

Article

Binocular Vision-Based Non-Singular Fast Terminal Control for the UVMS Small Target Grasp

Tao Jiang ¹, Yize Sun ¹, Hai Huang ^{1,*}, Hongde Qin ¹, Xi Chen ², Lingyu Li ¹, Zongyu Zhang ¹ and Xinyue Han ¹

¹ National Key Laboratory of Science and Technology on Underwater Vehicle, Harbin Engineering University, Harbin 150001, China; chuanbojiangtao@hrbeu.edu.cn (T.J.); sunyize@hrbeu.edu.cn (Y.S.); qinhd@hrbeu.edu.cn (H.Q.); lilingyu@hrbeu.edu.cn (L.L.); loner8701yuanfang@163.com (Z.Z.); hanxinyue24@foxmail.com (X.H.)

² College of Mechanical and Electrical Engineering, Heilongjiang Institute of Technology, Harbin 150050, China; chenxi_1113652@163.com

* Correspondence: haihus@163.com

Abstract: Autonomous underwater manipulation is very important for the robotic and intelligence operations of oceanic engineering. However, a small target often involves limited features and results in inaccurate visual matching. In order to improve visual measurement accuracy, this paper has proposed an improved unsharp masking algorithm to further enhance the weak texture region of blurred and low contrast images. Moreover, an improved ORB feature-matching method with adaptive threshold, non-maximum suppression and improved random sample consensus has also been proposed. To overcome unknown underwater disturbances and uncertain system parameters in the underwater robotic manipulations, an adaptive non-singular terminal sliding mode controller has been proposed with a quasi-barrier function to suppress the chattering problem and improve grasp accuracy for small target. Oceanic experiments have been conducted to prove the performance of the proposed method.

Keywords: binocular visual matching; underwater vehicle manipulator system; non-singular fast terminal control; underwater target grasp



Citation: Jiang, T.; Sun, Y.; Huang, H.; Qin, H.; Chen, X.; Li, L.; Zhang, Z.; Han, X. Binocular Vision-Based Non-Singular Fast Terminal Control for the UVMS Small Target Grasp. *J. Mar. Sci. Eng.* **2023**, *11*, 1905. <https://doi.org/10.3390/jmse11101905>

Academic Editor: Xianbo Xiang

Received: 8 August 2023

Revised: 20 September 2023

Accepted: 26 September 2023

Published: 30 September 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

For the time being, the implementation of underwater manipulation missions through underwater robots have attracted increasing attention worldwide [1–3]. Their missions and applications include environmental detection, oceanic engineering, structural inspections, seabed surveys, scientific explorations, mine extractions, wreck recoveries, etc. [4] However, most underwater manipulation tasks are still carried out through working class remotely operated vehicles (ROVs) and divers [5,6]. Moreover, the complicated and accurate manipulation operations depend not only on the pilots' skills and manipulation control [7], but also on the performance of the sensor equipment and their process algorithms [8]. Therefore, underwater autonomous or semi-autonomous operations have become one of the most important solutions for the accurate manipulation of an underwater vehicle manipulator system (UVMS) [9,10]. In order to realize the autonomous manipulation, the UVMS should not only perceive the target through visual recognition, tracking, and distance measurement, but also carry out motion planning and control for the multiple DOFs (degree of freedoms) UVMS [11,12]. Among these challenges, binocular vision distance measurement and manipulation control are very important for the integration and coordination between sensing and motions [13–15].

Regarding binocular vision distance measurement guided for manipulation [16,17], the underwater scenario is influenced by scattering and absorption with loss of color detail [18]. The distance measurement process includes binocular vision calibration [19],

rectification [20] and matching [21]. In order to realize visual-guided underwater missions, Guanying Huo et al. [22] presented a binocular vision-based underwater target detection and 3D reconstruction system. Guided by target detection, the article improved the semi-global binocular matching method through matching the valid target area and optimizing the basic disparity map. Although many articles have focused on binocular vision [23,24], the fast and precise visual matching for a small object is usually influenced by multiple refraction times and ineffective imaging from the underwater environment. Stereo matching is one of the major steps in determining the depth [21], but conventional image matching strategies are less effective because of underwater refractive distortions [25]. Yaqian Li et al. [26] proposed an adaptive window based on mean-shift segmentation to alleviate distortion. In order to realize visual-guided manipulation for small objects, fast and precise vision matching is very important for distance measurement. The precise distance measurement needs to pick up enough feature points for correct matching in real-time. With standard-oriented FAST (features from accelerated segment test) [27] detecting the key points, the ORB (oriented FAST and rotated BRIEF (binary robust independent elementary feature)) [28] method is a very fast and efficient binary descriptor. But small objects in the underwater environment often include limited features and cause unstable matching. A. Osipov et al. [29] achieved high-precision detection and classification through improved YOLOv4-tiny and CNN algorithms.

In order to realize motion control for underwater robots [30], Satja Sivčev et al. [17] presented a vision-based kinematics control method for a working class ROV. Bent Oddvar Arnesen Haugaløkken et al. [31] developed a sliding mode kinematic control for UVMS grasp on the basis of monocular vision. In order to alleviate control chattering and assure finite-time convergence, Mingxue Cai et al. [32] proposed an improved nonsingular terminal sliding mode control on the basis of an underwater biomimetic vehicle manipulator system. Li and Huang et al. [33] proposed a nonlinear adaptive model to estimate camera parameters and improve visual serving convergence, with a disturbance observer. In comparison with other sliding mode control methods, the nonsingular terminal sliding mode control has become more and more popular for high-precision robotic control [34]. But these control methods assume a prior knowledge of the upper bound, which may not easily be obtained due to the model complexity and uncertainties. Mohamed Boukattaya et al. [35] proposed an adaptive nonsingular fast terminal sliding-mode control for the unknown upper-bound estimation of uncertain dynamical systems; the concept of adaptive design in it is excellent, which has inspired us in our own design. Except for the vision-based manipulation control of UVMS, the view field is limited due to the close visible distance; moreover, vehicle navigation systems, such as magnetic compass and Doppler velocity log (DVL), cannot ensure high accuracy small boundary reciprocating motions for the long term. The controller should further suppress vibration problems to keep the target in sight. Therefore, a design obstacle function is formulated, combined with the concept of adaptive design [35], to develop control laws that are more suitable for UVMS operations.

The main contribution of this paper includes the following points:

- (1) On the underwater visual matching of blurred and low contrast images, an improved unsharp masking algorithm is proposed to further enhance the weak texture region. The improved ORB feature-matching method is developed with adaptive threshold, non-maximum suppression and improved random sample consensus (RANSAC). Adaptive threshold is adopted to extract FAST feature points, non-maximum suppression is performed to remove feature point blocks, which can reduce the number of invalid feature points and save matching time, and the RANSAC algorithm can strengthen the elimination of mismatch points.
- (2) A novel adaptive nonsingular terminal sliding mode controller is proposed, incorporating the obstacle function, to enhance its applicability in UVMS operations, specifically regarding unknown disturbances and uncertain parameters with unknown upper bounds. The obstacle function is designed using a quasi-potential barrier func-

tion to suppress chattering issues, and a feedforward strategy is employed to handle the coupling of the robotic arm.

- (3) Oceanic experiments have been conducted to prove the performance of the proposed algorithm and controller.

The rest of this article is organized as follows. The binocular vision matching method for underwater small objects will be outlined in Section 2. Section 3 will investigate the adaptive non-singular fast terminal controller. The results of our simulations and experiments will be discussed in Section 4. Section 5 will provide our conclusions.

2. Binocular Vision Matching Method for Underwater Small Object

Binocular vision can provide three-dimensional distance information for manipulation control [36]. Since underwater visual images are usually grayed, obscured and contrast reduced [37,38], it is very important to enhance the texture during image preprocessing. From the target position information obtained from binocular vision, the UVMS can realize position-based visual serving control for the manipulation. Figure 1 presents the block diagram of the system described in this paper.

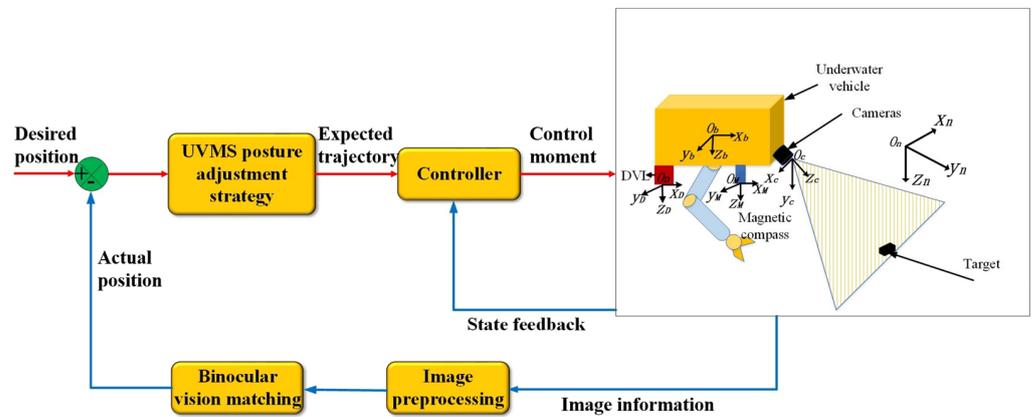


Figure 1. The process of binocular vision–based manipulation control.

2.1. Image Preprocessing

In order to preserve edge information and eliminate image noise for underwater small object visual matching, the fast guide filter algorithm was applied. The linear relationship between output image q_i and guiding image g_i in the window of ω_k is:

$$q_i = a_k g_i + b_k, \forall i \in \omega_k \tag{1}$$

To minimize the difference between the output image q_i and input image p_i , the cost function of window coefficient can be expressed as:

$$E(a_k, b_k) = \sum_{i \in \omega_k} ((a_k g_i + b_k - p_i)^2 + \epsilon a_k^2) \tag{2}$$

where the linear coefficient $a_k = \frac{cov(p_k, g_k)}{\sigma_k^2 + \epsilon}$ and $b_k = \overline{p_k} - \mu_k a_k$.

To further enhance the weak texture region, an improved unsharp masking algorithm has been proposed. After evaluating the high frequency component, enhancing the edge and texture details and establishing the gain functions, the equation for the enhanced image can be expressed as:

$$F(x_T, y_T) = I_T(x_T, y_T) + K_T \lambda_T(x_T, y_T) \times D_T(x_T, y_T) \tag{3}$$

$$\text{where } D_T(x_T, y_T) = \begin{cases} I_T(x_T, y_T)(1 - \frac{1}{2\pi_T\sigma_T^2}) & \text{if } |D_T(x_T, y_T)| \geq \text{Max}|D_T(x_T, y_T)| \\ 0 & \text{otherwise} \end{cases},$$

$I_T(x_T, y_T)$ is original image; $\lambda(x_T, y_T)$ is gain function; K is the coefficient to tune the high frequency information intensity.

$$\begin{cases} \lambda_T(x_T, y_T) = \beta_T \times L_C(x, y) + (1 - \beta_T) \times L_v(x_T, y_T) \\ L_C(x_T, y_T) = \frac{\sum_{k_T=0}^{L_T-1} \text{sgn}(k_T)}{G_{Lc}} \\ L_v(x_T, y_T) = \frac{\frac{1}{m \times n} (\sum_{i=x-(m-1)/2}^{i=x+(m-1)/2} \sum_{j=y-(n-1)/2}^{j=y+(n-1)/2} (f_T(i, j) - \bar{f}_T)^2)}{G_{Lv}}^{1/2} \end{cases} \quad (4)$$

where $L_C(x_T, y_T)$ is local complexity with G_{Lc} as global complexity; $L_v(x_T, y_T)$ is local standard deviation, with G_{Lv} as global standard deviation; \bar{f}_T is the average of domain $\Omega(m \times n)$ with center (x_T, y_T) ; β_T is the weight value between local standard deviation and local complexity.

2.2. ORB Algorithm for Binocular Vision Matching

In order to pick up feature points and guarantee matching speed for real time grasp, the ORB algorithm [38] was improved. The ORB algorithm includes four steps:

- (1) Scale space construction:

Gauss convolution is applied to pick up feature points with scale invariance. Octave and intra-octave layers are constructed through original image down sampling.

- (2) Candidate key points are picked up through FAST-9 and ascertained through Harris sort:

In order to select 16 pixels on a circle around every feature point, the FAST feature points are extracted. If more than 3/4 pixels are heavier or lighter than the threshold, the feature point is appropriate. Harris response values are calculated with the top N key response points.

- (3) Estimate the major orientation with intensity centroid method. The moment centroid is:

$$C_T = (\frac{m_{10}}{m_{00}}, \frac{m_{01}}{m_{00}}) \quad (5)$$

The main orientation from feature point to centroid is:

$$\theta_T = \arctan(m_{01} / m_{10}) \quad (6)$$

- (4) RBRIEF algorithm:

The RBRIEF algorithm is applied to generate feature descriptors in rotation invariant binary mode. The BRIEF descriptor is composed of 256 pairs 5×5 child window. The descriptors require the centroid direction of feature point to calculate the main BRIEF direction.

2.3. Improved ORB Matching Algorithm for Small Objects

Although the ORB algorithm is very rapid, its matching performance is still affected by obscure images. This section will improve the ORB algorithm with the key points selection process, RBRIEF descriptor and improved RANSAC to improve the underwater matching performance for small objects and eliminate mistake matching points.

- (1) Improvement of FAST-9 key points:

For the underwater uneven illumination environment, an adaptive threshold k_{ed} is adopted to improve the key points' selection process:

$$k_{ed} = a \times \left(\frac{1}{n} \sum_{i=1}^n g_{imax} - \frac{1}{n} \sum_{i=1}^n g_{imin} \right) \tag{7}$$

In the $L \times L$ feature point, the maximum gray value g_{imax} and minimum gray value g_{imin} are picked up to control the threshold value through the gray D-value.

(2) Apply the non-maximum value to remove repetition point:

The increase in feature description and matching depends on the feature points' congregation. The greater the repetition of the same feature points, the more feature description is required and time is consumed for the feature matching, which is disadvantageous for object positioning. After the comparisons of FAST values between feature points and their two layer pyramid, a non-maximum value is applied. FAST values are expressed with the absolute difference sum of pixel gray values:

$$v_s = \sum_{i=1}^{n_f} |g_p - g_i| \tag{8}$$

where n_f is the number of feature points; g_p is the extreme value of gray value.

(3) RBRIEF descriptor improvement:

The purpose of the improvement of the RBRIEF descriptor is to realize accurate descriptions, strong distinguishability and reduced time requirements. The norm value of three-pixel blocks will be applied instead of two pixels gray value in order to improve the descriptor robustness. Three 7×7 blocks of $p_{t,a}$, $p_{t,1}$ and $p_{t,2}$ are selected in the 48×48 square neighborhood of feature point to obtain each bit value of feature point $B_T(W_T, S_t)$:

$$B_T(W_T, S_t) = \begin{cases} 1 & \|p_{t,a} - p_{t,1}\|_F^2 > \|p_{t,a} - p_{t,2}\|_F^2 \\ 0 & otherwise \end{cases} \tag{9}$$

In order to obtain a large number of pixel block combinations, a supervised learning method has been applied to the pixel blocks grouping method. The training data set has firstly been constructed from local image descriptor data, a Harris operator has been applied to extract initial feature points and to match different images corresponding to the feature points. A total of $3 \times 56k$ pixel blocks will be generated from each feature point neighborhood to match each descriptors data. The 50k bit series will be obtained from the data set after matching. If the absolute correlation between candidate pixel block positions and all the previous positions is greater than the threshold, the candidate will be eliminated. The 256-bit binary descriptor is established at last.

(4) Improved RANSAC to eliminate mistake matching points:

The traditional RANSAC algorithm selects four matching pairs randomly, but with unstable iterative times. This paper will sort matching pairs according to their qualities and obtain the homography matrix for the top four pairs, in order to solve the optimal homography matrix and improve the matching success rate. The quality of matching pairs can be calculated as:

$$\gamma_T = \frac{1}{d_{min}^2 / d_{min2}} \tag{10}$$

where d_{min} and d_{min2} are the distances of the nearest and second nearest neighbor for the matching pair.

It is unnecessary to calculate all the matching pairs for the iteration of the homography matrix; one can randomly select half of all the matching pairs to calculate, if the ratio of

interior point is obviously lower than that of the current optimal model, then the model can be abandoned to save the iteration time Algorithm 1.

Algorithm 1. Binocular Vision Matching Method for Underwater Small Object

- ① Image preprocessing with improved unsharp masking algorithm.
 - ② Construct scale space.
 - ③ Pick up candidate key points through improved FAST-9, ascertain key points through Harris sort.
 - ④ Estimate the major orientation with intensity centroid method, and apply non-maximum value to get rid of repetition point.
 - ⑤ Generate feature descriptors in rotation invariant binary mode through improved RBRIEF descriptor.
 - ⑥ Eliminate mistake matching points through improved RANSAC.
- If the ratio of interior point is obviously lower than that of the current optimal model, go to ① and repeat sampling and matching.
Or generate parallax map.
-

3. Adaptive Non-Singular Fast Terminal Controller Design

3.1. Non-Singular Fast Terminal Sliding-Mode Controller

The dynamics model of the UVMS can be described with Euler–Lagrange equations of motion as:

$$M_{\zeta} \dot{v} + C_{\zeta} v + D_{\zeta} v + G_{\zeta} - \tau_{\zeta un} = \tau_{\zeta ctrl} + \tau_{\zeta mani} \tag{11}$$

where v denotes the velocity vector of vehicle; $M_{\zeta} \in \mathbb{R}^{(6+n) \times (6+n)}$ is the added mass, system inertial and mass matrix; $M_{\zeta} = M_0 + \Delta M_{\zeta}$, ΔM_{ζ} are the uncertainties of M_{ζ} ; n is the degree of freedoms of manipulator; $C_{\zeta} \in \mathbb{R}^{(6+n) \times (6+n)}$ is the centrifugal Coriolis matrix; $C_{\zeta} = C_0 + \Delta C_{\zeta}$, ΔC_{ζ} is the uncertainties of C_{ζ} ; $D_{\zeta} \in \mathbb{R}^{(6+n) \times (6+n)}$ denotes dissipative drag matrix caused by fluid viscosity; $D_{\zeta} = D_0 + \Delta D_{\zeta}$, ΔD_{ζ} are the uncertainties of D_{ζ} ; $G_{\zeta} \in \mathbb{R}^{(6+n) \times 1}$ is the gradational term; $G_{\zeta} = G_0 + \Delta G_{\zeta}$, ΔG_{ζ} are the uncertainties of G_{ζ} ; M_0 , C_0 , D_0 and G_0 are the known items; $\tau_{\zeta un} \in \mathbb{R}^{(6+n) \times 1}$ is the combination of external disturbance and model uncertainties; $\tau_{\zeta mani}$ denotes the coupling forces between the vehicle and manipulator [10]; $\tau_{\zeta ctrl}$ denotes the input vector of control forces.

The specific form of M_{ζ} is [39]:

$$M_{\zeta} = \begin{bmatrix} M_V & M_{Vm}^T(q) \\ M_{Vm}(q) & M_m(q) \end{bmatrix} \tag{12}$$

where M_V represents the inertia matrix of the vehicle body; $M_m(q) \in \mathbb{R}^{n \times n}$ represents the inertia matrix of the manipulator (including additional mass); $M_{Vm}(q) \in \mathbb{R}^{n \times 6}$ represents the coupling matrix of system inertia; vector $M_{Vm}^T(q)\ddot{q}$ and $M_{Vm}(q)\dot{v}$ represent the coupling forces (moments) generated between the manipulator and the vehicle body, and the vehicle body and the manipulator due to inertia.

In Equation (11), the specific form of C_{ζ} is [39]:

$$C_{\zeta} = \begin{bmatrix} C_V(v) & C_{Vm}^T(v, \dot{q}) \\ C_{Vm}(v, \dot{q}) & C_m(q, \dot{q}) \end{bmatrix} \tag{13}$$

where $C_V(v)$ represents the Coriolis force matrix of the vehicle body; $C_m(q, \dot{q}) \in \mathbb{R}^{n \times n}$ represents the Coriolis force matrix of the manipulator (including additional mass); $C_{Vm}(v, \dot{q}) \in \mathbb{R}^{n \times 6}$ represents the coupling matrix of system Coriolis forces; vector $C_{Vm}^T(v, \dot{q})\dot{q}$ and $C_{Vm}(v, \dot{q})v$, respectively, represent the coupling forces (moments) generated between the manipulator and the vehicle body, and the vehicle body and the manipulator due to Coriolis forces.

In Equation (11), the specific form of D_ζ is [39]:

$$D_\zeta = \begin{bmatrix} D_V(v) & 0 \\ 0 & D_m(q, \dot{q}) \end{bmatrix} \tag{14}$$

where $D_V(v)$ represents the damping matrix of the vehicle body; $D_m(q, \dot{q}) \in \mathbb{R}^{n \times n}$ represents the damping matrix of the manipulator.

For the transformation from the inertial coordinate system of the vehicle to the task space of the manipulator end effector, one has:

$$\begin{cases} \dot{\eta} = J(\eta)v \\ M_\eta \ddot{\eta} + C_\eta \dot{\eta} + D_\eta \dot{\eta} + G_\eta - \tau_{\eta un} = \tau_{\eta ctrl} - \tau_{\eta mani} \end{cases} \tag{15}$$

where

$$\begin{aligned} M_\eta &= J^{-T} M_\zeta J^{-1} \\ C_\eta &= J^{-T} C_\zeta J^{-1} \\ \tau_{\eta mani} &= J^{-T} \tau_{\zeta mani} \\ \tau_{\eta un} &= J^{-T} \tau_{\zeta un} \end{aligned} \tag{16}$$

$$\begin{aligned} C_\eta &= J^{-T} [C_\zeta - M_\zeta \dot{J}^{-1}] J^{-1} \\ G_\eta &= J^{-T} G_\zeta \\ \tau_{\eta un} &= J^{-T} \tau_{\zeta un} \end{aligned} \tag{17}$$

For the manipulation control, dynamic Equation (15) can be described as a second order system:

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = f(x) + d(t) + b(x)u(t) \end{cases} \tag{18}$$

where x_1 and x_2 are the system state variables vector; $f(x)$ and $b(x)$ are nonlinear functions; $d(t)$ is the external disturbance vector; $u(t)$ is the control input vector. The physical layout of the UVMS thruster positions studied in this article is shown in Figure 2, and the mathematical model diagram is shown in Figure 3.

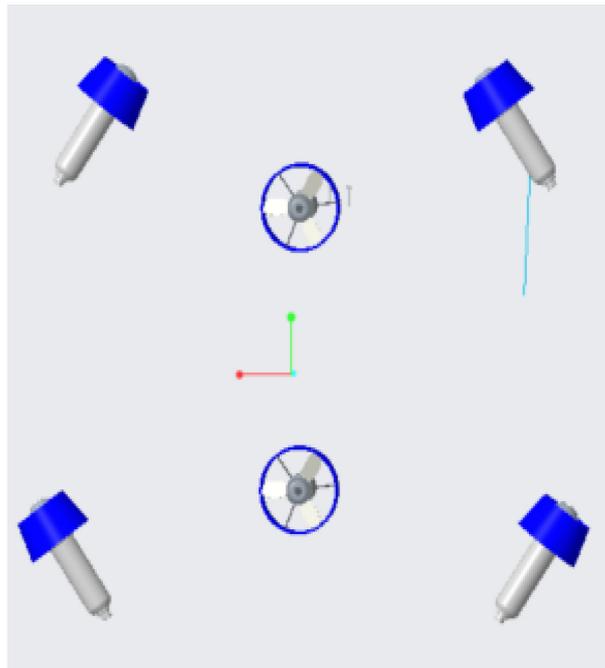


Figure 2. Physical layout diagram of the thruster positions.

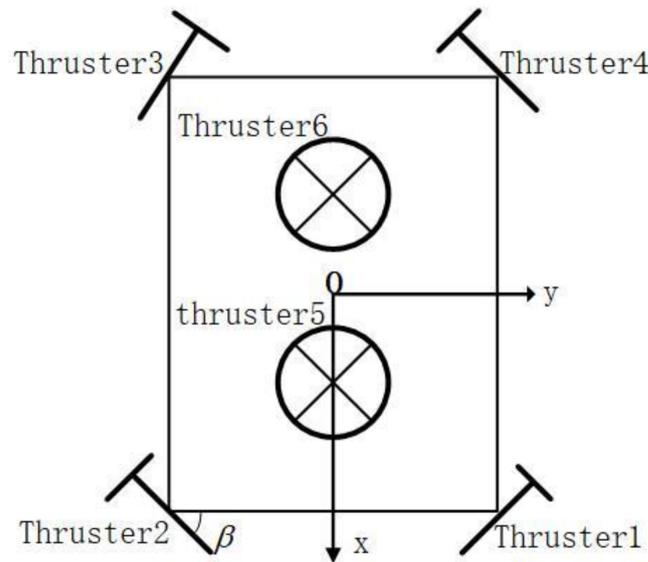


Figure 3. Physical layout diagram of the thruster positions.

The horizontal thrusters 1, 2, 3, and 4 and the forward and aft thrusters 5, 6 are arranged as shown in the figure; they will generate the control forces $\tau_{\eta ctrl}$ in the directions of forward and backward movement, lateral movement, ascent and descent, longitudinal pitch, and yaw rotation. The control force $\tau_{\eta ctrl}$ and the thrust of the thrusters $u(t)$ can be represented as shown in Equation (19):

$$\tau_{\eta ctrl} = B_v u(t) \tag{19}$$

In the equation, B_v is referred to as the thruster control matrix, and its specific form is given by Equation (20).

$$B_v = \begin{bmatrix} \sin \beta & \sin \beta & \sin \beta & \sin \beta & 0 & 0 \\ \cos \beta & -\cos \beta & -\cos \beta & \cos \beta & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & -x_5 & x_6 \\ y_1 & -y_2 & y_3 & -y_4 & 0 & 0 \end{bmatrix} \tag{20}$$

In the equation, the angular variable β , as illustrated in Figure 3, the values of $(y_1, y_2, y_3, y_4, x_5, x_6)$ represent the distances from the coordinate origin to the positions of the thrusters.

After obtaining the desired control forces using the control algorithm, it is necessary to calculate the thrust provided by each thruster. The conversion relationship is given by Equation (21), where B_v^+ is referred to as the pseudoinverse of the thruster control matrix.

$$\begin{aligned} B_v^+ &= B_v^T (B_v B_v)^{-1} \\ u(t) &= B_v^+ \tau_{ctrl} \end{aligned} \tag{21}$$

In order to realize manipulation control, the nonsingular fast terminal sliding-mode surface is [40]:

$$s = e + \frac{1}{\beta} \dot{e}^{p/q} \tag{22}$$

where $\beta > 0$, p and q are positive odd numbers; $e = x_1 - x_d$, x_d is the reference vector. One has the equivalent control law as:

$$u = b^{-1}(x) \left[\ddot{x}_d - f(x) + \beta \frac{q}{p} \dot{e}^{2-p/q} \text{sign}(\dot{e}) + (\lambda + Lg) \text{sign}(s) \right] \tag{23}$$

where $|d(t)| \leq L_g$. From (16) and (23), one has the controller for UVMS cruising as:

$$\tau_{\zeta ctrl} = J^T M_\eta(\eta) \left[\ddot{x}_d - f(x) + \beta \frac{q}{p} \dot{e}^{2-p/q} \text{sign}(\dot{e}) + (\lambda + L_g) \text{sign}(s) \right] \quad (24)$$

where $\tau_{\zeta mani} = 0$, since the manipulator stays in a fixed posture. Meanwhile, the controller for the manipulation process can be described as:

$$\tau_{\eta ctrl} = J^T M'_\eta(\eta) \left[\ddot{x}_d - f'(x) + \beta \frac{q}{p} \dot{e}^{2-p/q} \text{sign}(\dot{e}) + (\lambda + L_g) \text{sign}(s) \right] - \tau_{\eta mani} \quad (25)$$

3.2. Adaptive Non-Singular Fast Terminal Sliding-Mode Controller

Since the manipulation process should consider the coupling effects between vehicle and manipulator, thus M'_η and $f'(x)$ are different from M_η and $f(x)$, respectively.

For the practical application of (24) and (25), one should consider model uncertainties. The combination of uncertainties and disturbance can be defined as:

$$\tau_{\zeta un} = \Delta M_\zeta \dot{v} + \Delta C_\zeta v + \Delta D_\zeta v + \Delta G_\zeta + d(t) \quad (26)$$

If we assume,

$$|\tau_{\zeta un}| = |\Delta f(x) + d(t)| < \delta \quad (27)$$

where $\Delta f(x)$ denotes the model uncertainties of UVMS; δ represents the upper bound of $\tau_{\zeta un}$. If we suppose the control input $\tau_{\zeta ctrl}$ does not contain an acceleration signal, δ should be composed of the position and velocity feedback as:

$$\delta = d_0 + d_1 |e| + d_2 |\dot{e}| \quad (28)$$

where d_0, d_1 and d_2 are constants.

From (24), the controller without considering disturbances can be equivalent to:

$$\tau_{eq} = J^T M_\eta(\eta) \left[\ddot{x}_d - f(x) + \beta \frac{q}{p} \dot{e}^{2-p/q} \text{sign}(\dot{e}) \right] \quad (29)$$

The switching controller for δ can be designed as:

$$\dot{s} = k_s s + (\delta + \mu_s) \text{sign}(s) \quad (30)$$

where k_s is the positive constant; μ_s is the vector of very small parameter. In order to satisfy the sliding condition with disturbance and uncertainties, the switching controller can be designed as:

$$\begin{aligned} \tau_{sw}(t) &= J^T M_\eta(\eta) [k_s s + (\delta + \mu_s) \text{sign}(s)] \\ &= J^T M_\eta(\eta) [k_s s + (d_0 + d_1 \|e\| + d_2 \|\dot{e}\|^2 + \eta) \text{sign}(s)] \end{aligned} \quad (31)$$

The adaptive switching controller is:

$$\tau_{asw}(t) = J^T M_\eta(\eta) [k s + (\hat{d}_0 + \hat{d}_1 \|e\| + \hat{d}_2 \|\dot{e}\|^2 + \mu_s) \text{sign}(s)] \quad (32)$$

Therefore, the adaptive nonsingular fast terminal sliding-mode controller is designed as:

$$\tau_{ANTSMC} = J^T M_\eta(\eta) \left[\ddot{x}_d - f(x) + \beta \frac{q}{p} \dot{e}^{2-p/q} \text{sign}(\dot{e}) \right] + J^T M_\eta(\eta) [k s + (\hat{d}_0 + \hat{d}_1 \|e\| + \hat{d}_2 \|\dot{e}\|^2 + \mu_s) \text{sign}(s)] \quad (33)$$

where $\tilde{d}_0 = \hat{d}_0 - d_0$, $\tilde{d}_1 = \hat{d}_1 - d_1$, $\tilde{d}_2 = \hat{d}_2 - d_2$ are the adaptation errors. \hat{d}_0 , \hat{d}_1 and \hat{d}_2 can be updated through:

$$\begin{cases} \dot{\hat{d}}_0 = \lambda_0 |s| \|\dot{e}\|^{\beta-1} \\ \dot{\hat{d}}_1 = \lambda_1 |s| \|\dot{e}\|^{\beta-1} \|e\| \\ \dot{\hat{d}}_2 = \lambda_1 |s| \|\dot{e}\|^{\beta-1} \|e\|^2 \end{cases} \quad (34)$$

where λ_0 , λ_1 and λ_1 are arbitrary positive-tuning parameters.

3.3. The Overall Controller with Barrier Function

In the vision-based manipulation process, the vehicle should stay in its position so that the target can be kept in the field of sight. However, the external hydrodynamic disturbance and the coupling effects between vehicle and manipulator make the control system vibrate reciprocally and apt to lose the target.

In order to further suppress the vibration and ensure controller convergence in finite time, the barrier function was designed. The barrier function $K_{BF}(\sigma)$ is an even and positive function (see Figure 4). The interval parameter ε is defined as $\varepsilon > 0$. The boundaries of the barrier function are $\sigma \in (-\varepsilon, \varepsilon)$ and $K_{BF}(\sigma) \in (0, \infty)$. For the barrier function, one has:

$$\begin{cases} K_{BF}(\sigma) = 0 & \text{when } \sigma \rightarrow \pm\varepsilon \\ K_{BF}(\sigma) \rightarrow \infty & \text{when } \lim_{\sigma \rightarrow \pm\varepsilon} K_{BF}(\sigma) = +\infty \end{cases} \quad (35)$$

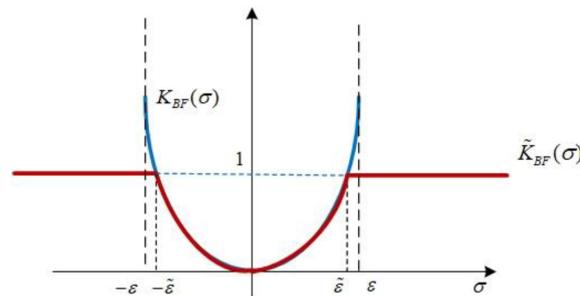


Figure 4. Basic barrier function and quasi barrier function.

The quasi-barrier function is proposed for the adaptation controller and vibration suppression from the barrier function.

$$\begin{cases} \tilde{K}_{BF}(\sigma) = K_{BF}(\sigma) & \text{when } -\tilde{\varepsilon} < \sigma < \tilde{\varepsilon} \\ \tilde{K}_{BF}(\sigma) = 1 & \text{when } \sigma \leq -\tilde{\varepsilon}, \sigma \geq \tilde{\varepsilon} \end{cases} \quad (36)$$

The basic barrier function and quasi barrier function can be defined as:

$$\begin{cases} K_{BF}(\sigma) = \frac{\bar{L}}{\varepsilon - |\sigma|} \\ \tilde{K}_{BF}(\sigma) = \frac{\bar{L} \cdot \text{sat}_{\tilde{\varepsilon}}(|\sigma|)}{(\varepsilon - \text{sat}_{\tilde{\varepsilon}}|\sigma|)^{\gamma}} \end{cases} \quad (37)$$

where $\lim_{\sigma \rightarrow \varepsilon} \tilde{K}_{BF}(\sigma) = 1$, $\bar{L} = \frac{\varepsilon - \tilde{\varepsilon}}{\tilde{\varepsilon}}$; the $\text{sat}_{\tilde{\varepsilon}}$ function is defined as: $\text{sat}_{\tilde{\varepsilon}}(\sigma) = \begin{cases} \sigma & |\sigma| < \tilde{\varepsilon} \\ \tilde{\varepsilon} \text{sign}(\sigma) & |\sigma| \geq \tilde{\varepsilon} \end{cases}$

Therefore, the quasi-barrier function concerning control errors and sliding surface can be described as:

$$\begin{cases} \tilde{K}_{BF1}(\dot{e}) = \tilde{L}_1 \frac{\text{sat}_{\tilde{\varepsilon}_1}(|\dot{e}|)}{(\varepsilon_1 - \text{sat}_{\tilde{\varepsilon}_1}|\dot{e}|)^{\gamma}} \\ \tilde{K}_{BF2}(s) = \tilde{L}_2 \frac{\text{sat}_{\tilde{\varepsilon}_2}(|s|)}{(\varepsilon_2 - \text{sat}_{\tilde{\varepsilon}_2}|s|)^{\gamma}} \end{cases} \quad (38)$$

where ε_1 and ε_2 are the bound of control errors and sliding surface, respectively; \tilde{L}_1 and \tilde{L}_2 are the gains of $\tilde{K}_{BF1}(\dot{e})$ and $\tilde{K}_{BF2}(s)$; $sat_{\tilde{\varepsilon}_1}(|\dot{e}|)$ and $sat_{\tilde{\varepsilon}_2}(|s|)$ are the saturation functions.

$$\begin{cases} \tilde{L}_1 = \frac{\varepsilon_1 - \tilde{\varepsilon}_1}{\tilde{\varepsilon}_1} \\ \tilde{L}_2 = \frac{\varepsilon_2 - \tilde{\varepsilon}_2}{\tilde{\varepsilon}_2} \end{cases} \text{ and } \begin{cases} sat_{\tilde{\varepsilon}_1}(\dot{e}) = \begin{cases} \dot{e} & \text{for } |\dot{e}| < \tilde{\varepsilon}_1 \\ \tilde{\varepsilon}_1 sign(\dot{e}) & \text{for } |\dot{e}| \geq \tilde{\varepsilon}_1 \end{cases} \\ sat_{\tilde{\varepsilon}_2}(s) = \begin{cases} s & \text{for } |s| < \tilde{\varepsilon}_2 \\ \tilde{\varepsilon}_2 sign(s) & \text{for } |s| \geq \tilde{\varepsilon}_2 \end{cases} \end{cases} \quad (39)$$

Therefore, from the frame of the overall controller in Figure 5, the overall controller is designed as:

$$\begin{cases} \tau_{\zeta ctrl} = J^T M_{\eta}(\eta) \left[\begin{aligned} &\ddot{x}_d - f(x) + \beta \frac{q}{p} |\dot{e}|^{2-p/q} \tilde{K}_{BF1}(\dot{e}) sign(\dot{e}) \\ &+ (\hat{d}_0 + \hat{d}_1 \|e\| + \hat{d}_2 \|\dot{e}\|^2 + \mu_s) * \tilde{K}_{BF2}(s) * sign(s) + ks \end{aligned} \right] \text{crui sin g} \\ \tau_{\zeta ctrl} = J^T M'_{\eta}(\eta) \left[\begin{aligned} &\ddot{x}_d - f'(x) + \beta \frac{q}{p} |\dot{e}|^{2-p/q} \tilde{K}_{BF1}(\dot{e}) sign(\dot{e}) \\ &+ (\hat{d}_0 + \hat{d}_1 \|e\| + \hat{d}_2 \|\dot{e}\|^2 + \mu_s) * \tilde{K}_{BF2}(s) * sign(s) + ks \end{aligned} \right] - \tau_{\zeta mani} \text{manipulating} \end{cases} \quad (40)$$

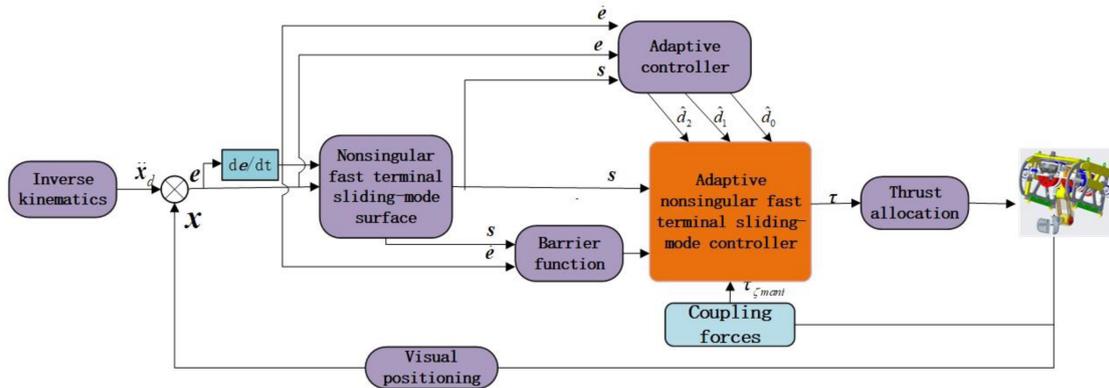


Figure 5. The frame of overall controller.

Theorem 1. For the control system (40), sliding surface is chosen as (26), the barrier function is defined as (38), the control system converges to the equilibrium point in finite time.

Proof. The Lyapunov function candidate about barrier function is selected as:

$$V_{BF} = |s| + \varepsilon_2^{-1} \tilde{L}_2 \tilde{K}_{BF2}(s) \quad (41)$$

Different V_{BF} with respect to time, one has:

$$\dot{V}_{BF} = \frac{s}{|s|} \dot{s} + \tilde{L}_2^2 \frac{sign(s)}{(\varepsilon_2 - s)} \dot{s} \quad (42)$$

If we split \dot{V}_{BF} into Equation (43)

$$\dot{V}_{BF} = \dot{V}_{BF1} + \dot{V}_{BF2} \quad (43)$$

where

$$\begin{cases} \dot{V}_{BF1} = \frac{s}{|s|} \dot{s} = \frac{s}{|s|} (\dot{e} + \frac{1}{\beta} \frac{p}{q} \dot{e}^{p/q-1} \ddot{e}) \\ \dot{V}_{BF2} = \tilde{L}_2^2 \frac{sign(s)}{(\varepsilon_2 - s)} \dot{s} = \tilde{L}_2^2 \frac{sign(s)}{(\varepsilon_2 - s)} (\dot{e} + \frac{1}{\beta} \frac{p}{q} \dot{e}^{p/q-1} \ddot{e}) \end{cases} \quad (44)$$

From (33) and (34), one has:

$$\begin{aligned} \dot{V}_{BF1} &= \frac{s}{|s|} \dot{s} = \text{sign}(s) \left(\dot{e} + \frac{1}{\beta} \frac{p}{q} \dot{e}^{p/q-1} \ddot{e} \right) = \text{sign}(s) \left(\dot{e} + \frac{1}{\beta} \frac{p}{q} \dot{e}^{p/q-1} (\ddot{x}_d - \ddot{x}) \right) \\ &= \text{sign}(s) \left(\dot{e} + \frac{1}{\beta} \frac{p}{q} \dot{e}^{p/q-1} [\ddot{x}_d - f(x) + \tau_{\zeta un} - \ddot{x}_d + f(x) - \right. \\ &\quad \left. \beta \frac{q}{p} |\dot{e}|^{2-p/q} \tilde{K}_{BF1}(\dot{e}) \text{sign}(\dot{e}) - \frac{1}{\beta} \frac{p}{q} \dot{e}^{p/q-1} (\hat{d}_0 + \hat{d}_1 \|e\| + \hat{d}_2 \|\dot{e}\|^2 + \mu_s) \right. \\ &\quad \left. * \tilde{K}_{BF2}(s) * \text{sign}(s) - \frac{1}{\beta} \frac{p}{q} \dot{e}^{p/q-1} k_s s \right) \\ &= \text{sign}(s) \left[\left(\dot{e} - \tilde{K}_{BF1}(\dot{e}) + \frac{1}{\beta} \frac{p}{q} \dot{e}^{p/q-1} \tau_{\zeta un} \right) \right. \\ &\quad \left. - \frac{1}{\beta} \frac{p}{q} \dot{e}^{p/q-1} (\hat{d}_0 + \hat{d}_1 \|e\| + \hat{d}_2 \|\dot{e}\|^2 + \mu_s) * \tilde{K}_{BF2}(s) - \frac{1}{\beta} \frac{p}{q} \dot{e}^{p/q-1} k_s |s| \right] \end{aligned}$$

and

$$\begin{aligned} \dot{V}_{BF2} &= \tilde{L}_2^2 \frac{\text{sign}(s)}{(\varepsilon_2 - |s|)^2} \dot{s} = \tilde{L}_2^2 \frac{\text{sign}(s)}{(\varepsilon_2 - |s|)^2} \left(\dot{e} + \frac{1}{\beta} \frac{p}{q} \dot{e}^{p/q-1} \ddot{e} \right) \\ &= \tilde{L}_2^2 \frac{\text{sign}(s)}{(\varepsilon_2 - |s|)^2} \left(\dot{e} + \frac{1}{\beta} \frac{p}{q} \dot{e}^{p/q-1} (\ddot{x}_d - \ddot{x}) \right) \\ &= \tilde{L}_2^2 \frac{\text{sign}(s)}{(\varepsilon_2 - |s|)^2} \left(\dot{e} + \frac{1}{\beta} \frac{p}{q} \dot{e}^{p/q-1} [\ddot{x}_d - f(x) + \tau_{\zeta un} - \ddot{x}_d + f(x) - \right. \\ &\quad \left. \beta \frac{q}{p} |\dot{e}|^{2-p/q} \tilde{K}_{BF1}(\dot{e}) \text{sign}(\dot{e}) - \frac{1}{\beta} \frac{p}{q} \dot{e}^{p/q-1} (\hat{d}_0 + \hat{d}_1 \|e\| + \hat{d}_2 \|\dot{e}\|^2 + \mu_s) \right. \\ &\quad \left. * \tilde{K}_{BF2}(s) * \text{sign}(s) - \frac{1}{\beta} \frac{p}{q} \dot{e}^{p/q-1} k_s s \right) \\ &= \tilde{L}_2^2 \frac{\text{sign}(s)}{(\varepsilon_2 - |s|)^2} \left[\left(\dot{e} - \tilde{K}_{BF1}(\dot{e}) + \frac{1}{\beta} \frac{p}{q} \dot{e}^{p/q-1} \tau_{\zeta un} \right) \right. \\ &\quad \left. - \frac{1}{\beta} \frac{p}{q} \dot{e}^{p/q-1} \frac{\tilde{L}_2^2}{(\varepsilon_2 - |s|)^2} * (\hat{d}_0 + \hat{d}_1 \|e\| + \hat{d}_2 \|\dot{e}\|^2 + \mu_s) \right. \\ &\quad \left. * \tilde{K}_{BF2}(s) - \frac{1}{\beta} \frac{p}{q} \dot{e}^{p/q-1} \frac{\tilde{L}_2^2}{(\varepsilon_2 - |s|)^2} k_s |s| \right] \end{aligned}$$

If we define $\zeta_1 = \frac{1}{\beta} \frac{p}{q} \dot{e}^{p/q-1} (\hat{d}_0 + \hat{d}_1 \|e\| + \hat{d}_2 \|\dot{e}\|^2 + \mu_s)$ and $\zeta_2 = (\dot{e} - \tilde{K}_{BF1}(\dot{e}) + \frac{1}{\beta} \frac{p}{q} \dot{e}^{p/q-1} \tau_{\zeta un})$ one has:

$$\begin{aligned} \dot{V}_{BF} &= \dot{V}_{BF1} + \dot{V}_{BF2} \\ &= \text{sign}(s) \zeta_2 - \zeta_1 * \tilde{K}_{BF2}(s) - \frac{1}{\beta} \frac{p}{q} \dot{e}^{p/q-1} k_s |s| \\ &\quad + \text{sign}(s) \frac{\tilde{L}_2^2}{(\varepsilon_2 - |s|)^2} \zeta_2 - \zeta_1 \frac{\tilde{L}_2^2}{(\varepsilon_2 - |s|)^2} * \tilde{K}_{BF2}(s) - \frac{1}{\beta} \frac{p}{q} \dot{e}^{p/q-1} \frac{\tilde{L}_2^2}{(\varepsilon_2 - |s|)^2} k_s |s| \end{aligned}$$

If we define $\text{sign}(\dot{e}) \zeta_2 \leq |\zeta_2|$, one has:

$$\dot{V}_{BF} \leq -\zeta_1 * \tilde{K}_{BF2}(s) - k_1 \frac{\tilde{L}_2^2}{(\varepsilon_2 - |s|)^2} * \tilde{K}_{BF2}(s) + \Delta_1 - \frac{1}{\beta} \frac{p}{q} \dot{e}^{p/q-1} k |s| - \frac{1}{\beta} \frac{p}{q} \dot{e}^{p/q-1} \frac{\tilde{L}_2^2}{(\varepsilon_2 - |s|)^2} k |s|$$

where $\Delta_1 = |\zeta_2| + \frac{\tilde{L}_2^2}{(\varepsilon_2 - |s|)^2} |\zeta_2|, \Delta_1 \geq 0$. Thus, one has:

$$\begin{aligned} \dot{V}_{BF} &= -\zeta_1 \frac{\tilde{L}_2^2}{(\varepsilon_2 - |s|)^2} |s| - \frac{1}{\beta} \frac{p}{q} \dot{e}^{p/q-1} k_s |s| - \frac{1}{\beta} \frac{p}{q} \dot{e}^{p/q-1} \frac{\tilde{L}_2^2}{(\varepsilon_2 - |s|)^2} k_s |s| - \zeta_1 \frac{\tilde{L}_2}{(\varepsilon_2 - |s|)} |s| + \Delta_1 \\ &= -\zeta_1 \frac{\tilde{L}_2^2}{(\varepsilon_2 - |s|)^2} |s| - \frac{1}{\beta} \frac{p}{q} \dot{e}^{p/q-1} k_s |s| - \frac{1}{\beta} \frac{p}{q} \dot{e}^{p/q-1} \frac{\tilde{L}_2^2}{(\varepsilon_2 - |s|)^2} k_s |s| - \frac{1}{(\varepsilon_2 - |s|)} (\tilde{L}_2 \zeta_1 |s| - (\varepsilon_2 - |s|) \Delta_1) \\ &= -\zeta_1 \frac{\tilde{L}_2^2}{(\varepsilon_2 - |s|)^2} |s| - \frac{1}{\beta} \frac{p}{q} \dot{e}^{p/q-1} k_s |s| - \frac{1}{\beta} \frac{p}{q} \dot{e}^{p/q-1} \frac{\tilde{L}_2^2}{(\varepsilon_2 - |s|)^2} k_s |s| - \frac{\Delta_1}{(\varepsilon_2 - |s|)} (\frac{\tilde{L}_2 \zeta_1}{\Delta_1} |s| - (\varepsilon_2 - |s|)) \\ &= -\zeta_1 \frac{\tilde{L}_2^2}{(\varepsilon_2 - |s|)^2} |s| - \frac{1}{\beta} \frac{p}{q} \dot{e}^{p/q-1} k_s |s| - \frac{1}{\beta} \frac{p}{q} \dot{e}^{p/q-1} \frac{\tilde{L}_2^2}{(\varepsilon_2 - |s|)^2} k_s |s| - \frac{\Delta_1}{(\varepsilon_2 - |s|)} (\frac{\tilde{L}_2 \zeta_1}{\Delta_1} |s| + |s| - \varepsilon_2) \end{aligned}$$

where $\frac{\tilde{L}_2\tilde{\zeta}_1}{\Delta_1}|\mathbf{s}| + |\mathbf{s}| - \varepsilon_2 \geq 0$. If $|\mathbf{s}| \geq \varepsilon_2$, one has $\frac{\tilde{L}_2\tilde{\zeta}_1}{\Delta_1}|\mathbf{s}| + |\mathbf{s}| - \varepsilon_2 > 0$. If $|\mathbf{s}| \rightarrow \varepsilon_2$, one has $\Delta_1 = 0$ and $\dot{V}_{BF} < 0$. Therefore, the control law (32) is asymptotically stable. \square

4. Simulations and Experiments

4.1. Image Processing and Binocular Visual Matching Experiments

In order to verify the proposed algorithm, underwater images in the tank and oceanic environment have been processed with fast guide filter algorithm and improved unsharp masking algorithm. From the processing results, the proposed algorithm can better preserve image marginal information and smooth background details. The comparison of image processing algorithms is shown in Figure 6.

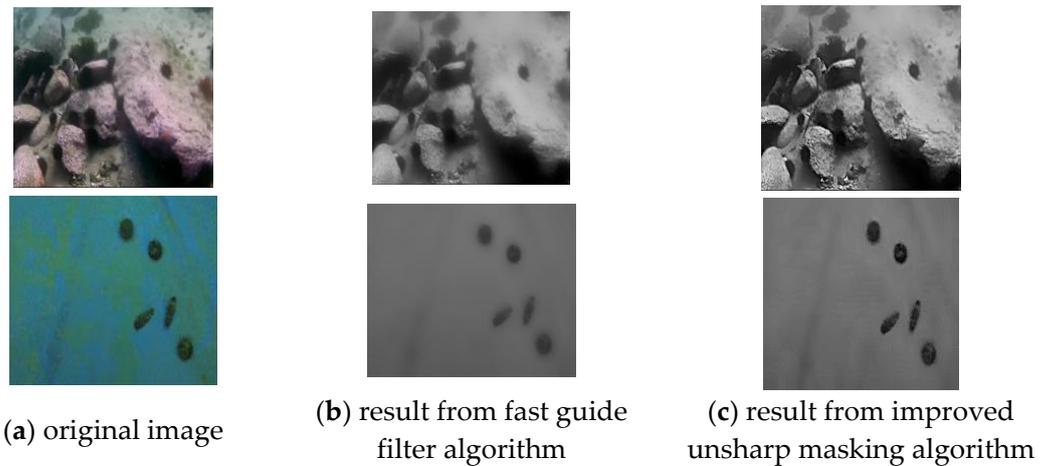
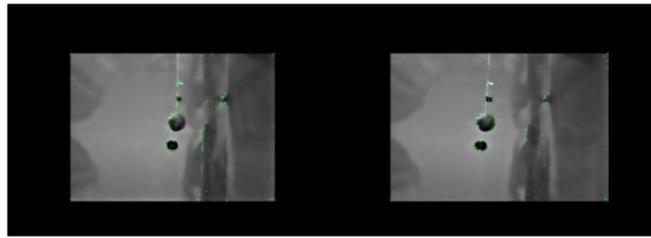
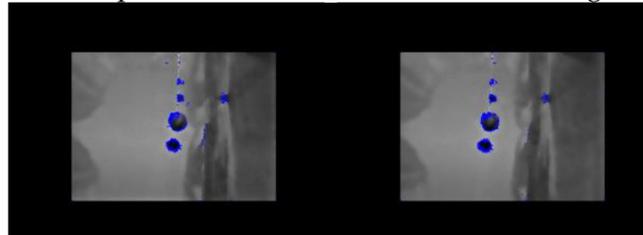


Figure 6. Comparisons for image processing algorithm.

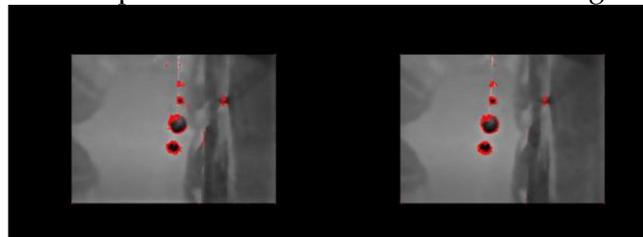
To prove the robustness and accuracy of the visual matching algorithm of this paper, the classical SIFT [41] (scale invariant feature transform) and ORB have been compared with on feature point extraction and matching (see Figures 7 and 8). The feature extraction points from SIFT algorithm are comparatively sparse, resulting in few matching pairs of the small objects. For the ORB algorithm, unclear margin and texture could result in few or unstable matching pairs. On the other hand, the feature extraction points from the proposed algorithm can obtain dense and accurate matching pairs with strong distinguishable characters. On the matching efficiency (see Table 1), non-maximum value and improved RBRIEF descriptor have got rid of the repetition and ineffective matching points, which has reduced the time required and improved the matching accuracy. The matching efficiency of the improved method is significantly better than the ORB algorithm and slightly better than the SIFT algorithm. In terms of real-time performance, the algorithm proposed in this paper takes slightly longer than the ORB algorithm in the feature point extraction process. Although the algorithm proposed in this paper improves the feature description compared to ORB, increasing the algorithm complexity, the time consumed in the final feature point description process is still smaller than that of the ORB algorithm. Therefore, in terms of real-time performance, the algorithm proposed in this paper is superior to both the ORB algorithm and the SIFT algorithm.



(a) Feature point extraction results from SIFT algorithm

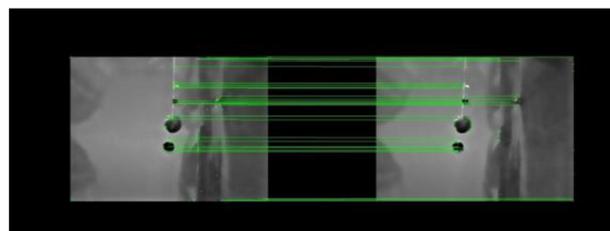


(b) Feature point extraction results from ORB algorithm

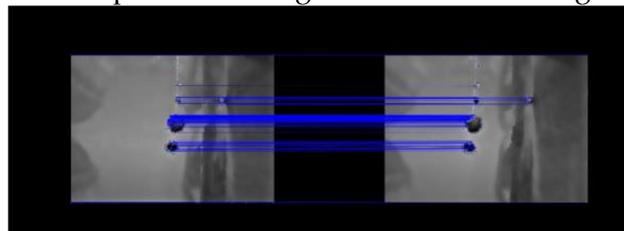


(c) Feature point extraction results from the proposed algorithm

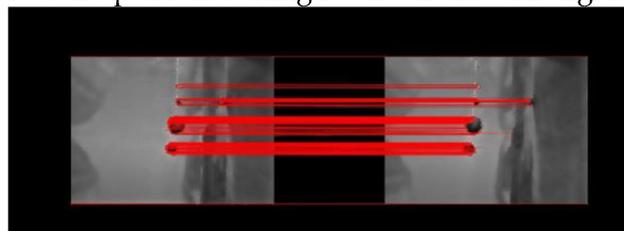
Figure 7. Comparison of results for the feature point extraction.



(a) Feature point matching results from SIFT algorithm



(b) Feature point matching results from ORB algorithm



(c) Feature point matching results from the proposed algorithm

Figure 8. Comparison of results for the feature point matching.

Table 1. Time required and numbers for feature point extraction.

Descriptors	Number of Left Map Feature Point	Number of Right Map Feature Point	Feature Point Extraction Time (ms)	Feature Point Description Time (ms)	Numbers of Thick Matching Point Pairs	Ratio of Feature Point Matching
SIFT	125	174	5970	233	30	24%
OFAST	1865	1909	127	387	158	8.47%
Proposed algorithm	1591	1627	144	184	484	30.42%

To prove the underwater visual localization accuracy of the proposed method in this paper, the localization results of the three algorithms were compared with the ground truth values obtained in the underwater experiment, as shown in Table 2. For the three tested targets: sea cucumber, scallop, and sea urchin, the improved algorithm proposed in this paper shows significantly higher accuracy than ORB and SIFT algorithms, and the coordinate values after transformation by this algorithm are within the allowable error range.

Table 2. Comparison of image target coordinates.

Object	Descriptors	True Value(mm)	Direct Calibration Calculation Value (mm)	Relative Error	Calibration Calculation Value after Conversion (mm)	Relative Error
Sea cucumber	Proposed algorithm				(93.84, -119.79, 1186.23)	(3.69%, -2.02%, 0.921%)
	SIFT	(90.50, -117.42, 1175.40)	(100.51, -107.53, 1303.09)	(11.06%, 8.42%, 10.86%)	(94.73, -114.414, 1189.12)	(4.68%, -2.56%, 1.17%)
	OFAST				(102.49, -108.90, 1214.28)	(13.25%, -7.25%, 3.31%)
Pectinid	Proposed algorithm				(92.57, 15.86, 1187.78)	(-2.21%, -3.52%, 1.05%)
	SIFT	(94.62, 16.44, 1175.40)	(83.73, 13.96, 1259.79)	(-11.21%, -15.08%, 7.18%)	(91.97, 15.71, 1191.04)	(-2.80%, -4.46%, 1.33%)
	OFAST				(87.11, 14.36, 1219.72)	(-7.94%, -12.64%, 3.77%)
Sea urchin	Proposed algorithm				(62.45, 127.78, 1187.02)	(1.51%, 2.58%, 1.02%)
	SIFT	(61.52, 124.56, 1175.40)	(70.89, 142.26, 1244.32)	(15.23%, 13.41%, -5.86%)	(62.70, 128.63, 1190.59)	(1.91%, 3.27%, 1.29%)
	OFAST				(64.86, 136.10, 1218.16)	(5.42%, 9.27%, 3.66%)

To prove the improved RANSAC algorithm for mismatching elimination, Figure 9 illustrate the comparison of results for RANSAC and improved RANSAC algorithm in the matching map. From Figure 9 and Table 3, the iteration time for RANSAC is apparently unstable with fewer accurate matching points than the improved RANSAC algorithm of this paper. The results show the proposed algorithm is more accurate with more correct matching points.

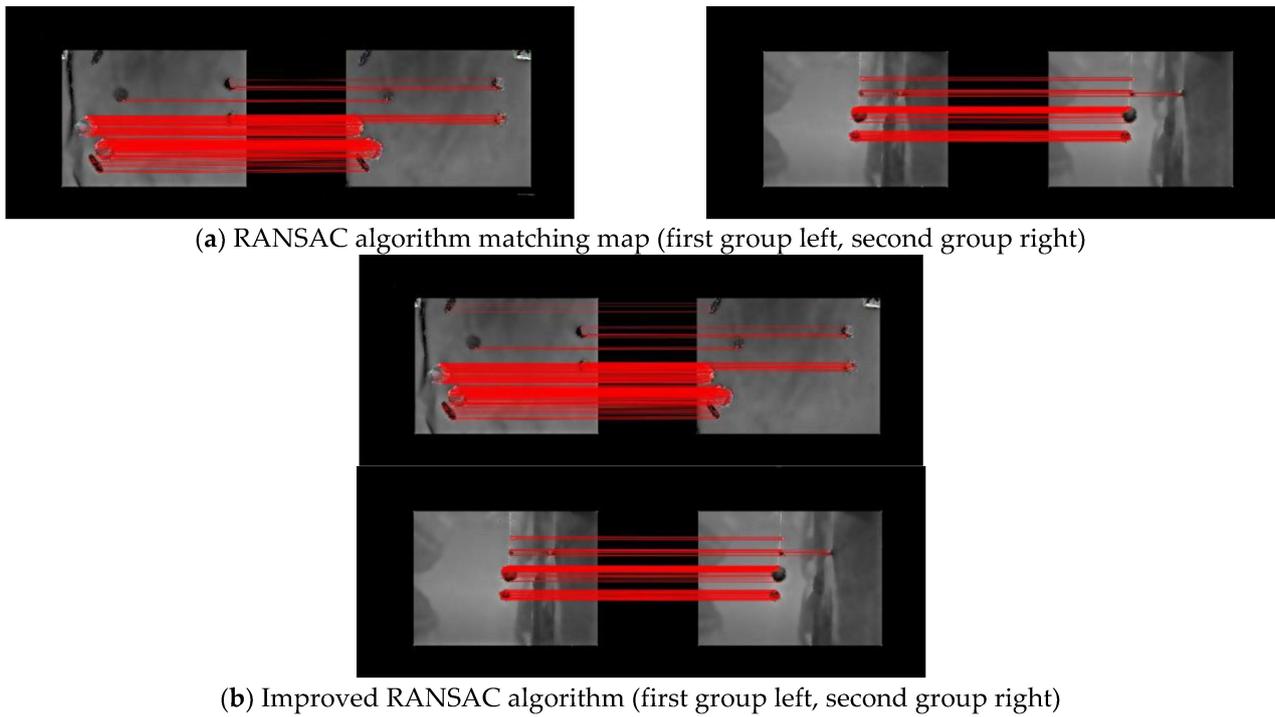


Figure 9. Comparison of results for the feature point accurate matching.

Table 3. Comparisons between proposed and RANSAC algorithm.

	Algorithm	Number of Thick Matching Point	Number of Accurate Matching Point	Time for Eliminate Mistake Matching Time (ms)	Ratio of Correct Matching Point
First Group	RANSAC	554	466	5.4	84.11%
	Proposed algorithm	554	489	2.7	88.27%
Second Group	RANSAC	484	339	9.2	70.04%
	Proposed algorithm	484	368	3.5	76.03%

4.2. Simulations on Proposed Non-Singular Fast Terminal Sliding-Mode Controller

This numerical simulation is for the autonomous grasping task of the underwater vehicle manipulator system after discovering a target. This paper focuses on the control problem of target grasping for underwater robots based on visual-assisted localization. It mainly addresses the issues of pose accuracy and frequency during the target grasping process of the underwater vehicle manipulator system (UVMS), as well as the kinematic planning and UVMS dynamic control for the target grasping task after obtaining the relative distance between the robot and the target.

The simulated process of the target grasping for the UVMS based on visual-assisted localization is as follows: When the target is detected through the visual system, the initial accurate value is obtained from binocular ranging and the desired pose is obtained for both the vehicle and the manipulator. Subsequent binocular ranging values will be used for the localization system; based on the current position state, we can calculate the boat's position error e and sliding mode surface s by setting a target position or reference position; by utilizing the position error e , error derivative \dot{e} , and sliding mode surface s , we can obtain the adaptive update law; calculations can be performed to obtain $\bar{K}_{BF1}(\dot{e})$ based on the position error \dot{e} and a predefined ϵ_1 , as well as $\bar{K}_{BF2}(s)$ based on the sliding mode surface s and a predefined ϵ_2 ; the joint drivers rotate the individual joints of the robotic arm and calculate the coupling forces $\tau_{\zeta mani}$; the UVMS (underwater vehicle manipulator

system) coordination controller for target grasping based on the quasi-obstacle function is obtained through the obtained errors e , sliding mode surface s , and quasi-obstacle function $\bar{K}_{BF1}(e), \bar{K}_{BF2}(s)$; the thrust allocation is used to calculate the required thrust for each thruster.

In order to analyze the proposed non-singular fast terminal sliding-mode controller on the UVMS, the non-singular fast terminal sliding-mode controller (NTSMC) in the first part of Section 3, adaptive non-singular fast terminal sliding-mode controller (ANTSMC) in the second part of Section 3 and quasi barrier function based adaptive non-singular fast terminal sliding-mode controller (QBF-ANTSMC) in the third part of Section 3 have been compared. The platform of the Oceanic Experiments with three DOFs manipulator has been applied for the simulations. The Denavit–Hartenberg parameters of the manipulator are illustrated in Table 4. The quality and center of gravity positions of UVMS are displayed in Table 5. The dynamic parameters are displayed in Table 6. The correlated parameters of the controller are set out in Tables 7–9. Figure 10 illustrates the single dimensional DOF control results in longitudinal, vertical and heading directions.

Table 4. Denavit–Hartenberg parameters of the manipulator.

Joint	a_i (m)	α_i (rads)	d_i (m)	q_i (rads)	\hat{a}_i (m)
1	0	$\pi/2$	0	q_1	0
2	0.6	0	0	q_2	0.5
3	0.3	0	0	q_3	0.5

Table 5. Quality and center of gravity position of UVMS.

m (kg)	x_g (mm)	y_g (mm)	Z_g (mm)
82.13	-3.04×10^{-2}	2.79×10^2	3.19×10^2

Table 6. Main dynamic parameters of the system.

	Vehicle	Link 1	Link 2	Link 3
M (kg)	82.13	2.60	3.52	3.16
I_{xx} · (kgm ²)	4.95	0.03	0	0
I_{yy} · (kgm ²)	7.36	0.03	0.16	0.16
I_{zz} · (kgm ²)	8.66	0	0.16	0.16

Table 7. Parameters of NTSMC.

q	p	β	λ	Lg
3	5	2	0.4	6

Table 8. Parameters of ANTSMC.

μ_0	0.1	μ_1	0.1	μ_2	0.1
k_1	20	k_2	20	k_3	20
k_4	200	k_5	200	k_6	200
η_1	0.11	η_2	0.11	η_3	0.11
η_4	1	η_5	1	η_6	1

Table 9. Parameters of QBF-ANTSMC.

ϵ_1	$\tilde{\epsilon}_1$	ϵ_2	$\tilde{\epsilon}_2$
0.98	0.9	0.98	0.7

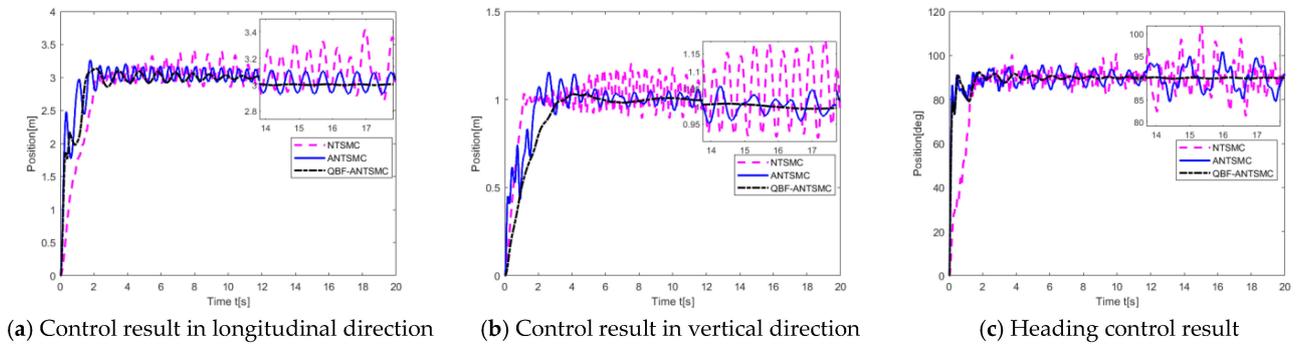


Figure 10. The single dimensional DOF control result.

Since the motion control of UVMS sometimes requires multiple degrees of freedom to move together, this paper conducts simulation experiments on multiple degrees of freedom of UVMS. The desired trajectory for the simultaneous motion of four degrees of freedom in the simulation process is as follows:

$$\eta_d = [r \sin(wt), r - r \cos(wt), -r \sin(wt), 0, 0, wt]^T \tag{45}$$

where $r = 2, w = \frac{2\pi}{40} rad/s$. We conducted simulation experiments to control the UVMS system using three different methods: NTSMC, ANTSMC and QBF-ANTSMC.

Figure 11 presents the comparisons of vehicle three-dimensional trajectory tracking. From the results, the vehicle can reach the desired position in finite time with the NTSMC controller, but should assume a disturbance upper limit. Moreover, the upper limit should be accurately set, to eliminate preliminary tracking errors. Under the ANTSMC control, the vehicle can reach the desired position in finite time without a disturbance upper limit assumption, but still with great chattering phenomena. On the other hand, QBF-ANTSMC can obtain a trajectory that strictly coincides with the desired trajectory, which means the proposed controller can obtain accurate control results and suppress chattering phenomena.

By utilizing binocular vision to obtain target information and acquiring the desired pose of the robot, subsequent simulation experiments are conducted for trajectory tracking. The initial position of the robot, target position and desired pose are shown in Table 10 for this simulation. Figures 12 and 13 illustrate the position-based manipulation control process.

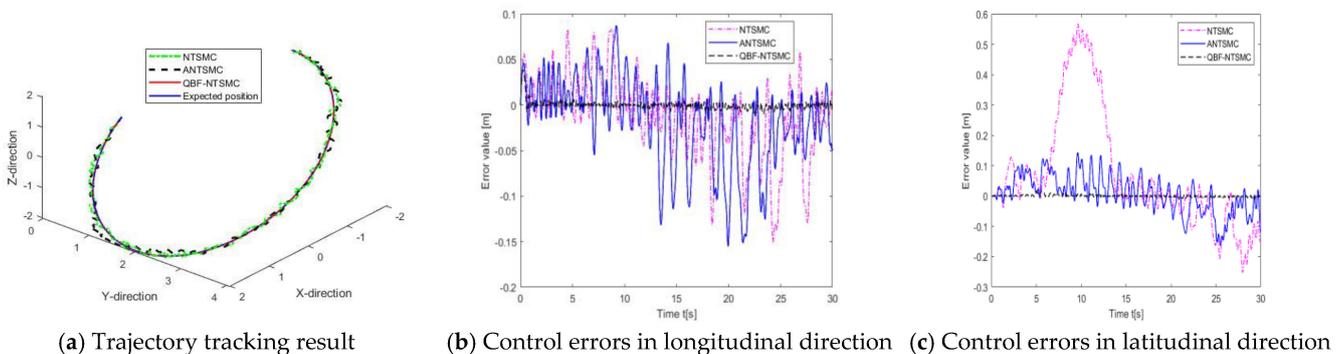
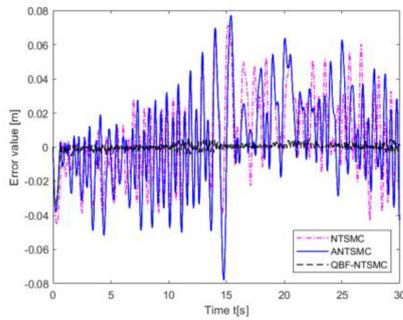
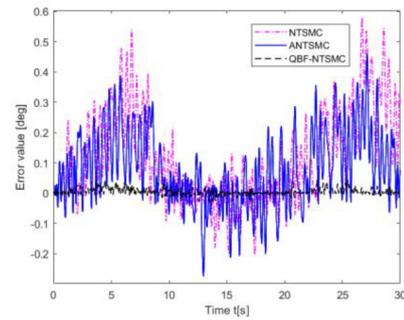


Figure 11. Cont.



(d) Control errors in vertical direction



(e) Control errors in yaw direction

Figure 11. Three–dimensional control results for comparison.

Table 10. The task information for the simulation experiment.

The Initial Position of the Robot (m)	Target Position (m)	Desired Vehicle Position (m)	Desired Vehicle Attitude (rad)	Desired Manipulator Attitude (rad)
(0, 0, 0)	(1.5, 0.5, 2)	(0.9518, 0.5, 1.5845)	(0, 0.1231, 0)	(0, −0.2159, 0.0929)

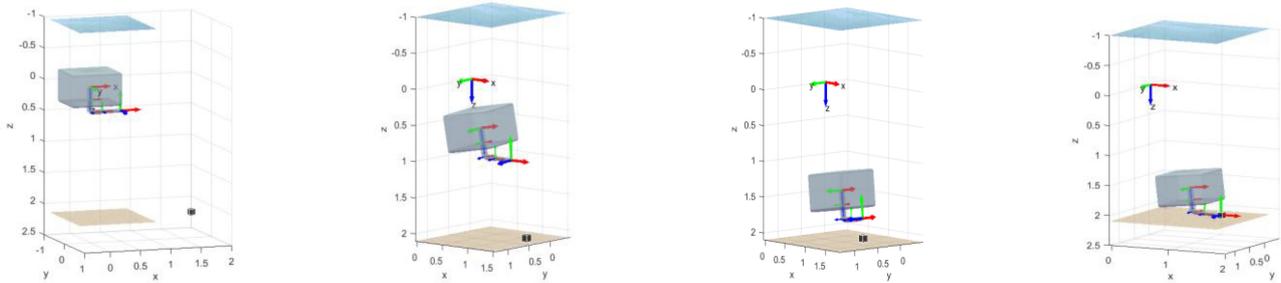
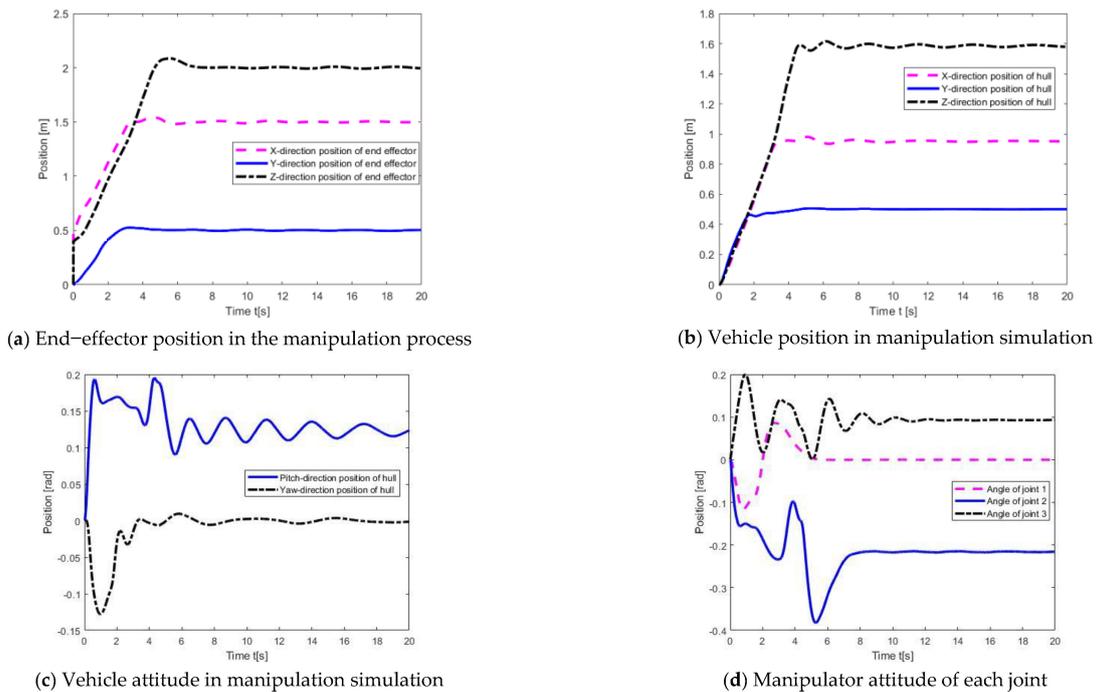


Figure 12. Manipulation simulation process.



(a) End-effector position in the manipulation process

(b) Vehicle position in manipulation simulation

(c) Vehicle attitude in manipulation simulation

(d) Manipulator attitude of each joint

Figure 13. Manipulation simulation result.

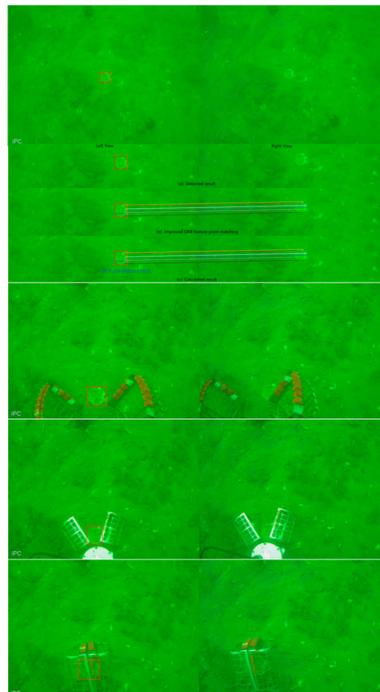
4.3. Oceanic Experiments on Small Object Positioning and Grasp

Based on the effective numerical simulation using physical parameter data, we conducted oceanic experiments to verify the effectiveness of the proposed binocular visual algorithm and controller. A UVMS platform (see Figure 14) with three joints manipulator, binocular vision, DVL, magnetic compass, four horizontal vector thrusters and two vertical thrusters was equipped. The experiments were performed in the oceanic environment of the Chinese Yellow Sea region close to the Zhangzi Dao Island of Liao Ning Province.



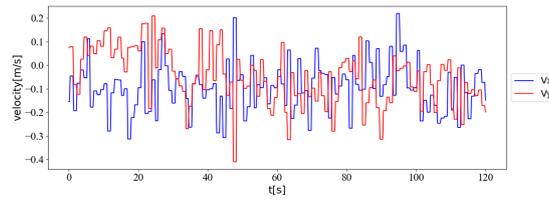
Figure 14. The UVMS platform in the oceanic experiment.

Figure 15 illustrates the process of the oceanic experiments (see Figure 15a). At first, the UVMS is cruising according to the planned trajectory and looking for the sea organism object. When the object has been detected, the UVMS will position the object relative distance from the binocular visual matching of the proposed improved ORB algorithm. Then, the vehicle begins to float approaching the object according to the binocular matching and dead reckoning feedback. Figure 15b shows the current disturbance obtained from DVL during the capture process which is about 0.5 knot in a single direction. Figure 15c–g illustrate the approach and capture control process of the vehicle and end effector; the proposed QBF-ANTSMC controller can realize accurate and stable control under complicated current disturbance and suppress chattering phenomenon. Figure 15h illustrates the process from the diver’s perspective with GoPro underwater camera.

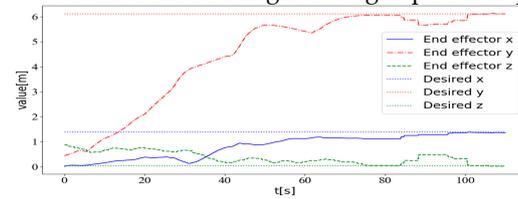


(a) Grasp process in the camera; from the first camera, the object is detected, positioned, the vehicle makes an approaching control in current disturbance and finally makes a stable and accurate capture.

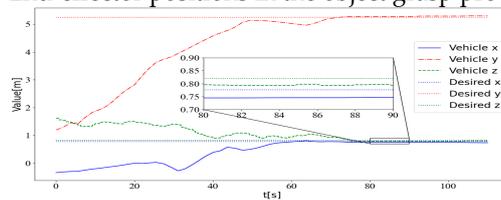
Figure 15. Cont.



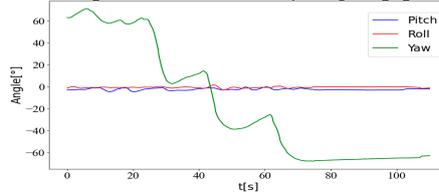
(b) Current disturbance during UVMS grasp control process.



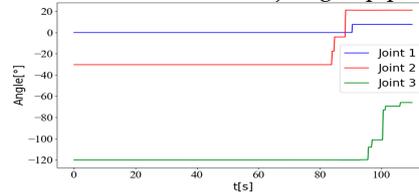
(c) End-effector positions in the object grasp process.



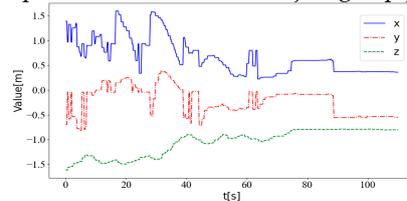
(d) Vehicle positions in the object grasp process.



(e) Vehicle attitude in the object grasp process.



(f) Manipulator attitude in the object grasp process.



(g) Object 3-dimensional distance relative to the binocular camera.



(h) Approaching and grasp process in the GoPro camera from the diver.

Figure 15. Oceanic experiment results.

5. Conclusions

In order to realize vision-based grasp control for small targets, this study has proposed an improved binocular visual measurement and a novel QBF-ANTSMC for UVMS position-based manipulation control. At first, an improved unsharp masking algorithm has been proposed to further enhance the weak texture region for the blurred and low contrast images. Secondly, an improved ORB feature-matching method has been developed with adaptive threshold, non-maximum suppression and improved random sample consensus. Thirdly, an adaptive non-singular terminal sliding mode controller with a quasi-barrier function has been proposed for the unknown bound of unknown disturbance and to suppress the chattering problem. Oceanic experiments with binocular vision measurement and QBF-ANTSMC manipulation control were conducted to prove the performance of the proposed algorithm and controller. However, due to the complexity of the underwater environment, there are still certain errors in the binocular ranging values. In the next step, we will improve the accuracy of the visual binocular ranging by studying new algorithms. The simulation environment and sea trial environment for the control algorithm in this article were relatively flat seabed environments, but obstacles often exist in many situations. In the next step, we will improve the control algorithm for situations where obstacles exist around the target object.

Author Contributions: Conceptualization, T.J. and H.H.; data curation, Y.S., Z.Z. and X.H.; formal analysis, L.L. and X.H.; funding acquisition, H.H., H.Q. and X.C.; investigation, Y.S.; methodology, T.J. and L.L.; project administration, H.H. and H.Q.; resources, H.H., H.Q. and X.C.; validation, Z.Z.; writing—original draft, T.J.; writing—review and editing, T.J. and Y.S. All authors have read and agreed to the published version of the manuscript.

Funding: This project is supported by the National Natural Science Foundation of China (No. U21A20490, 52025111, 61633009, 51779058), and funded by Joint guidance project of Heilongjiang Natural Science Foundation (LH2020F026, YQ2020E033). All of the support provided is highly appreciated.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Prats, M.; Ribas, D.; Palomeras, N.; García, J.C.; Nannen, V.; Wirth, S.; Fernández, J.J.; Beltrán, J.P.; Campos, R.; Ridao, P.; et al. Reconfigurable AUV for intervention missions: A case study on underwater object recovery. *Intell. Serv. Robot.* **2012**, *5*, 19–31. [[CrossRef](#)]
2. Carrera, A.; Palomeras, N.; Hurtós, N.; Kormushev, P.; Carreras, M. Cognitive system for autonomous underwater intervention. *Pattern Recognit. Lett.* **2015**, *67*, 91–99. [[CrossRef](#)]
3. Yang, L.; Zhao, S.; Wang, X.; Shen, P.; Zhang, T. Deep-Sea Underwater Cooperative Operation of Manned/Unmanned Submersible and Surface Vehicles for Different Application Scenarios. *J. Mar. Sci. Eng.* **2022**, *10*, 909. [[CrossRef](#)]
4. Prats, M.; Garc, J.C.; Wirth, S.; Ribas, D.; Sanz, P.J.; Ridao, P.; Gracias, N.; Oliver, G. Multipurpose Autonomous Underwater Intervention: A Systems Integration Perspective. In Proceedings of the 20th Mediterranean Conference on Control & Automation (MED), Spain, Barcelona, 3–6 July 2012; pp. 1379–1384.
5. Choi, J.-K.; Yokobiki, T.; Kawaguchi, K. ROV-Based Automated Cable-Laying System: Application to DONET2 Installation. *IEEE J. Ocean. Eng.* **2018**, *43*, 665–676. [[CrossRef](#)]
6. Wang, Y.; Wang, S.; Wei, Q.; Tan, M.; Zhou, C.; Yu, J. Development of an Underwater Manipulator and Its Free-Floating Autonomous Operation. *IEEE/ASME Trans. Mechatron.* **2016**, *21*, 815–824. [[CrossRef](#)]
7. Razzanelli, M.; Casini, S.; Innocenti, M.; Pollini, L. Development of a Hybrid Simulator for Underwater Vehicles with Manipulators. *IEEE J. Ocean. Eng.* **2020**, *45*, 1235–1251. [[CrossRef](#)]
8. Lynch, B.; Ellery, A. Efficient Control of an AUV-Manipulator System: An Application for the Exploration of Europa. *IEEE J. Ocean. Eng.* **2014**, *39*, 552–570. [[CrossRef](#)]
9. Youakim, D.; Cieslak, P.; Dornbush, A.; Palomer, A.; Ridao, P.; Likhachev, M. Multirepresentation, Multiheuristic A* search-based motion planning for a free-floating underwater vehicle-manipulator system in unknown environment. *J. Field Robot.* **2020**, *37*, 925–950. [[CrossRef](#)]

10. Huang, H.; Tang, Q.; Li, H.; Liang, L.; Li, W.; Pang, Y. Vehicle-Manipulator System Dynamic Modeling and Control for Underwater Autonomous Manipulation. *Multibody Syst. Dyn.* **2017**, *41*, 367–390. [[CrossRef](#)]
11. Rizzini, D.L.; Kallasi, F.; Aleotti, J.; Oleari, F.; Caselli, S. Integration of a stereo vision system into an autonomous underwater vehicle for pipe manipulation tasks. *Comput. Electr. Eng.* **2016**, *58*, 560–571. [[CrossRef](#)]
12. Ridaou, P.; Carreras, M.; Ribas, D.; Sanz, P.J.; Oliver, G. Intervention AUVs: The next challenge. *Annu. Rev. Control.* **2015**, *40*, 227–241. [[CrossRef](#)]
13. Taryudi; Wang, M.-S. Eye to hand calibration using ANFIS for stereo vision-based object manipulation system. *Microsyst. Technol.* **2018**, *24*, 305–3177. [[CrossRef](#)]
14. Chang, J.-W.; Wang, R.-J.; Wang, W.-J.; Huang, C.-H. Implementation of an Object-Grasping Robot Arm Using Stereo Vision Measurement and Fuzzy Control. *Int. J. Fuzzy Syst.* **2015**, *17*, 193–205. [[CrossRef](#)]
15. Chang, W.-C. Robotic assembly of smartphone back shells with eye-in-hand visual servoing. *Robot. Comput. Manuf.* **2018**, *50*, 102–113. [[CrossRef](#)]
16. Peñalver, A.; Pérez, J.; Fernández, J.; Sales, J.; Sanz, P.; García, J.; Fornas, D.; Marín, R. Visually-guided manipulation techniques for robotic autonomous underwater panel interventions. *Annu. Rev. Control.* **2015**, *40*, 201–211. [[CrossRef](#)]
17. Sivčev, S.; Rossi, M.; Coleman, J.; Dooly, G.; Omerdić, E.; Toal, D. Fully automatic visual servoing control for work-class marine intervention ROVs. *Control. Eng. Pract.* **2018**, *74*, 153–167. [[CrossRef](#)]
18. Lin, Y.H.; Shou, K.P.; Huang, L.J. The initial study of LLS-based binocular stereo-vision system on underwater 3D image reconstruction in the laboratory. *J. Mar. Sci. Technol.* **2017**, *22*, 513–532. [[CrossRef](#)]
19. Li, T.; Liu, C.; Liu, Y.; Wang, T.; Yang, D. Binocular stereo vision calibration based on alternate adjustment algorithm. *Opt.-Int. J. Light Electron Opt.* **2018**, *173*, 13–20. [[CrossRef](#)]
20. Hu, Y.; Chen, Q.; Feng, S.; Tao, T.; Asundi, A.; Zuo, C. A new microscopic telecentric stereo vision system- Calibration, rectification, and three-dimensional reconstruction. *Opt. Lasers Eng.* **2019**, *113*, 14–22. [[CrossRef](#)]
21. Park, J.-S.; Kim, H.-E.; Kim, H.-Y.; Lee, J.; Kim, L.-S. A vision processor with a unified interest point detection and matching hardware for accelerating stereo matching algorithm. *IEEE Trans. Circuits Syst. Video Technol.* **2016**, *26*, 2328–2343. [[CrossRef](#)]
22. Huo, G.; Wu, Z.; Li, J.; Li, S. Underwater Target Detection and 3D Reconstruction System Based on Binocular Vision. *Sensors* **2018**, *18*, 3570. [[CrossRef](#)] [[PubMed](#)]
23. Negahdaripour, S.; Firoozfam, P. An ROV Stereovision System for Ship-Hull Inspection. *IEEE J. Ocean. Eng.* **2006**, *31*, 551–564. [[CrossRef](#)]
24. Ttofis, C.; Kyrkou, C.; Theocharides, T. A Low-Cost Real-Time Embedded Stereo Vision System for Accurate Disparity Estimation Based on Guided Image Filtering. *IEEE Trans. Comput.* **2016**, *65*, 2678–2693. [[CrossRef](#)]
25. Zhuang, S.; Zhang, X.; Tu, D.; Zhang, C.; Xie, L. A standard expression of underwater binocular vision for stereo matching. *Meas. Sci. Technol.* **2020**, *31*, 115012. [[CrossRef](#)]
26. Li, Y.; Zhang, Y.; Li, H.; Zhang, W.; Zhang, Q. Epipolar geometry and stereo matching algorithm for underwater fish-eye images. *Int. J. Adv. Robot. Syst.* **2018**, *15*, 1729881418764715. [[CrossRef](#)]
27. Rosten, E.; Porter, R.; Drummond, T. Faster and Better: A Machine Learning Approach to Corner Detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **2010**, *32*, 105–119. [[CrossRef](#)]
28. Rublee, E.; Rabaud, V.; Konolige, K.; Bradski, G. ORB: An efficient alternative to SIFT or SURF. In Proceedings of the 2011 IEEE International Conference on Computer Vision, Barcelona, Spain, 6–13 November 2011; pp. 2564–2571.
29. Osipov, A.; Shumaev, V.; Ekielski, A.; Gataullin, T.; Suvorov, S.; Mishurov, S.; Gataullin, S. Identification and Classification of Mechanical Damage during Continuous Harvesting of Root Crops Using Computer Vision Methods. *IEEE Access* **2022**, *10*, 28885–28894. [[CrossRef](#)]
30. Li, D.; Du, L. AUV Trajectory Tracking Models and Control Strategies: A Review. *J. Mar. Sci. Eng.* **2021**, *9*, 1020. [[CrossRef](#)]
31. Haugaløkken, B.O.A.; Skaldebo, M.B.; Schjølberg, I. Monocular vision-based gripping of objects. *Robot. Auton. Syst.* **2020**, *131*, 103589. [[CrossRef](#)]
32. Cai, M.; Wang, S.; Wang, Y.; Wang, R.; Tan, M. Coordinated Control of Underwater Biomimetic Vehicle–Manipulator System for Free Floating Autonomous Manipulation. *IEEE Trans. Syst. Man Cybern. Syst.* **2020**, *51*, 4793–4803. [[CrossRef](#)]
33. Li, J.; Huang, H.; Xu, Y.; Wu, H.; Wan, L. Uncalibrated visual servoing for underwater vehicle manipulator systems with an eye in hand configuration camera. *Sensors* **2019**, *24*, 5469. [[CrossRef](#)] [[PubMed](#)]
34. Chen, H.-T.; Song, S.-M.; Zhu, Z.-B. Robust Finite-time Attitude Tracking Control of Rigid Spacecraft under Actuator Saturation. *Int. J. Control. Autom. Syst.* **2018**, *16*, 1–15. [[CrossRef](#)]
35. Boukattaya, M.; Mezghani, N.; Damak, T. Adaptive nonsingular fast terminal sliding-mode control for the tracking problem of uncertain dynamical systems. *ISA Trans.* **2018**, *77*, 1–19. [[CrossRef](#)]
36. Kong, S.; Fang, X.; Chen, X.; Wu, Z.; Yu, J. A NSGA-II-Based Calibration Algorithm for Underwater Binocular Vision Measurement System. *IEEE Trans. Instrum. Meas.* **2020**, *69*, 794–803. [[CrossRef](#)]
37. Lwin, K.N.; Minami, M.; Mukada, N.; Myint, M.; Yamada, D.; Yanou, A.; Matsuno, T.; Saitou, K.; Godou, W.; Sakamoto, T. Visual Docking against Bubble Noise with 3-D Perception Using Dual-Eye Cameras. *IEEE J. Ocean. Eng.* **2020**, *45*, 247–270. [[CrossRef](#)]
38. Marani, G.; Yuh, J. *Introduction to Autonomous Manipulation*; Springer: Berlin/Heidelberg, Germany, 2014.
39. Antonelli, G. *Underwater Robots*; Springer: Berlin/Heidelberg, Germany, 2003.

40. Chen, Y.; Yan, L.; Xu, B.; Liu, Y. Multi-Stage Matching Approach for Mobile Platform Visual Imagery. *IEEE Access* **2019**, *7*, 160523–160535. [[CrossRef](#)]
41. Kamyshova, G.; Osipov, A.; Gataullin, S.; Korchagin, S.; Ignar, S.; Gataullin, T.; Terekhova, N.; Suvorov, S. Artificial Neural Networks and Computer Vision's-Based Phytoindication Systems for Variable Rate Irrigation Improving. *IEEE Access* **2022**, *10*, 8577–8589. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.