Vision Based Object Following for Underwater Vehicles

(Part 2: Collaborative Control with Pan and Tilt Camera Mechanism)

by Arjuna Balasuriya* Tamaki Ura*, *Member*

Summary

This paper proposes a vision based guidance system for Underwater Vehicles (UVs) to follow moving targets by collaborating with a pan and tilt camera mechanism. The proposed system can be divided into three main sections. The vision section provides the position information of the UV with respect to the object being tracked. A commercially available CCD camera mounted on the pan and tilt mechanism attached to the UV provides the environmental information as a two-dimensional image and the proposed vision algorithm detects the position of the interested target in that image. By the pan and tilt calibration, a transformation matrix is derived for the image pixel to pan/tilt angle relationship. The second section consisting of the pan and tilt mechanism, brings the interested object in the image to a predefined pixel position. The pan action is derived from the horizontal pixel movement and the tilt action from the vertical pixel movement. As a result of this action, the camera unit will be looking in a direction other than the UVs forward direction. The third section consisting of the UV controller, aligns the camera unit in the forward direction of the UV by controlling the vehicle. An experiment is conducted to demonstrate the performance of the proposed algorithm using the test-bed UV available at the University of Tokyo called the Twin-Burger 2. The vision algorithm was implemented on the MaxVideo 250 image processing board. On these hardware, the proposed system successfully managed to track objects moving with a maximum speed of 4 cm/sec.

1. Introduction

Underwater object following becomes a key issue for underwater vehicles in tasks like manipulation work on underwater structures, refuelling/battery charging of Autonomous Underwater Vehicles (AUVs), study of different species, cooperation between other agents etc.. To make UVs intelligent enough to carry out such tasks autonomously and reliably, it is necessary to gather as much information as possible from the highly complex, unstructured underwater environment. This paper proposes an algorithm based on the data captured by

commercially available Charge Couple Device (CCD) camera to navigate UVs autonomously. This is a continuation of the previous work reported in references 2) and 6), where UV navigation with respect to a stationary object was considered.

Here, an algorithm is proposed to track an object which moves with respect to the vehicle in an untextured environment¹⁾. A fast tracking performance is necessary when there is relative movement of the object with respect to the UV. This algorithm assumes that the features of the interested object is definable. The special features in the 2D image frame are used to generate 3D control targets for the navigation of the UV. This algorithm acquires the image using a single CCD camera, extracts the image features, recognizes the desired object, and interprets the pose (relative position) of the object which is used to control the UV.

The objective of the proposed algorithm is to track an

*Institute of Industrial Science, University of Tokyo

Received 9th Jan. 1998

Read at the Spring meeting 14, 15th May 1998

object with a rectangular shape. The UV is navigated by deriving the heave, heading and surge control targets from the image of the interested object. The interested object is defined in the image by the following two features 1) shape and 2) movement. The difference in the pixel position of the center of the interested object is used to control the pan and tilt mechanism of the camera. The pan and tilt angles are then used to derive the **yaw** and **heave** control targets, respectively. The area within the shape of the object in the image is used to derive the **surge** control target of the UV as shown in Fig. 1.

Fig. 1: The Flow Diagram of the Feature Based Tracking System

The paper is organized as follows; section 2 describes the vision algorithm proposed to detect the target in the image. This algorithm generates the position of the target and the area covered by the target in the image. Section 3 explains the criterion used to control the pan and tilt mechanism of the camera to bring the location of the target in the image to a defined position. Due to panning and tilting, the camera points to a different direction than the forward direction of the UV. Section 4 describes the method used to bring the camera to the forward looking position by controlling the UV. Here, pan and tilt corresponds to the yaw and heave motion of the UV. The area covered by the object in the image corresponds to the surge motion of the UV. Section 5 presents the experimental results obtained by implementing the proposed system in the test-bed vehicle Twin-Burger 2⁵⁾ to track a moving cylindrical object.

2. Vision System

As shown in Fig. 1, the image captured by the CCD camera branches in to two directions, namely; feature selection and movement detection. The proposed algorithm looks into two main properties in the image; 1) geometrical features and 2) dynamic behaviour. The two branches simultaneously extracts these properties from the image. In the feature selector, the image is pre-filtered to make sure that the features of the interested object are highlighted and unwanted features are removed. At first, the gain of the camera is adjusted by grayscale comparison as shown in Fig. 2.

Fig. 2: Image Gain Tuning

Next, the image is pre-filtered using the Signum of Laplacian of Gaussian(SoLoG) filter¹⁾²⁾³⁾⁴⁾. The SoLoG operation on the image produces a binary map. The mission is initiated, by placing the interested object in a search window in the image. Using the modified Hough transformation technique, the location of the target is selected as shown in Fig. 3.

Fig. 3: Search Window for Feature Detection

The search window is divided into four sections, symbolized as S1, S2, S3 and S4 shown in Fig. 3(b). The projection of a cylindrical object on to a 2 dimensional plane is a rectangle. The four sections of the search window determines the four sides of the projection. To avoid mis-interpretations, a geometrical check is done on the sides detected by the proposed algorithm. If the lengths of the sides detected are denoted by L1, L2, L3 and L4, then,

a) $L1 = L2, L3 = L4,$

b) $L1/L3 = L2/L4 = \text{ratio of the physical}$

dimensions of the object $\pm$ small deviation.

Once these conditions are satisfied, the location of the target, centroid of the rectangle (X_{Gr}, Y_{Gr}), and the area(A) covered by the rectangle in the image are calculated.

In movement detection, the motion-energy of the pixels enclosed by the search window is extracted. Movement is derived by considering two consecutive images. The flow chart of the motion-energy based method proposed, is shown in Fig. 4. Unlike the feature detection algorithm explained earlier, this algorithm tracks any moving object regardless of its size or shape. Only the position of the object is extracted in the image. As shown in Fig. 4, A and B are consecutive images where B is the current image and A is the previous image. Motion energy is calculated by subtracting A from B. To extract the target, an AND operation is performed between the difference in image and the edge map obtained by the SoLoG filter. The resulting image consists of only the interested target. By moment calculations, the centroid (X_{Gm}, Y_{Gm}) of the target in the image is calculated.

Fig. 4: Motion-Energy Based Object Tracking

Feature based tracking system provides three parameters (X_{Gr}, Y_{Gr}, A) and if the detected features satisfy the geometrical check, then these parameters are used for the next stage of the algorithm. If not, the (X_{Gm}, Y_{Gm}) values are used to control the pan and tilt mechanism and surge target is set to zero.

3. Pan and Tilt Control

Calibration of the pan and tilt movement with respect to the pixel displacement in the image is carried out on a target placed at different distances from the camera. Depending on the required distance at which the object following is carried out, the set of mapping parameters which relates the pixel displacement to pan/tilt angles are calculated. By denoting the mapping parameters as K_{pan} and K_{tilt} , the derivation of the pan and tilt angles can be written as;

$$\text{pan angle} = K_{pan} \Delta X_G \quad \text{and} \quad \text{tilt angle} = K_{tilt} \Delta Y_G \quad (1)$$

Where, (X_G, Y_G) is the centroid value derived from the vision algorithm and $\Delta X_G = (X_{target} - X_G)$, $\Delta Y_G = (Y_{target} - Y_G)$ as shown in Fig. 5. The (X_{target}, Y_{target}) is the desired target location in the image.

Fig. 5: Vision Feedback Pan/Tilt Control

4. AUV Control

Yaw and heave movement of the UV brings the pan and tilt angles to zero positions respectively. The surge movement keeps the area of the object appearing in the image at a constant value range. In turn, the relative distance between the UV and the target is controlled. A calibration is done to derive the parameters which relates the pan/tilt angles to yaw/heave values and the pixel area in the image covered by the target at the required relative distance between the UV and the target is recorded. By denoting the yaw and heave parameters as $K_{heading}$ and K_{heave} , yaw and heave angles can be calculated as given in Eq. (2).

$$\Delta \text{heading} = K_{heading} \times (\text{pan angle}) \quad \text{and},$$

$$\Delta \text{heave} = K_{heave} \times (\text{tilt angle}) \quad (2)$$

Therefore,

$$\text{yaw target} = \text{yaw angle} + \Delta \text{heading} \quad \text{and},$$

$$\text{heave target} = \text{depth} + \Delta \text{heave} \quad (3)$$

The change of the area covered by the detected object in the image is used to trigger the surge control of the UV.

A switching controller is used for the surge action of the UV. The three speeds, forward, stop and backward are activated depending on the size of the object appearing in the image as shown in Fig. 6.

surge = forward; if area of object < desired area - ϵ
 surge = stop; if area of object < desired area + ϵ ,
 and,
 area of object > desired area - ϵ
 surge = backward; if area of object > desired area + ϵ

Fig. 6: Surge Target Generation

5. Experiment

An experiment is carried out to demonstrate the performance of the proposed algorithm. A cylindrical object (9 cm f x 9 cm) shown in Fig. 7 is used as the target. The vision algorithm discussed in section 2 is implemented on the MaxVideo 250 image processing board and the test-bed AUV "Twin-Burger 2" available at the University of Tokyo⁶⁾, shown in Fig. 8, is used to implement the control algorithms.

Fig. 7: Tracking Target

The hardware used and the experimental set up is shown in Fig. 9. The (X_G, Y_G, A) from the vision algorithm executed in the MaxVideo system, are sent through the RS232 line to the on-line monitoring computer in which the Twin-Burger server program is executed. The root transputer in the on-line monitoring computer is connected to the RS422 line via the converter box. The

RS422 line carries data to and from the Twin-Burger. The image data, captured by the CCD camera placed in front of the Twin-Burger, are sent to the MaxVideo system through a NTSC line.

Fig. 8: The Twin-Burger 2

Fig. 9: Hardware System Used in the Experiment

The experiment is carried out in the testing pool (5m x 20m x 1.5m) at the Institute of Industrial Science, University of Tokyo. This experiment demonstrates the yaw and surge control performance with the vision algorithm proposed. The heave target is kept constant due to the space limitation in the experimental tank. Therefore the object is moved in a horizontal plane. The movement of the object is done manually. The trajectory of the object is recorded by a position sensor attached to it. Due to the contrast limitation in the underwater environment, the separation between the camera and the tracking object is selected to be 1m. The area covered by the object in the image when it is 1m away from the camera is around 4000 square pixels.

The size of the search window is selected as 100 x 100 pixels. The maximum pixel displacement for this window size is therefore 50 pixels. From camera calibration it is found that 50 pixel displacement in the image of an object, moving in a vertical plane 1m away from the camera perpendicular to the camera axis, corresponds to 15 cm movement. The processing time of the proposed algorithm on the MaxVideo hardware is achieved at 0.1

sec. Therefore, this image processing system on MaxVideo 250 with a search window size of 100×100 pixels can track objects moving at a maximum speed of 1.5 m/sec in a vertical plane perpendicular to the camera axis 1m from the camera. The desired location of the target is selected as $X_{\text{target}} = 230$ pixels and $Y_{\text{target}} = 200$ pixels.

The pan and tilt mechanism on the Twin-Burger 2 has a local hardware servo, to which the feedback is connected from the potentiometers mounted on the pan/tilt rotating axis as shown in Fig. 10.

Fig. 10: Twin-Burger CCD Camera on the Pan and Tilt Mechanism

This camera mechanism has a pan span of 22° and a tilt span of 80° . The maximum pan angle is selected as 11° since the pan zero position is at the middle of the pan span. From calibration, it is found that this mechanism can track an object moving with a maximum speed of 19 cm/sec in a vertical plane perpendicular to the camera axis 1m away from the camera. From calculations, $K_{\text{pan}} = 0.02$. The pan angle is calculated using Eq. (1).

A Proportional Derivative (PD) controller is used to control the heading of the Twin-Burger 2. Considering the dynamics of the Twin-Burger 2, the parameter $K_{\text{heading}} = 0.07$ and corresponding maximum speed of tracking is 4 cm/sec. Therefore in this experiment the object is moved at a speed less than 4 cm/sec. Heading target is calculated using the Eq. (3).

The heave target is kept at 0.8m and the surge target is generated as explained in Fig. 6. The surging speed is selected as ± 0.1 m/s. The area of the object in the image at the desired distance is about 4000 pixel². Therefore, a ± 800 is added as the differential gap to avoid too frequent on/off operations of the switching controller.

6. Results

Since the mission is carried out at a constant depth of 0.8m, Twin-Burger 2 initially dives using the heave controller to that depth by keeping its heading direction in the positive X - direction along the $Y = 0$ line and the initial position of the tracking object at $X = 0$, $Y = 0$. Once the operational depth is reached, the Twin-Burger starts the tracking operation. The following results are obtained at 200 sec. At the end of the mission time, the Twin-Burger surfaces using the heave controller. PD controllers are used for heading and heaving.

Fig. 11: Step Response of the Pan Mechanism

Figure 11 shows the response of the camera pan mechanism of the Twin-Burger 2. It can be seen that the pan mechanism has a short settling time enabling quick response, which is advantageous for fast tracking operations. The response of the Twin-Burger 2 for different K_{heading} values are shown in Fig. 12.

Fig. 12: Step Response of the Heading Controller

It can be seen that the response of the Twin-Burger is stable with the parameter $K_{\text{heading}} = 0.07$. With this value, Twin-Burger has a settling time of about 12sec for a 0.2 rad yaw target change.

Fig. 13: XY Trajectory of the Interested Object

Figure 13 shows the trajectory of the target object during the mission. The speed of the object by manual movement is shown in Fig. 14.

Fig. 14: Speed of the Interested Object

The speed of the target is kept below 4 cm/sec, which is the maximum speed of tracking of the Twin-Burger. The dynamic behaviour of the Twin-Burger 2 during the mission is shown in Figs. 15 and 16.

During the tracking of the target, Twin-Burger 2 maintains a constant depth using a PD controller. Since X coordinate of the image plane corresponds to the pan and heading control, pixel variation of the centroid of the target in the image is shown on the top graph of Fig. 15. It can be seen that the pan action keeps the target around the desired pixel position in the image. The corresponding pan action is shown in the middle graph

Fig. 15: Yaw Control of the Twin-Burger 2 During Tracking

Fig. 16: Surge Control of the Twin-Burger During Tracking

of Fig. 15. The camera position with respect to the pan movement is kept at the desired pan angle by the control of the Twin-Burger as shown in the lower graph of Fig. 15. As shown in this figure, the PD controller performed well to keep the Twin-Burger on track with the derived heading targets.

The top graph of Fig. 16 shows the variation of the area covered by the target object in the image, during the object tracking mission. The derivation of surge control targets and the surging speeds achieved by the Twin-Burger 2 is shown in the lower graph of Fig. 16. The peaks are due to the noise introduced by the flow sensors in the surge feedback. The sampling time of the control loop is kept at 0.1 sec (10 Hz).

7. Conclusions

Vision based autonomous tracking of a moving underwater

object with an UV is proposed and demonstrated on the test-bed system available at the University of Tokyo. The system demonstrates how to extract useful information for navigation of an UV in the three dimensional world, using the two dimensional image data captured by a commercially available single CCD camera. This also demonstrates the collaborative control between the UV and the pan/tilt camera mechanism. From the results it can be seen that the speed of tracking of the proposed system is hardware dependent. In the experiment conducted, an object which is at a 1m distance is considered due to the contrast difficulties. Therefore, it can be concluded that visual sensing is suitable for short range sensing. On the present hardware, the proposed algorithm is suitable to track objects which move very slowly with respect to the UV. A typical example of this nature can be encountered in underwater manipulation tasks. In this system, the features of the tracking object should be definable and a cylindrical object is considered. Authors intend to improve the system to cater objects whose features are undefinable and changes shape as they move in the environment.

- 6) M. Takai, B.A.A.P. Balasuriya, C. Lam, T. Ura, and Y. Kuroda, "Visual Feedback Navigation for Cable Tracking by Autonomous Underwater Vehicles", J. of the society of Naval Architects of Japan, vol. 181, pp. 213-219, June 1997.(in Japanese).

Acknowledgement

The authors would like to thank Mr. Segawa and Mr. Nose for the assistance given during the experiments conducted.

References

- 1) B.A.A.P. Balasuriya, "Computer Vision for Autonomous Underwater Vehicle Navigation", Doctor Thesis, Univ. of Tokyo, March 1998 (under review).
- 2) B.A.A.P. Balasuriya and Tamaki Ura, "Vision Based Object Following for Underwater Vehicles", J. of the society of Naval Architects of Japan, vol. 180, pp. 163-668, December 1996.
- 3) B.A.A.P. Balasuriya, M. Takai, W.C. Lam, T. Ura and Y. Kuroda, "Vision Based Autonomous Underwater Vehicle Navigation: Underwater Cable Tracking", Proc. of OCEANS'97, Nova Scotia, Canada, pp. 1418-1424, October 1997.
- 4) D. Marr and E. Hildreth, "Theory of Edge Detection", Proc. of Royal Society London, pp. 187-217, 1980.
- 5) T. Fujii, T. Ura and Y. Kuroda, "Mission Execution Experiments with a Newly Developed AUV "Twin-Burger"", Proc. of UUST'93, Durham, New Hampshire, September 1993.