not. Additionally, the exhaust from the refuse-removal vessel may cause secondary pollution. We aim to address these limitations by developing a zero-emissions water-cleaning robot system that can collect objects recognized as garbage.

Above all, the intelligent water surface cleaner robot is a compositive robotic system with vision module, motion control module, and grasping module, which can sequentially accomplish three tasks (TTs), i.e., cruise and detection, tracking and steering, and grasping and collection. In the first task, the robot moves on the water surface following the preplanned path and uses its vision module to detect garbage. When an object is targeted for removal, the second task begins. The vision module tracks the target and measures the relative position between the robot and the target. In the meantime, the motion control module utilizes the position information from the vision module to correct the yaw angle error seamlessly ensuring an accurate target-approach angle. In the final task, the grasping module determines the grasp timing, and then it commands the manipulator to grasp and collect the objects. Notably, there are three major technical challenges standing in front of accomplishing the aforementioned TTs.

1) *Challenge for Cruise and Detection:* To distinguish the garbage despite the robot moving, the object detection algorithm of the cleaner robot should be accurate and real time.
2) *Challenge for Tracking and Steering:* There are unpredictable dynamic factors in the aquatic environment, like wind and waves, that make it necessary.
3) *Challenge for Grasping and Collection:* The dynamic conditions on the surface of water means that objects will move, including the robot itself. This poses a

Manuscript received August 2, 2019; accepted December 10, 2019. This work was supported in part by the National Natural Science Foundation of China under Grant U1909206, Grant 61725305, Grant 61633004, and Grant 61633020, and in part by the Pre-Research Fund of Equipments of China under Grant 61403120108. This article was recommended by Associate Editor S. Song. *(Corresponding author: Junzhi Yu.)*

Shihan Kong, Changlin Qiu, and Zhengxing Wu are with the State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China, and also with the School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing 100049, China (e-mail: kongshihan2016@ia.ac.cn; qiuchanglin2019@ia.ac.cn; zhengxing.wu@ia.ac.cn).

Manjun Tian is with the School of Information Engineering, Minzu University of China, Beijing 100081, China (e-mail: tianmanjun2018@163.com).

Junzhi Yu is with the State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China, and also with the State Key Laboratory for Turbulence and Complex Systems, Department of Mechanics and Engineering Science, BIC-ESAT, College of Engineering, Peking University, Beijing 100871, China (e-mail: junzhi.yu@ia.ac.cn).

Color versions of one or more of the figures in this article are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSMC.2019.2961687

challenge for grasping an object that is likely to be moving.

Related theories and methods, which are appropriate to be employed in the cleaner robot to solve the aforementioned problems, have been widely studied in recent years. For the aspect of object detection, deep neural networks have been applied to object detection in images with great success [11]. There are two major frameworks for detection, i.e., the two-stage framework and the one-stage framework. The two-stage framework is based on R-CNN with a region proposal network [12]–[15]. With the advantage of high computing speed, the one-stage framework converts the object detection to a bounding box regression problem, such as SSD [16], [17] and YOLO [18]–[20] networks. For techniques of the motion control module, the sliding-mode controller (SMC) has been successfully applied on autonomous underwater vehicles (AUVs) owing to its insensitivity to model parameters and external disturbances [21]–[26]. Additionally, grasping strategies with grasping detection based on deep neural network are studied extensively nowadays [27]–[32], by which the graspable position can be determined accurately. However, these data-driven strategies take too much time to execute in real time and, therefore, are not suitable for the dynamic grasping process.

In this article, IWSCR, a prototype of intelligent water surface cleaner robots, is designed, which can work in the experimental tank environment and accomplish the TTs with the purpose of collecting a plastic bottle, a plastic bag, and styrofoam floating on the water surface. IWSCR is a small underwater vehicle that is equipped with a binocular camera, a manipulator, and a collection box. To tackle the challenges, the object detection algorithm is based on the YOLOv3 network is employed for improving the accuracy and speed of detection by being trained on our proposed dataset; the SMC-based control law with a satisfactory capability of resisting disturbances is designed for the vision-based steering to guide IWSCR toward the target; based on the stability of objects in fluid, a feasible grasping strategy for floating objects is proposed to solve the problem of dynamic grasping. Here, we report our experimental results, which demonstrate that IWSCR has the ability to accomplish TTs and that it is competent to complete the work of water surface cleaner. The primary contributions of this article are twofold.

1) IWSCR is a novel design for floating garbage collection. IWSCR accomplishes TTs autonomously due to our integration of computational strategies for vision-based detection, identification, motion control, and grasping. To the best of our knowledge, there are few, if any, published examples of autonomous water surface cleaning robots. We suspect that IWSCR will attract interest from researchers.
2) To overcome the related technical difficulties, the YOLOv3 detection framework, SMC for the vision-based steering, and a feasible grasping strategy based on the stability of objects in the fluid are employed in IWSCR. The experimental results indicate that the performances of these technical methods are satisfactory. Especially, the proposed grasping strategy inspired by the characteristic of floating objects provides a novel implementation for the dynamic grasping in fluid. In addition, IWSCR might serve as a platform to test related techniques in the future.

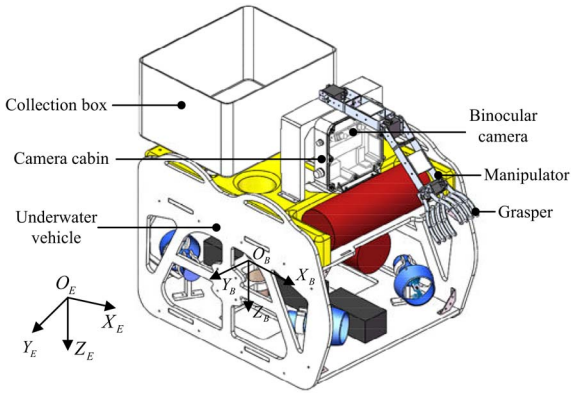

Fig. 1. Configuration of IWSCR.

The remainder of this article is organized as follows. In Section II, the design of IWSCR is overviewed. The object detection framework based on the YOLOv3 network, SMC for vision-based steering, and grasping strategy are detailed in Sections III–V, respectively. Next, the experimental results and discussion are offered in Section VI. Finally, the conclusion and future work are summarized in Section VII.

II. Prototype Design of IWSCR

A. Configuration of IWSCR

The configuration of IWSCR is illustrated in Fig. 1. The system is based on an electrically powered underwater vehicle that can sail on the water surface under the loading condition. The underwater vehicle is about 63.5-cm long, 48.5-cm wide, and 46.5-cm high and it weights approximately 25 kg. A camera cabin is installed in the front of the top of the vehicle, in which a binocular camera is fixed on the clapboard. A manipulator with 3-DOF is placed on the vehicle, which is composed of three servo motors. Note that the scope of joint angles is 270° to ensure that the tail end of manipulator can move to the inner of the collection box. Thus, the size of IWSCR is diminutive enough to be applied in the small waters, and its source of energy is clean to avoid the secondary pollution.

B. Framework of Control System

As shown in Fig. 2, the personal computer (PC) processes with the information from the binocular camera in order to control IWSCR. There are three communication modes between PC and IWSCR, i.e., USB video class (UVC), TCP/IP, and Bluetooth. The image processor in PC obtains images from IWSCR and estimates the object position. The estimated position flows in two different ways. The first way includes steering error generator, motion controller, and thruster allocation. In this way, the computed error is fed into the motion controller to obtain the relative force and moment. According to the thruster allocation, the force and moment

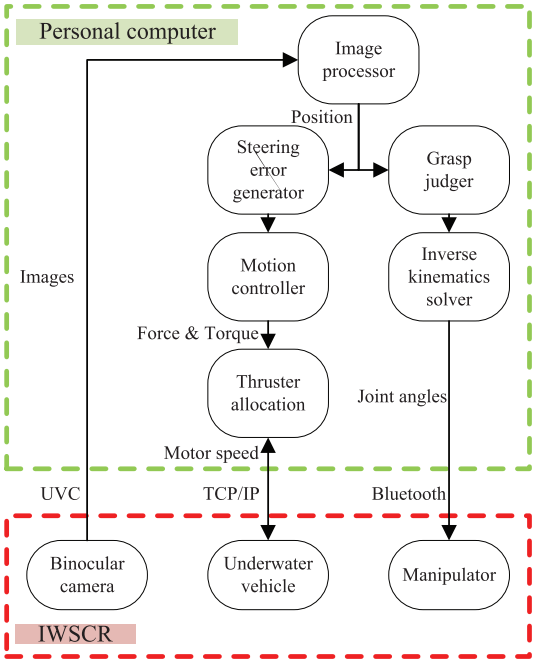

Fig. 2. Framework of control system.

are converted to motor speed, which can control the underwater vehicle directly. The second way includes grasp judger and inverse kinematics solver. In this way, grasping judger determines the timing and graspable position, and the inverse kinematics solver computes the joint angles to control the manipulator. Notice that the control system for IWSCR is inspired by human studies of grasping and object recognition. There is a considerable evidence to suggest that humans (and nonhuman primates) possess different cortical networks for recognizing objects and for planning movements to grasp them [33], [34].

III. Accurate and Real-Time Garbage Detection

Vision techniques are important for IWSCR to perceive the environment, which provide the basic information for subsequent steering and grasping. The deep learning-based vision techniques perform eminently well in object detection. In this article, the YOLOv3 network is employed for garbage detection, which is evolved from the YOLO and YOLOv2 networks [18]–[20]. Compared with R-CNN and its ramification networks, the YOLO framework transforms the detection problem into a regression problem, which does not require the module of proposal region but generates both of the bounding box coordinates and probabilities of each class directly through regression. Therefore, the YOLO framework has greatly higher detection speed compared to faster R-CNN [18], [19], which meets the requirement of real timeness for IWSCR.

The framework of YOLOv3 based on Darknet-53, shown in Fig. 3, is applied to detect three classes of objects floating on the water surface, including plastic bottles, plastic bag, and styrofoam. Different from the YOLO and YOLOv2 networks, YOLOv3 uses multiscale prediction to detect targets and is effective at detecting small objects. For each grid in different scale, YOLOv3 will predict three bounding boxes. The output in each scale is composed of: 1) $\{w, h, x, y, \text{confidence}\}$ for each bounding boxes and 2) probabilities of the three classes. Therefore, the total channels of output is $3 \times (5 + 3) = 24$. Among the predicted bounding boxes, the nonmaximum suppression (NMS) method is used to select the best bounding box. Besides, to ensure the garbage detection accuracy of the trained network, we have established a floating garbage data set including almost 1000 pieces of floating garbage images under different background conditions and illumination intensities. In this context, the total mean average precision (mAP) reaches over 0.90 in the verification stage as shown in Section VI.

After detection, an object is targeted for removal. Then, KCF [36] and triangulation work to update the bounding box continually and measure the position of the object, respectively.

IV. Sliding-Mode Controller for Vision-Based Steering

In this section, based on the dynamic model of the underwater vehicle, an SMC is designed for vision-based steering. The proposed control law is robust for disturbance affecting the input.

A. Dynamic Model of Underwater Vehicle

There are a body-fixed frame $\mathcal{B} = \{X_B, Y_B, Z_B\}$ attached to the vehicle's center of gravity and an inertial frame $\mathcal{I} = \{X_E, Y_E, Z_E\}$ located at the predefined position, as shown in Fig. 1. Following the standard modeling techniques [35], [37], [38], the dynamic model of the underwater vehicle in frame $\mathcal{B}$ will be derived according to the general Newton–Euler motion equation in fluid as follows:

$$\begin{aligned} \mathbf{M}\dot{v} + \mathbf{C}(v)v + \mathbf{D}(v)v + \mathbf{g}(\eta) &= \tau_E + \tau \\ \dot{\eta} &= \mathbf{J}(\eta)v \end{aligned} \tag{1}$$

where:

1) $\eta = [\eta_1\ \eta_2]^T \in \mathbb{R}^6$ denotes the pose vector expressed in $\mathcal{I}$, which involves the position vector $\eta_1 = [x\ y\ z]$ and the orientation vector $\eta_2 = [\phi\ \theta\ \psi]$;
2) $v = [v_1\ v_2]^T \in \mathbb{R}^6$ represents the velocity vector expressed in $\mathcal{B}$, which involves the linear velocity vector $[u\ v\ w]$ and the angular velocity vector $v_2 = [p\ q\ r]$;
3) $\tau = [\tau_X\ \tau_Y\ \tau_Z\ \tau_K\ \tau_M\ \tau_N]^T \in \mathbb{R}^6$ means the total propulsion vector, i.e., forces τ_X, τ_Y, τ_Z and torques τ_K, τ_M, τ_N generated by thrusters and expressed in frame $\mathcal{B}$;
4) $\tau_E \in \mathbb{R}^6$ is the total environmental force/torque vector expressed in frame $\mathcal{B}$, which can be regarded as the external disturbance;
5) $\mathbf{M} = \mathbf{M}_{RB} + \mathbf{M}_A$, where $\mathbf{M}_{RB} \in \mathbb{R}^{6\times6}$ and $\mathbf{M}_A \in \mathbb{R}^{6\times6}$ are the rigid body and added mass inertia matrices, respectively;
6) $\mathbf{C}(v) = \mathbf{C}_{RB}(v) + \mathbf{C}_A(v)$, where $\mathbf{C}_{RB}(v) \in \mathbb{R}^{6\times6}$ and $\mathbf{C}_A(v) \in \mathbb{R}^{6\times6}$ are the Coriolis and centripetal matrix which are resulted by inertial mass and added mass, respectively;
7) $\mathbf{D}(v)$ denotes the drag matrix;

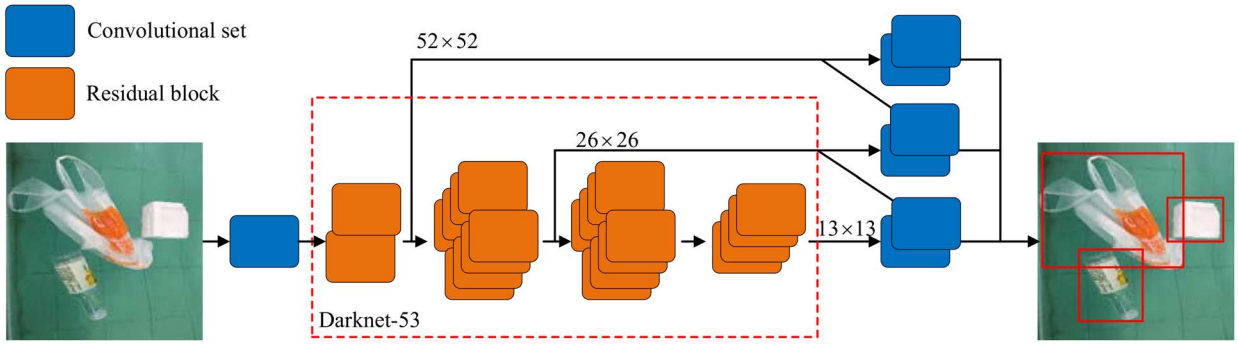

Fig. 3. Framework of YOLOv3 for garbage detection.

8) $\mathbf{g}(\eta)$ represents the hydrostatic force vector;
9) $\mathbf{J}(\eta) = \text{diag}\{\mathbf{J_1}(\eta_2), \mathbf{J_2}(\eta_2)\}$ is the Jacobian matrix transforming velocities from frame $\mathcal{B}$ to frame $\mathcal{I}$, where $\mathbf{J_1}(\eta_2)$ denotes the rotation matrix and $\mathbf{J_2}(\eta_2)$ stands for lumped transformation matrix.

Considering that the underwater vehicle has three symmetric planes, and that the buoyancy center coincides with the gravity center, the expanded expression of $\mathbf{M}$, $\mathbf{C}(v)$, $\mathbf{D}(v)$, and $\mathbf{g}(\eta)$ are separately described as follows:

$$\mathbf{M} = \begin{bmatrix} m - X_{\dot{u}} & 0 & 0 \\ 0 & m - Y_{\dot{v}} & 0 \\ 0 & 0 & m - Z_{\dot{w}} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ I_x - K_{\dot{p}} & 0 & 0 \\ 0 & I_y - M_{\dot{q}} & 0 \\ 0 & 0 & I_z - N_{\dot{r}} \end{bmatrix} \tag{2}$$

$$\mathbf{C}(v) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & mw - Z_{\dot{w}}w & -mv + Y_{\dot{v}}v \\ -mw + Z_{\dot{w}}w & 0 & mu - X_{\dot{u}}u \\ mv - Y_{\dot{v}}v & -mu + X_{\dot{u}}u & 0 \end{bmatrix} \begin{bmatrix} 0 & mw + Z_{\dot{w}}w & -mv + Y_{\dot{v}}v \\ -mw + Z_{\dot{w}}w & 0 & mu - X_{\dot{u}}u \\ mv - Y_{\dot{v}}v & -mu + X_{\dot{u}}u & 0 \\ 0 & I_z r - N_{\dot{r}}r & -I_y q + M_{\dot{q}}q \\ -I_z r + N_{\dot{r}}r & 0 & I_x p - K_{\dot{p}}p \\ I_y q - M_{\dot{q}}q & -I_x p + K_{\dot{p}}p & 0 \end{bmatrix} \tag{3}$$

$$\begin{aligned} \mathbf{D}(v) = -\text{diag}\big\{&X_u - X_{u|u|}|u|, Y_v - Y_{v|v|}|v| \\ &Z_w - Z_{w|w|}|w|, K_p - K_{p|p|}|p|, \\ &M_q - M_{q|q|}|q|, N_r - N_{r|r|}|r|\big\} \end{aligned} \tag{4}$$

$$\mathbf{g}(\eta) = \begin{bmatrix} (mg - B)\sin\theta \\ -(mg - B)\cos\theta\sin\phi \\ -(mg - B)\cos\theta\cos\phi \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{5}$$

where

1) m, g, and B denote the mass, gravity, and buoyancy of the underwater vehicle, respectively;
2) $X_{\dot{u}}$, $Y_{\dot{v}}$, $Z_{\dot{w}}$, $K_{\dot{p}}$, $M_{\dot{q}}$, and $N_{\dot{r}}$ are the added mass terms, respectively;
3) I_x, I_y, and I_z are the moments of inertia along the related axes, respectively;
4) X_u, Y_v, Z_w, K_p, M_q, and N_r denote the first-order drag parameters. $X_{u|u|}$, $Y_{v|v|}$, $Z_{w|w|}$, $K_{p|p|}$, $M_{q|q|}$, and $N_{r|r|}$ denote the second-order drag parameters, respectively.

In this article, the motion of IWSCR is only on the water surface and $[\phi\ \theta]$ is negligible. Therefore, the motion equation of the underwater vehicle is simplified to 3-DOF (i.e., u, v, and r) as follows:

$$\begin{cases} \dot{u}(m - X_{\dot{u}}) - vr(m - Y_{\dot{v}}) - u(X_u + X_{u|u|}|u|) \\ \qquad = \tau_X + \tau_{EX} \\ \dot{v}(m - Y_{\dot{v}}) + ur(m - X_{\dot{u}}) - v(Y_v + Y_{v|v|}|v|) \\ \qquad = \tau_Y + \tau_{EY} \\ \dot{r}(I_z - N_{\dot{r}}) + uv(m - Y_{\dot{v}}) + uv(-m + X_{\dot{u}}) \\ \qquad -r(N_r + N_{r|r|}|r|) = \tau_N + \tau_{EN}. \end{cases} \tag{6}$$

As for thrusters allocation, each thruster produces forces and torques with respect to frame $\mathcal{B}$ as the following equation:

$$^i\tau = \begin{bmatrix} ^iF \\ ^iQ \end{bmatrix} = \begin{bmatrix} ^ie \\ (^iL \times ^ie) \end{bmatrix} ^iT \tag{7}$$

where ^ie denotes the orientation of the ith thruster with respect to frame $\mathcal{B}$, ^iL is the position of attack for ^iT with respect to frame $\mathcal{B}$, and ^iT denotes the thrust of the ith thruster. For the prototype, the distribution of four horizontal thrusters is shown in Fig. 4, where $^1e = [\cos\beta\ \ \sin\beta\ \ 0]^T$, $^2e = [\cos\beta\ \ -\sin\beta\ \ 0]^T$, $^3e = [-\cos\beta\ \ -\sin\beta\ \ 0]^T$, $^4e = [-\cos\beta\ \ \sin\beta\ \ 0]^T$, $^1L = [l\cos\alpha\ \ -l\sin\alpha\ \ 0]^T$, $^2L = [l\cos\alpha\ \ l\sin\alpha\ \ 0]^T$, $^3L = [-l\cos\alpha\ \ l\sin\alpha\ \ 0]^T$, and $^4L = [-l\cos\alpha\ \ -l\sin\alpha\ \ 0]^T$. By substituting the aforementioned values, the total vector of propulsion forces and torques on the water surface is as follows:

$$\tau_X = \begin{bmatrix} \cos\beta \\ \cos\beta \\ -\cos\beta \\ -\cos\beta \end{bmatrix}^T \begin{bmatrix} ^1T \\ ^2T \\ ^3T \\ ^4T \end{bmatrix} \tag{8}$$

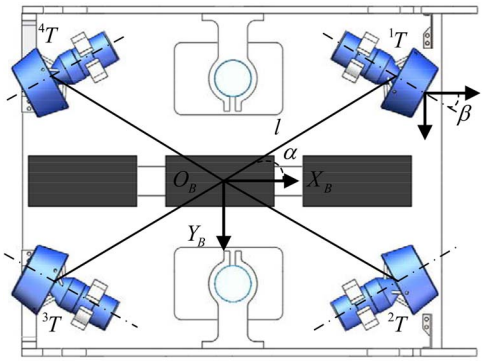

Fig. 4. Distribution of four horizontal thrusters in IWSCR.

TABLE I
RULES OF THRUSTER ALLOCATION

Force / Torque	Condition
τ_X^+	$(^1T = {}^2T) > (^3T = {}^4T)$
τ_X^-	$(^1T = {}^2T) < (^3T = {}^4T)$
τ_Y^+	$(^1T = {}^4T) > (^2T = {}^3T)$
τ_Y^-	$(^1T = {}^4T) < (^1T = {}^4T)$
τ_N^+	$(^1T = {}^3T) > (^2T = {}^4T)$
τ_N^-	$(^1T = {}^3T) < (^2T = {}^4T)$

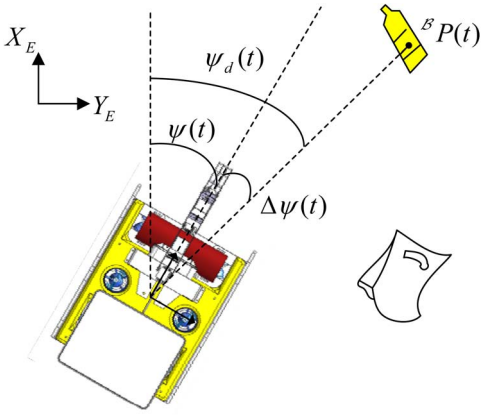

Fig. 5. Formulation of the control objective.

$$\tau_Y = \begin{bmatrix} \sin\beta \\ -\sin\beta \\ -\sin\beta \\ \sin\beta \end{bmatrix}^T \begin{bmatrix} ^1T \\ ^2T \\ ^3T \\ ^4T \end{bmatrix} \tag{9}$$

$$\tau_N = \begin{bmatrix} l\cos\alpha\sin\beta + l\cos\beta\sin\alpha \\ -l\cos\alpha\sin\beta - l\cos\beta\sin\alpha \\ l\cos\alpha\sin\beta + l\cos\beta\sin\alpha \\ -l\cos\alpha\sin\beta - l\cos\beta\sin\alpha \end{bmatrix}^T \begin{bmatrix} ^1T \\ ^2T \\ ^3T \\ ^4T \end{bmatrix}. \tag{10}$$

It is apparent that τ_X, τ_Y, and τ_N can be applied to the underwater vehicle independently with an appropriate combination of thrusts. The rules of thruster allocation are shown in Table I.

B. Formulation of the Vision-Based Steering

As shown in Fig. 5, inspired by [39], the steering objective on the water surface is expressed as follows:

$$\lim_{t\to\infty} \Delta\psi(t) \to 0$$
$$\Delta\psi(t) = \psi_d(t) - \psi(t)$$

where $\psi(t)$ and $\psi_d(t)$ denote the yaw angle of IWSCR and the target yaw angle at time t, respectively. Based on the technique of triangulation for binocular vision, the position of the object $^{\mathcal{C}}P(t)$ in the camera-fixed frame $\mathcal{C}$ can be obtained continually. When $^{\mathcal{C}}P(t)$ is transformed to frame $\mathcal{B}$ as $^{\mathcal{B}}P(t)$, the difference yaw angle $\Delta\psi(t)$ is derived as follows:

$$\Delta\psi(t) = \arctan\frac{^{\mathcal{B}}P(t)_y}{^{\mathcal{B}}P(t)_x}. \tag{11}$$

The estimated values of velocity u, v, and r are described as follows:

$$\begin{bmatrix} u \\ v \\ r \end{bmatrix} = \begin{bmatrix} -^{\mathcal{B}}\dot{P}(t)_x \\ -^{\mathcal{B}}\dot{P}(t)_y \\ -\Delta\dot{\psi}(t) \end{bmatrix}. \tag{12}$$

C. Design and Stability Analysis of the Sliding-Mode Controller

There are numerous disturbances, which are difficult to identify, against robot system in aquatic environment. Besides, the inaccurate measurement of vision system will lead disturbance to the input of the system. The SMC has the advantages of insensitivity about both of the variance of parameters and disturbance, which is appropriate for IWSCR. First, the sliding surface S is chosen as follows:

$$S = \lambda_0\Delta\psi(t) + \Delta\dot{\psi}(t) \tag{13}$$

where λ_0 is a positive constant. For succinct expression, we define $M_1 = I_z - N_{\dot{r}}$, $M_2 = m - Y_{\dot{v}}$, $M_3 = m - X_{\dot{u}}$, and $M_4 = N_r + N_{r|r|}|r|$. From the third equation of (6), the control law is constructed as follows:

$$\tau_N = M_1\left[\frac{(M_2 - M3)}{M_1}uv - \frac{M_4}{M_1} - \lambda_0 r\right] + C_0\text{sat}(S) \tag{14}$$

where C_0 is a positive value meeting $C_0 \geq |\tau_{EN}|$, and sat(S) is a saturation function to remove the chattering effect. The expression of saturation function is shown as follows:

$$\text{sat}(S) = \begin{cases} 1 & S > \sigma \\ \frac{S}{\sigma} & |S| < \sigma \\ -1 & S < -\sigma \end{cases} \tag{15}$$

where $\sigma \leq C_0$, a positive value, denotes the boundary layer thickness.

Next, the Lyapunov function $V = (1/2)S^2$ is employed to analyze the stability of the designed controller. The derivative Lyapunov candidate function is derived as follows:

$$\dot{V} = S\dot{S}. \tag{16}$$

Then, by substituting the expanded expression of $\dot{S}$, $\dot{V}$ can be written as follows:

$$\begin{aligned} \dot{V} &= S\left[\lambda_0\Delta\dot{\psi}(t) + \Delta\ddot{\psi}(t)\right] \\ &= -\frac{S}{M_1}[M_1\lambda_0 r - (M_2 - M3)uv \\ &\quad + M_4 r + (\tau_N + \tau_{EN})]. \end{aligned} \tag{17}$$

By substituting (14) into (17), $\dot{V}$ can be derived as follows:

$$\dot{V} = \frac{S}{M_1}[-C_0\text{sat}(S) - \tau_{EN}]. \tag{18}$$

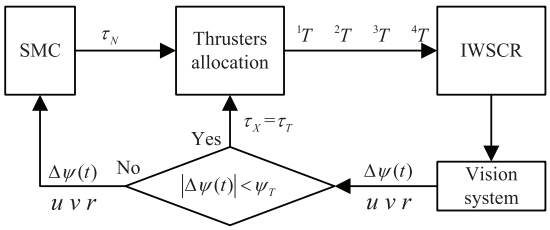

Fig. 6. Process of engineering realization.

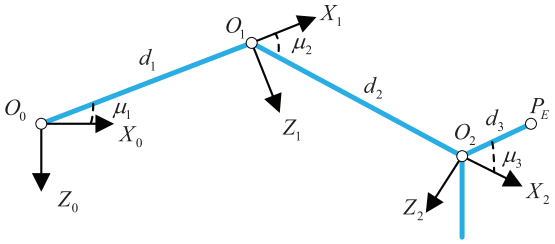

Fig. 7. Simplified schematic of link frames.

Consider the saturation function (15):

1) when $S > \sigma$, $\dot{V} = (S/M_1)[-C_0 - \tau_{EN}] \le 0$;
2) when $S < -\sigma$, $\dot{V} = (S/M_1)[C_0 - \tau_{EN}] \le 0$;
3) when $|S| < \sigma$, we suppose $|\tau_{EN}| \le \gamma$ and $\dot{V}$ can be derived as follows:

$$\begin{aligned}\dot{V} &= -\frac{C_0}{M_1\sigma}S^2 - \frac{S}{M_1}\tau_{EN} \\ &\le -\frac{C_0}{M_1\sigma}S^2 + \frac{|S|}{M_1}\gamma \\ &\le -\frac{C_0}{M_1\sigma}S^2 + \frac{1}{2M_1}S^2 + \frac{1}{2M_1}\gamma^2 \\ &= -\left(\frac{C_0}{M_1\sigma} - \frac{1}{2M_1}\right)S^2 + \frac{1}{2M_1}\gamma^2 \\ &= -\left(\frac{2C_0}{M_1\sigma} - \frac{1}{M_1}\right)V + \frac{1}{2M_1}\gamma^2 \\ &= -kV + \frac{1}{2M_1}\gamma^2 \end{aligned} \tag{19}$$

where $k > 0$.

Above all, the proposed controller (14) is uniformly ultimately bounded (UUB) [40], [41]. Therefore, the system is robust for the bounded unknown disturbance.

In addition, the process of engineering realization for vision-based steering is illustrated in Fig. 6. Here, ψ_T is the threshold value for whether the SMC would be enabled. If $|\Delta\psi(t)| \ge \psi_T$, SMC is applied to generate τ_N in order to reduce $|\Delta\psi(t)|$. If $|\Delta\psi(t)| < \psi_T$, the force τ_X is set to a constant τ_T to drive the underwater vehicle forward.

V. Dynamic Grasping Strategy for Floating Bottles

A. Kinematics and Inverse Kinematics of Manipulator

The manipulator of IWSCR is 3-DOF which can move on a longitudinal vertical plain parallel to the $X_BO_BZ_B$ plain in frame $\mathcal{B}$. The simplified schematic is shown in Fig. 7. The position of the end effector ${}^0P_E(p_x, p_y, 1)$ in the frame $O_0X_0Z_0$ can be obtained by

$$ {}^0P_E = A_1(\mu_1, d_1) \cdot A_2(\mu_2, d_2)^2P_E \tag{20}$$

where

$$A_i(\mu_i, d_i) = \begin{bmatrix} \cos\mu_i & \sin\mu_i & d_i \\ -\sin\mu_i & \cos\mu_i & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$${}^2P_E = \begin{bmatrix} d_3\cos\mu_3 \\ -d_3\sin\mu_3 \\ 1 \end{bmatrix}.$$

In practice, μ_1 and μ_3 are set to constants, so the inverse kinetic equation is as follows:

$$A_1^{-1}(\mu_1, d_1){}^0P_E = A_2(\mu_2, d_2)^2P_E. \tag{21}$$

According to the derivation of (21), μ_2 can be determined by an equation in the following form:

$$Q_{2\times2}\begin{bmatrix}\sin\mu_2 \\ \cos\mu_2\end{bmatrix} = V_{2\times1}. \tag{22}$$

Additionally, the range of μ_2 is $[-\pi, \pi]$, which ensures the unique solution of (22). In the process of grasp, when we obtain the position of object, the position and orientation of the manipulator are determined by (22) simultaneously.

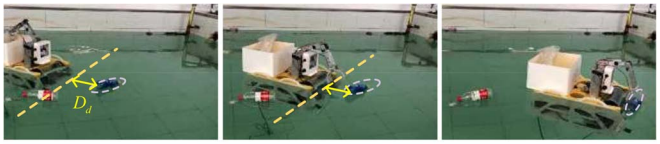

Fig. 8. Process of the proposed grasping strategy. (a) IWSCR navigates toward the bottle. (b) Manipulator prepares to grasp, when the distance equals D_d. (c) IWSCR continues moving and grasps the bottle.

B. Description of the Feasible Grasping Strategy

In the grasp task, plastic bag and styrofoam are easy to grasp due to their light weight and soft material. Therefore, the bag and styrofoam have many graspable positions. However, plastic bottle is difficult to grasp due to its cylindrical shape, which means the feasible graspable position of plastic bottle locates near the middle of the long axis of the bottle. In order to solve this problem, we proposed a pragmatic grasping strategy for plastic bottles based on the analysis of objects' stability in the fluid. When the vehicle moves toward the bottle, the direction of fluid velocity is from the vehicle to the bottle. In this situation, the bottle will rotate to a stable orientation, which is vertical to the direction of fluid velocity. Therefore, as shown in Fig. 8, IWSCR should prepare to grasp in advance, when the distance between IWSCR and the object equals decision distance D_d. Note that D_d is an empirical value related to the present surge speed u and the preparation time T_p, which can be obtained as follows:

$$D_d = uT_p. \tag{23}$$

Then, IWSCR continues moving and grasps the bottle eventually.

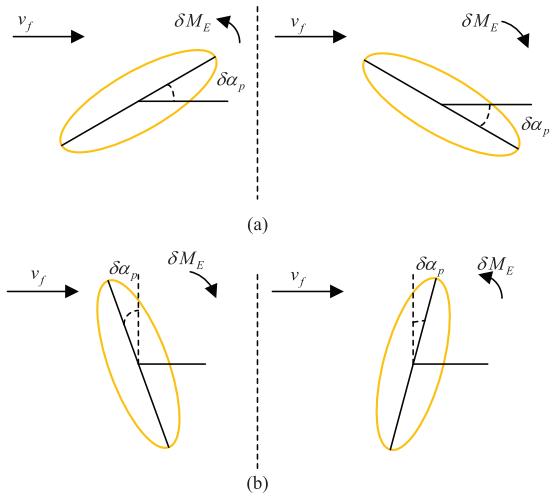

Fig. 9. Stability of the two equilibrium states. (a) $\alpha_p = 0$. (b) $\alpha_p = (1/2)\pi$.

Next, the proposed strategy is expounded in theory. Suppose that the plastic bottle is abstracted as a prolate ellipse on the water surface, that the water velocity v_f is constant, and that the fluid is incompressible. Given the included angle between the long axis of ellipse and the direction of water velocity α_p, refer to [42], the total torque applied on the ellipse is derived as follows:

$$M_E = \frac{1}{2}\pi\rho\left(a^2 - b^2\right)v_f^2 \sin 2\alpha_p \tag{24}$$

where ρ denotes the fluid density, and a and b are the semi-major axis and the semiminor axis, respectively. Note that $M_E = 0$ when $\alpha_p = 0$ and $\alpha_p = \pi/2$, which represents that the bottle locates at the equilibrium position. However, the stabilities of the two equilibrium states are different.

As to $\alpha_p = 0$, when a disturbance results into an infinitesimal angle $\delta\alpha_p$ adding to the included angle α_p, as shown in Fig. 9(a), the torque applied on the ellipse is as follows:

$$\begin{aligned}\delta M_E|_{\alpha_p=0} &= \pi\rho\left(a^2 - b^2\right)v_f^2\delta\alpha_p \cos 2\alpha_p|_{\alpha_p=0} \\ &= \pi\rho\left(a^2 - b^2\right)v_f^2\delta\alpha_p\end{aligned} \tag{25}$$

where $\delta M_E \cdot \delta\alpha_p > 0$. It means that α_p will increase, so the equilibrium is not stable. Similarly, as $\alpha_p = \pi/2$, shown in Fig. 9(b), the torque applied on the ellipse is as follows:

$$\begin{aligned}\delta M_E|_{\alpha_p=\frac{1}{2}\pi} &= \pi\rho\left(a^2 - b^2\right)v_f^2\delta\alpha_p \cos 2\alpha_p|_{\alpha_p=\frac{1}{2}\pi} \\ &= -\pi\rho\left(a^2 - b^2\right)v_f^2\delta\alpha_p\end{aligned} \tag{26}$$

where $\delta M_E \cdot \delta\alpha_p < 0$. It means that α_p will decrease, so the equilibrium is stable.

Above all, the long axis of plastic bottle will be vertical to the direction of the water velocity, which demonstrates the feasibility of the proposed grasping strategy.

VI. Experimental Results and Discussion

In this section, experiments of garbage detection and SMC for vision-based steering are concretely introduced. The whole experimental results of IWSCR to accomplish TTs are exhibited here.

TABLE II
Results of Garbage Detection

Class	Accuracy (mAP)	Speed (fps)
Plastic bottle	0.8799	–
Plastic bag	0.9376	–
Styrofoam	0.9182	–
Total	0.9119	45.9036

TABLE III
Model Parameters for Simulation

Parameter	Value	Parameter	Value
m	25	$X_{\dot{u}}$	–28.5
I_z	20.4	$Y_{\dot{v}}$	–32.5
X_u	–64.3	$X_{u\|u\|}$	–23.6
Y_v	–70.5	$Y_{v\|v\|}$	–42.5
N_r	–55.1	$N_{\dot{r}}$	–10.4
$N_{r\|r\|}$	–18.4	–	–

A. Experimental Results of Garbage Detection

To realize the garbage detection, a floating garbage dataset (FGD) is established, which includes 1000 images covering various plastic bottles, plastic bags, and styrofoam under different illumination. FGD is divided into a train dataset (TD) and a verification dataset (VD). The GPU in this experiment is NVDIA-1080. YOLOv3 is trained on TD, and the results on VD are shown in Table II. The accuracy is described with the mAP, and the computing speed is evaluated by frames per second (fps). Note that the detection accuracy is commendable and the speed meets the requirement of real-timeness. However, the detection accuracy of the plastic bottle is bit lower than accuracy of bag and styrofoam, for the packagings of plastic bottles are various.

B. Experimental Results of SMC for Vision-Based Steering and Achievement of TTs

The major advantage of SMC is that it is robust to disturbance led to input. Therefore, related simulation experiments based on the underwater vehicle model for the process of steering on the water surface are carried out. In the simulation experiments, the model parameters of IWSCR are listed in Table III. Note that $\Delta\psi(0) = 30°$ is set in the series experiments, and that the disturbances τ_{EN} are sine waves with different amplitudes led to the input. The comparison experimental results about the antijamming capability between SMC and PI are shown in Fig. 10. When no disturbance impacts the system, the proposed controller and PI have good steady-state performance. However, the PI wastes longer convergence time than the proposed controller. When a sinusoidal disturbance occurs on the system, the proposed controller is more robust than PI. To describe the antijamming capability of the two controllers more intuitively, we use the root-mean-square error (RMSE) to evaluate the steady-state error. As shown in Fig. 10(c), with the disturbance intensities increasing, the RMSEs of the two controllers increase. It is apparent that the RMSE of the proposed controller increases slowly, which demonstrates the SMC's insensibility to disturbance. Additionally, it is difficult for the PI controller to avoid overshoot, which is a fatal problem for vision-based steering,

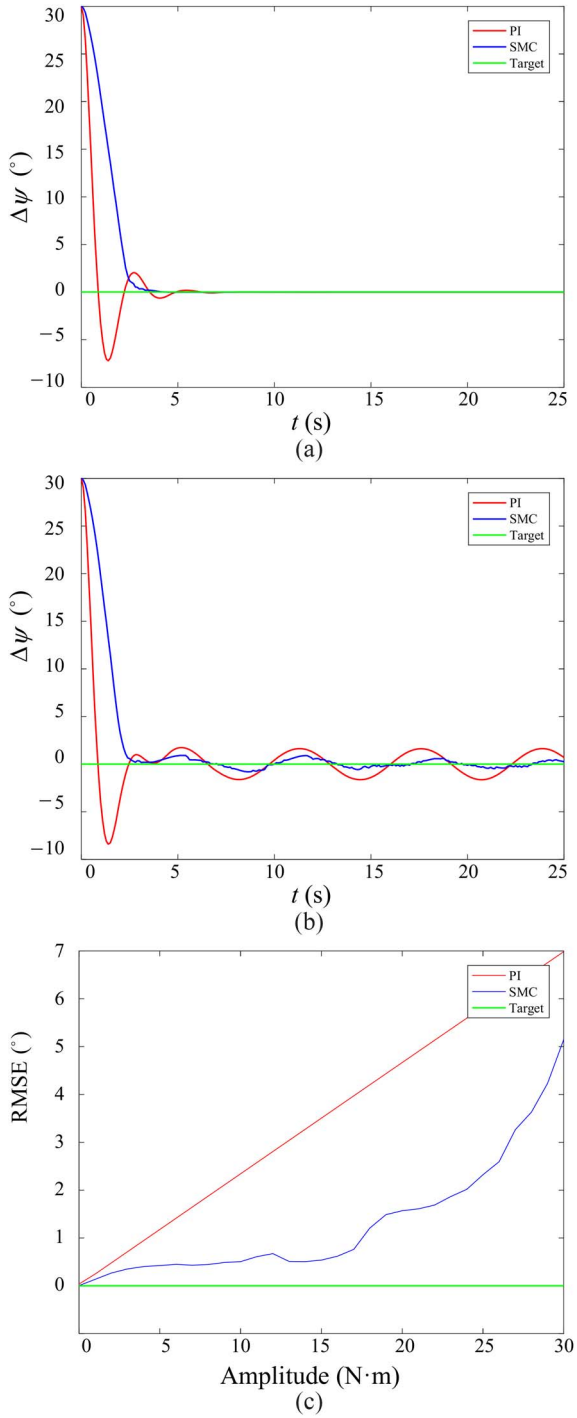

Fig. 10. Comparison experiments between SMC and PI. (a) Without disturbance. (b) Amplitude of disturbance is 5 N·m. (c) Relationship between RMSE and amplitudes of disturbance.

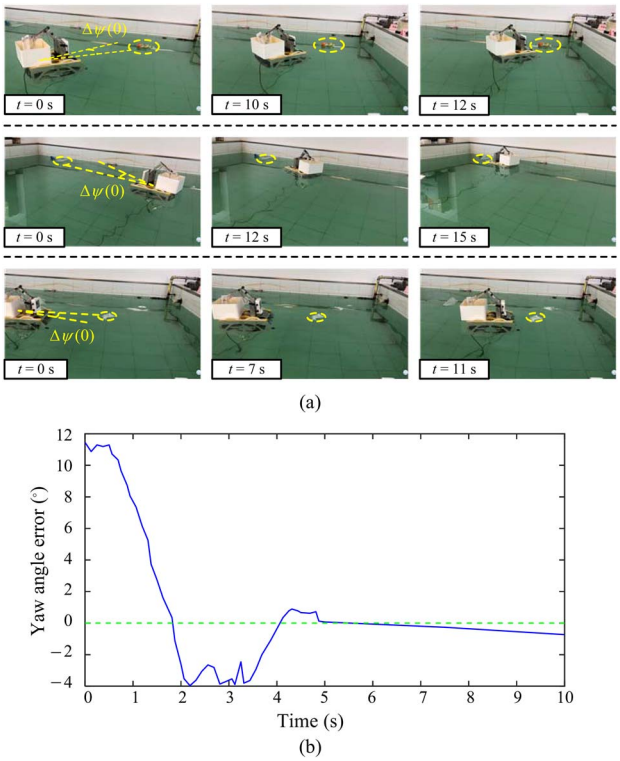

Fig. 11. Experiments on vision-based steering in real environment. (a) Video snaps of the vision-based steering with three different initial relative angles. (b) Tendency curve of yaw angle error deviation of the third experiment.

because the overshoot may cause the loss of targets in practice. Above all, the results from the simulation experiments demonstrate the superiority of SMC applied on IWSCR.

Based on the prior simulations, three experiments of vision-based steering of IWSCR in the real laboratory environment are carried out as shown in Fig. 11. The target in each experiment locates by different initial angles relative to IWSCR as shown in Fig. 11(a), and the tendency of the yaw angle error deviation obtained from the binocular vision sensor for the third experiment is described in Fig. 11(b). Notice that the overshoot exists in the experiment due to the vision measurement errors and the imprecise model parameters. In conclusion, the capability of IWSCR to approach the target demonstrates the feasibility of the proposed control law.

By applying the detection framework, SMC, and the feasible grasping strategy on IWSCR, the performance of autonomously accomplishing TTs is shown in Fig. 12. For a better comprehension, the video is captured by two cameras, i.e., an external camera and IWSCR's binocular camera. For this experiment, the target garbage is a plastic bottle. As shown in Fig. 12(a), IWSCR is cruising on the water surface and detecting garbage by bounding boxes in real time. Due to the objects are plastic bags and styrofoam, IWSCR does not conduct the tracking and steering task. When the plastic bottle is detected, the target is determined. Fig. 12(b) shows that the vision-based steering starts after the target has been locked. Finally, the process of grasp and collection is recorded in Fig. 12(c). The video snapshots demonstrate the feasibility of IWSCR to clean the water surface.

To further evaluate the adaptability of IWSCR, a field experiment of autonomous garbage cleaning was conducted in a reservoir. The performance of IWSCR is shown in Fig. 13, where Fig. 13(a) records the video snapshots of the cleaning process and Fig. 13(b) plots the tendency curves for yaw angle error deviation and the distance deviation measured by the binocular vision system. Notice that at $t = 11$ s, IWSCR got into the grasping preparation stage, thus the visual feedback data was interrupted. An interesting phenomenon happened

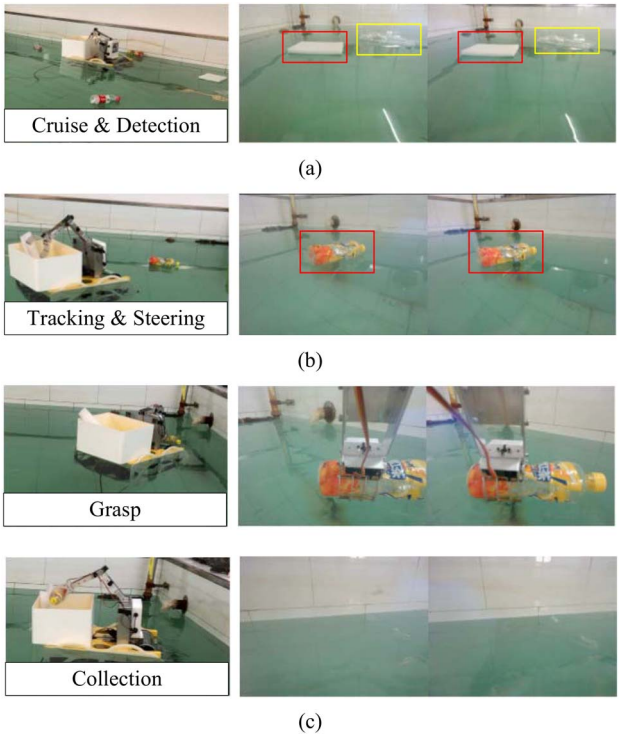

Fig. 12. Performance of accomplishing TTs. (a) Cruise and detection. (b) Tracking and steering. (c) Grasping and collection.

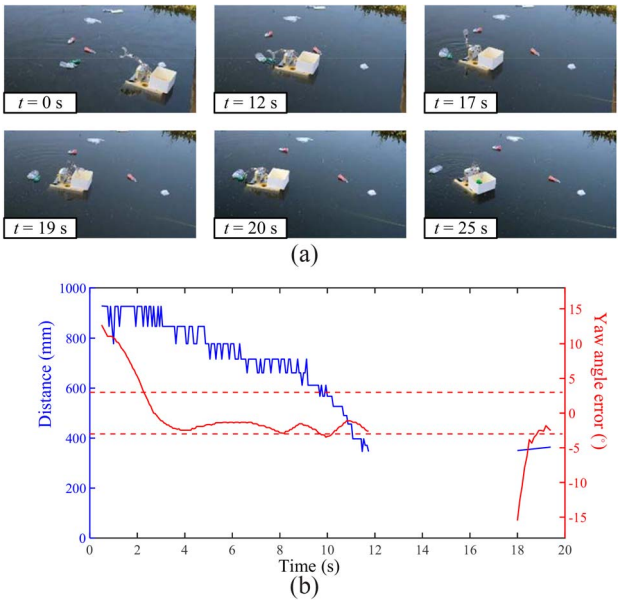

Fig. 13. Performance of the field experiment. (a) Video snaps of the whole cleaning process. (b) Tendency curve of the yaw angle error and distance deviations.

in the cleaning process, that the first grasping attempt failed but IWSCR detected the target again soon at $t = 18$ s and performed the second grasping successfully. Consequently, the field environmental results demonstrate a good utilization prospect of IWSCR for the aquatic environment protection.

C. Discussion

According to the aforementioned experimental results of IWSCR, our prototype can successfully realize garbage detection, vision-based steering, and water surface grasp in the laboratory environment. During implementation, we overcome three major problems. The first one is how to increase the speed of detection, for which we employ the YOLOv3 network to achieve a high accurate and real-time detection whose mAP and speed are up to 0.91 and 45.9 frames/s, respectively. The second one is how to resist the disturbance in vision-based steering, as to which a control law derived by SMC is proposed. The third one is how to grasp plastic bottles floating on the surface.

The development of an intelligent robot for cleaning water surface is a sophisticated multidisciplinary integration systematic engineering project with numerous challenges. With respect to the visual techniques, we employed here for detecting and recognizing objects as garbage, which is the basis of the cleaning, we established specific dataset for garbage floating on the water surface. However, the generality of this dataset is questionable. Furthermore, there are still many aspects to be improved, such as tracking accuracy and match accuracy for triangulation. With respect to controller design, we only tested IWSCR in the laboratory and relatively calm field environments. Whereas, the expanse water possesses more noises and poses more physical disturbances and other unpredictable influences. Therefore, designing a more robust controller is necessary for achievement in more complicated real-world settings. With respect to grasping, the proposed strategy is just for specific floating objects. It is necessary to investigate the dynamic grasping on the water surface further to form a general method.

VII. Conclusion

In this article, IWSCR has been creatively designed to collect plastic garbage on the water surface. To tackle the three challenges declared in Section I, we proposed and implemented feasible methodology on IWSCR. First, the YOLOv3 network is trained on the proposed FGD for garbage detection to ensure the high accuracy and high speed. Second, an SMC-based control law is applied to improve the capacity of resisting disturbance for IWSCR. Third, a feasible grasping strategy inspired from the stability of floating object in fluid is proposed, which provides a novel pathway for dynamic grasping on the water surface. With the aforementioned three techniques, IWSCR possesses the capability to accomplish TTs, and it becomes a pragmatic water surface cleaner.

In the future, IWSCR will be tested in larger scale and more complicated field environments, which needs much more robust and stable techniques for vision module, motion control module, and grasping module.

References

[1] J. R. Jambeck *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, J. W. Damsteeg, B. Slat, A. Andrady, and J. Reisser, "River plastic emissions to the world's oceans," *Nat. Commun.*, vol. 8, Jun. 2017, Art. no. 15611.

[3] J. Palacin, J. A. Salse, I. Valganon, and X. Clua, "Building a mobile robot for a floor-cleaning operation in domestic environments," *IEEE Trans. Instrum. Meas.*, vol. 53, no. 5, pp. 1418–1424, Oct. 2004.

[4] M.-C. Kang, K.-S. Kim, D.-K. Noh, J.-W. Han, and S.-J. Ko, "A robust obstacle detection method for robotic vacuum cleaners," *IEEE Trans. Consum. Electron.*, vol. 60, no. 4, pp. 587–595, Nov. 2014.

[5] Y. Fu-Cai, H. Shi-Jian, S. Hai-Liang, and W. Li-Zhu, "Design of cleaning robot for swimming pools," in *Proc. Int. Conf. Manage. Sci. Ind. Eng.*, Harbin, China, Jan. 2011, pp. 1175–1178.

[6] C. Su, W. Dongxing, L. Tiansong, R. Weichong, and Z. Yachao, "An autonomous ship for cleaning the garbage floating on a lake," in *Proc. 2nd Int. Conf. Intell. Comput. Technol. Autom.*, Changsha, China, Oct. 2009, pp. 471–474.

[7] H. Zhang, J. Zhang, G. Zong, W. Wang, and R. Liu, "Sky cleaner 3: A real pneumatic climbing robot for glass-wall cleaning," *IEEE Robot. Autom. Mag.*, vol. 13, no. 1, pp. 32–41, Mar. 2006.

[8] G. Ferri, A. Manzi, P. Salvini, B. Mazzolai, C. Laschi, and P. Dario, "DustCart, an autonomous robot for door-to-door garbage collection: From DustBot project to the experimentation in the small town of Peccioli," in *Proc. IEEE Int. Conf. Robot. Autom.*, Shanghai, China, May 2011, pp. 655–660.

[9] S. Watanasophon and S. Ouitrakul, "Garbage collection robot on the beach using wireless communications," in *Proc. 3rd Int. Conf. Informat. Environ. Energy Appl.*, Singapore, Mar. 2014, pp. 92–96.

[10] J. Bai, S. Lian, Z. Liu, K. Wang, and D. Liu, "Deep learning based robot for automatically picking up garbage on the grass," *IEEE Trans. Consum. Electron.*, vol. 64, no. 3, pp. 382–389, Aug. 2018.

[11] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015.

[12] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Columbus, OH, USA, Jun. 2014, pp. 580–587.

[13] R. Girshick, "Fast R-CNN," in *Proc. IEEE Int. Conf. Comput. Vis.*, Santiago, Chile, Dec. 2015, pp. 1440–1448.

[14] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," in *Proc. Adv. Neural Inf. Process. Syst.*, Montreal, QC, Canada, Dec. 2015, pp. 91–99.

[15] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask R-CNN," in *Proc. Int. Conf. Comput. Vis.*, Venice, Italy, Oct. 2017, pp. 2961–2969.

[16] W. Liu *et al.*, "SSD: Single shot multibox detector," in *Proc. Eur. Conf. Comput. Vis.*, Amsterdam, The Netherlands, Oct. 2016, pp. 21–37.

[17] X. Chen, Z. Wu, and J. Yu, "TSSD: Temporal single-shot detector based on attention and LSTM," in *Proc. IEEE/RSJ Int. Conf. Intell. Robot. Syst.*, Madrid, Spain, Oct. 2018, pp. 5758–5763.

[18] J. Redmon, S. K. Divvala, R. B. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Las Vegas, NV, USA, Jun. 2016, pp. 779–788.

[19] J. Redmon and A. Farhadi, "YOLO9000: Better, faster, stronger," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Honolulu, HI, USA, Nov. 2017, pp. 6517–6525.

[20] J. Redmon and A. Farhadi, "YOLOv3: An incremental improvement," *arXiv:1804.02767*, 2018. [Online]. Available: https://arxiv.org/abs/1804.02767

[21] R. Cui, X. Zhang, and D. Cui, "Adaptive sliding-mode attitude control for autonomous underwater vehicles with input nonlinearities," *Ocean Eng.*, vol. 123, pp. 45–54, Sep. 2016.

[22] Z.-P. Yan, H. Ju, and S.-P. Hou, "Diving control of underactuated unmanned undersea vehicle using integral-fast terminal sliding mode control," *J. Central South Univ.*, vol. 23, no. 5, pp. 1085–1094, May 2016.

[23] H. Zhou, K. Liu, Y. Li, and S. Ren, "Dynamic sliding mode control based on multi-model switching laws for the depth control of an autonomous underwater vehicle," *Int. J. Adv. Robot. Syst.*, vol. 12, no. 7, pp. 1–10, Jul. 2015.

[24] R. Cui, L. Chen, C. Yang, and M. Chen, "Extended state observer-based integral sliding mode control for an underwater robot with unknown disturbances and uncertain nonlinearities," *IEEE Trans. Ind. Electron.*, vol. 64, no. 8, pp. 6785–6795, Aug. 2017.

[25] J. Yu, J. Liu, Z. Wu, and H. Fang, "Depth control of a bioinspired robotic dolphin based on sliding-mode fuzzy control method," *IEEE Trans. Ind. Electron.*, vol. 64, no. 3, pp. 2429–2438, Mar. 2018.

[26] G. P. Incremona, G. D. Felici, A. Ferrara, and E. Bassi, "A supervisory sliding mode control approach for cooperative robotic system of systems," *IEEE Syst. J.*, vol. 9, no. 1, pp. 263–272, Mar. 2015.

[27] I. Lenz, H. Lee, and A. Saxena, "Deep learning for detecting robotic grasps," *Int. J. Robot. Res.*, vol. 34, nos. 4–5, pp. 705–724, Mar. 2015.

[28] S. Kumra and C. Kannan, "Robotic grasp detection using deep convolutional neural networks," in *Proc. IEEE/RSJ Int. Conf. Intell. Robot. Syst.*, Vancouver, BC, Canada, Sep. 2017, pp. 769–776.

[29] J. Redmon and A. Angelova, "Real-time grasp detection using convolutional neural networks," in *Proc. Int. Conf. Robot. Autom.*, Seattle, WA, USA, May 2015, pp. 1316–1322.

[30] 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, Jun. 2017.

[31] E. Johns, S. Leutenegger, and A. J. Davison, "Deep learning a grasp function for grasping under gripper pose uncertainty," in *Proc. IEEE/RSJ Int. Conf. Intell. Robot. Syst.*, Daejeon, South Korea, Oct. 2016, pp. 4461–4468.

[32] U. Asif, J. Tang, and S. Harrer, "GraspNet: An efficient convolutional neural network for real-time grasp detection for low-powered devices," in *Proc. Int. Joint Conf. Artif. Intell.*, Stockholm, Sweden, Jul. 2018, pp. 4875–4882.

[33] M. A. Goodale, J. P. Meenan, H. H. Bülthoff, D. A. Nicolle, K. J. Murphy, and C. I. Racicot, "Separate neural pathways for the visual analysis of object shape in perception and prehension," *Current Biol.*, vol. 4, no. 7, pp. 604–610, Jul. 1994.

[34] R. L. Whitwell, A. D. Milner, and M. A. Goodale, "The two visual systems hypothesis: New challenges and insights from visual form agnosic patient DF," *Front. Neurol.*, vol. 5, p. 255, Dec. 2014.

[35] T. Fossen, *Guidance and Control of Ocean Vehicles*. New York, NY, USA: Wiley, 1994.

[36] J. F. Henriques, R. Caseiro, P. Martins, and J. Batista, "High-speed tracking with kernelized correlation filters," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 3, pp. 583–596, Mar. 2015.

[37] G. C. Karras, P. Marantos, C. P. Bechlioulis, and K. J. Kyriakopoulos, "Unsupervised online system identification for underwater robotic vehicles," *IEEE J. Ocean. Eng.*, vol. 44, no. 3, pp. 642–663, Jul. 2019.

[38] Z. Peng and J. Wang, "Output-feedback path-following control of autonomous underwater vehicles based on an extended state observer and projection neural networks," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 48, no. 4, pp. 535–544, Apr. 2018.

[39] F. Ke, Z. Li, H. Xiao, and X. Zhang, "Visual servoing of constrained mobile robots based on model predictive control," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 47, no. 7, pp. 1428–1438, Jul. 2017.

[40] J. Sun, Z. Pu, and J. Yi, "Conditional disturbance negation based active disturbance rejection control for hypersonic vehicles," *Control Eng. Pract.*, vol. 84, pp. 159–171, Mar. 2019.

[41] J. Sun, J. Yi, Z. Pu, and X. Tan, "Fixed-time sliding mode disturbance observer-based nonsmooth backstepping control for hypersonic vehicles," *IEEE Trans. Syst., Man, Cybern., Syst.*, to be published, doi: 10.1109/TSMC.2018.2847706.

[42] L. M. Milne-Thomson, *Theoretical Hydrodynamics*. New York, NY, USA: Dover, 1996.

Shihan Kong received the B.E. degree in automation from the School of Control Science and Engineering, Shandong University, Jinan, China, in 2016. He is currently pursuing the Ph.D. degree in control theory and control engineering with the Institute of Automation, Chinese Academy of Sciences, Beijing, China.

His current research interests include underwater robotics and underwater robot vision.

Manjun Tian received the B.E. degree in Internet of Things from the School of Computer and Information Technology, Northeast Petroleum University, Daqing, China, in 2018. He is currently pursuing the master's degree in computer science and technology with the School of Information Engineering, Minzu University of China, Beijing, China.

His current research interests include robotic vision and image processing.

Changlin Qiu received the B.E. degree in automation from the School of Control Science and Engineering, Shandong University, Jinan, China, in 2019. He is currently pursuing the Ph.D. degree in control theory and control engineering with the Institute of Automation, Chinese Academy of Sciences, Beijing, China.

His current research interests include underwater robotics and underwater robot vision.

Zhengxing Wu received the B.E. degree in logistics engineering from the School of Control Science and Engineering, Shandong University, Jinan, China, in 2008, and the Ph.D. degree in control theory and control engineering from the Institute of Automation, Chinese Academy of Sciences (IACAS), Beijing, China, in 2015.

He is currently an Associate Professor with the State Key Laboratory of Management and Control for Complex Systems, IACAS. His current research interests include fast maneuvers of bio-inspired robotic fish and gliding motions of robotic dolphins.

Junzhi Yu (Senior Member, IEEE) received the B.E. degree in safety engineering and the M.E. degree in precision instruments and mechanology from the North University of China, Taiyuan, China, in 1998 and 2001, respectively, and the Ph.D. degree in control theory and control engineering from the Institute of Automation, Chinese Academy of Sciences, Beijing, China, in 2003.

From 2004 to 2006, he was a Postdoctoral Research Fellow with the Center for Systems and Control, Peking University, Beijing. He was an Associate Professor with the Institute of Automation, Chinese Academy of Sciences in 2006, where he was a Full Professor in 2012. In 2018, he joined the College of Engineering, Peking University as a Tenured Full Professor. His current research interests include intelligent robots, motion control, and intelligent mechatronic systems.