VISION BASED AUTONOMOUS UNDERWATER VEHICLE NAVIGATION: Underwater Cable Tracking

B.A.A.P. Balasuriya, M. Takai, W.C. Lam, Tamaki Ura and Y. Kuroda,
Ura Lab, Institute of Industrial Science, University of Tokyo,
7-22-1, Roppongi, Minato-ku, Tokyo 106, Japan
Email: arjuna@manta.iis.u-tokyo.ac.jp

Abstract

In this paper, a vision based Autonomous Underwater Vehicle (AUV) navigation system is proposed and its performance is demonstrated by an Autonomous Underwater Cable Tracking (AUCT) mission at the lake Biwa using the test bed AUV the "Twin-Burger" available at the University of Tokyo. The proposed system for the AUCT consist of three levels of controllers; namely, Higher level controller, Middle level controller and the Lower level controller to navigate the AUV. A single CCD camera and an acoustic sensor is used to determine the relative position of the underwater cable with respect to the AUV in a 3 dimensional space. Visual data provides two dimensions and acoustic data provides the third dimension. An image filtering technique to reduce some undesirable features which are unique to underwater images, is used based on the Laplacian of Gaussian (LoG) operator. A real-time algorithm is developed to determine the position of the cable in the image plane by cascading an estimator governed by AUV dynamics, with the Hough transformation technique. At the lake Biwa mission, the proposed system performed very well other than in situations where the cable was covered by waterweeds for a long distance. The proposed system can be used as a solution to the problem of underwater positioning especially in pre-determined areas.

I. Introduction

Autonomous Underwater Vehicles (AUVs) offer a cost-effective alternative to the Remotely Operated Vehicles (ROVs) for deep ocean work. The maneuverability of an AUV is high compared to a ROV due to the untetheredness. AUVs can be used in deep ocean work platforms and ocean bottom exploration, sampling and underwater surveying systems etc.[2]. At present, research work is carried out in different fields to improve the reliability of the autonomous underwater systems. This paper focuses on a system, based on visual data captured by commercially available CCD cameras to boost the reliability and the intelligence of AUV navigation. The performance of the proposed system is demonstrated in an Autonomous Underwater Cable Tracking (AUCT) mission at lake Biwa using the AUV, the "Twin-Burger"[3].

Humans and most of the autonomous land and space systems use visual data for navigation as it carries much information on the environment[6]. So far there have been only a few applications underwater which use visual data for autonomous navigation. It is due to three factors: (1) In the past, vision processing hardware occupied a large space which wouldn't fit the small pressure hulls used by AUVs. (2) The number of applications which needed close range data was limited and the number of AUVs were small. (3) Vision processing algorithms available for land/space systems cannot be applied directly for underwater operations due to the special features introduced underwater which don't satisfy most of the assumptions made by these algorithms.

In this paper, a filtering technique is developed based on the Laplacian of Gaussian (LoG) operator to reduce the effects of the undesirable features in underwater images enabling vision processing for underwater applications[1]. The undesirable features are divided into three main categories: (1) Non-uniformity in lighting due to the use of highly directional active light sources. (2) Marine snow effect due to suspended particles in water. (3) Limited range due to light absorption by water. In the frequency domain, these features appear in the low and high frequency ranges. Therefore the LoG filter is designed in such a way that it acts as a band pass filter.

Unlike land or space systems, underwater positioning is a very difficult task. In this paper we look at the use of an image sensor (CCD camera) and an acoustic sensor for the purpose of positioning an AUV with respect to an object (underwater cable). The cable is used to indicate the desired path for the AUV which will navigate autonomously by deriving its control commands from the relative distance with respect to the cable. Such a system would be important in routine maintenance operations or data accumulation etc. avoiding complicated path planning algorithms. In addition, maintaining the interested object itself (eg. pipe-lines, communication cables etc.) could also be achieved. In the proposed algorithm, the interested object is located in the image frame (2D) captured by the CCD camera using the Hough transformation method. The image feature extraction was achieved in real-time by using an estimator based on AUV dynamics. The relative distance to the interested object is then found by mapping the 2D data from the image to 3D (vehicle's coordinates) data by using an acoustic sensor and a camera model designed for the CCD camera.

The relative positon of the cable with respect to the AUV is used to generate the necessary targets to navigate the AUV. The proposed system for the AUCT consist of three levels of controllers: (1) A higher level controller to decide different

0-7803-4108-2/97/$10.00 ©1997 IEEE

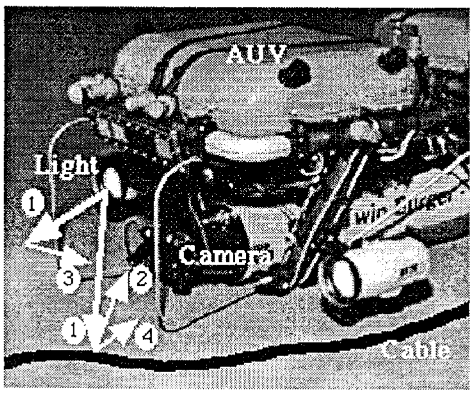
steering modes including a search mode in case the cable is invisible to the AUV. (ii) A middle level controller to generate steering targets for the modes in (1) by fusing sensor data captured by the AUV system. (iii) A lower level controller for the actuators in the AUV system to achieve the targets in (ii).

II. Underwater Visual Sensing

Visual sensing plays a main role in autonomous land systems as it is unobtrusive and accurate, and has a high bandwidth. There are, however, limitations due to both the environment and currently available technology. Due to the little or no ambient light underwater, the visibility is poor so that active light sources have to be used to illuminate the working environment resulting in non-uniform lighting. In case of AUVs, it is almost impossible to place large hardware systems in their small pressure hulls.

A. Underwater Image Filter

An underwater image consists of several light sources reflected from the environment as shown in Fig.1.



1 Incident ray
2 Direct component
3 Backscatter
4 Forward scatter (blur)

Fig. 1: Underwater Optical Properties

The linear superposition of the three reflected quantities shown in Fig. 1 are considered in the final image formed in the camera plane;

$$I_t(x,y) = I_d(x,y) + I_{bk}(x,y) + I_b(x,y) \qquad (1)$$

where,
I_t = the total intensity,
I_d = the direct component: the light reflected from the target,
I_{bk} = the backscatter component: light scattered by water and suspended particles with no illumination from the object, and
I_b = the blur component (forward scatter): the light reflected from the target can also be scattered on its way to the camera.

By considering each function[5];

$$I_d(x,y) = \frac{I_r}{R'^2}(x,y,z)\ K(\text{camera properties})\ e^{-cR}$$

$$I_b(x,y) = \frac{I_d}{e^{-cR}}(e^{-GR'} - e^{-cR'}) * F^{-1}(e^{-BR'f}) \qquad (2)$$

$$I_{bk}(x,y) = \sum_{n=1}^{Z/\Delta Z} I_{bp}(x,y,z)$$

where

I_r - reflectance intensity from the target
I_{bp} - the intensity at the backscatter plane
K - is a function depending on the camera used
c - total attenuation coefficient
R - distance from the light source to the target
R' - distance from the target to the camera
G - empirical constant related to the power in the scattered component
B - empirical constant related to the angular attenuation
f - angular frequency in radians
F^{-1} - inverse Fourier transform
$*$ - convolution
Z - distance from the light source to the camera
ΔZ - increment in the z-direction

By considering the frequency response of the above optical properties it is clear that underwater images have spatial attenuation of high frequencies (blur) and predominance of low frequencies (backscatter). A filter should be designed to filter out some of the low frequency components and to boost some high frequency components which were attenuated.

The above mentioned optical properties can be catogorized into main three physical features;

1) Non-uniform lighting

Due to the light absorbed by water, far ranges will appear to be darker than the close ranges and highly directional active light sources make regions in the image to appear darker and brighter. This will result in a non-uniform lighting condition in the underwater image which is undesirable for most of the available computer vision algorithms.

2) Suspended particles in water

The backscattering is mostly due to the randomly moving particles suspended in water. As a result, the image consists of many bright spots. In case of marine snow the image could be improved with the use of filtering techniques. But in case of high turbidity, quality of the image will be difficult to improve.

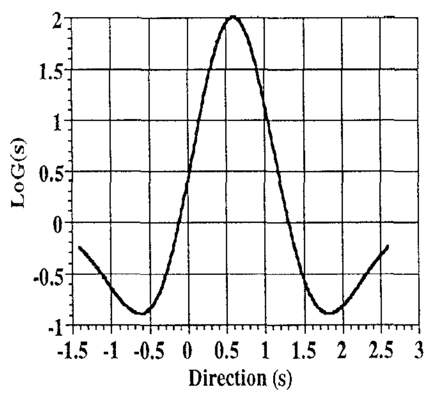
3) Limited range

Due to absorption of light, the range of visibility underwater is limited.

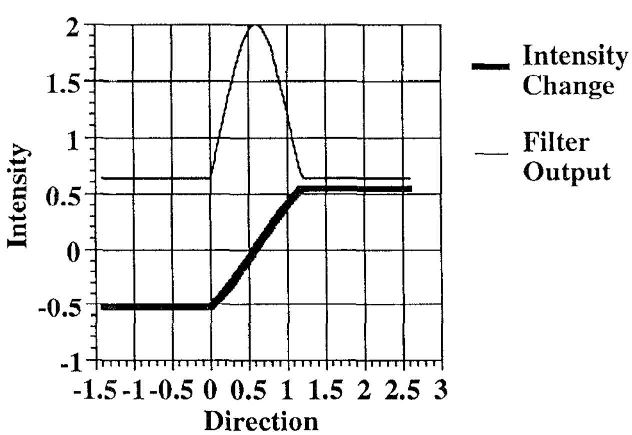
In this paper, a band pass filter based on the LoG operator[4] is introduced to minimize the effects of the above mentioned features which dominate in the high and low frequency domains. In this filter, the Laplacian portion will act as a high pass filter and the Gaussian as a low pass filter. Main advantage of this filter is that it can be implemented in real-time applications. The LoG mask in two dimensions is given in (3).

$$\nabla^2 G(x,y) = \left[K\left(2 - \frac{(x^2+y^2)}{\sigma^2}\right) e^{-\frac{(x^2+y^2)}{2\sigma^2}} \right] \tag{3}$$

where σ is the space constant of the Gaussian, K is a scale factor and (x,y) is the pixel position in the image. The size of the LoG operator, $s = 8.5\sigma$.



(a) LoG Mask



(b) Response on an Intensity Variation

Fig. 2: LoG Operator

The characteristic curve of the LoG mask is shown in Fig. 2(a) and the filter response on an intensity variation in the image, in a particular direction is shown in Fig. 2(b). Note how the filter evaluates a zero at the local maxima. The zero crossings are locally more stable in the presence of noise than most other properties, thus making them more attractive for underwater images. Therefore, a signum function to extract the zero crossings from the LoG filtered image is proposed. The position of the underwater cable is then determined from the edge map generated from the signum of LoG operation.

B. Underwater Cable Recognition

The image of an underwater cable produces a high pixel concentration in a particular direction due to the fact that there are very few well defined geometrical shapes in an underwater environment. In this paper, the direction of the high pixel concentration in the image is modelled as the underwater cable. The position of the model is extracted by transforming the pixel data into a parametric space which is know as the Hough transform. The parameters are defined as shown in Fig. 3.

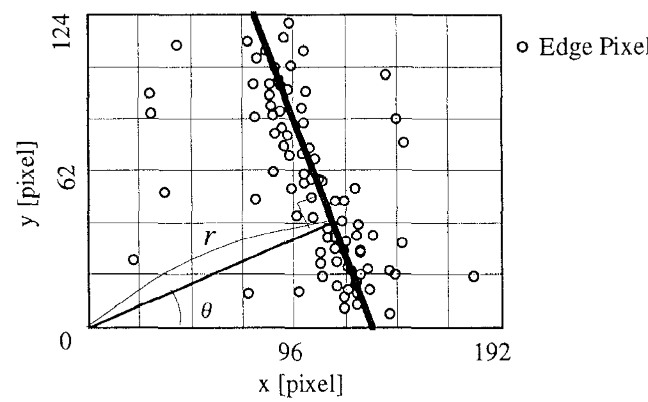


Fig. 3: Edge Map of an Image Plane

A particular direction in an image plane (2D) is a straight line and the parameter r is the perpendicular distance falling onto the line from the origin of the image plane and θ is the angle made by the perpendicular with the positive direction of the x-axis as shown in Fig. 3. Once the edge pixels are transformed into the parametric plane (Hough plane r, θ), the parameters corresponding to the direction of highest pixel concentration can be extracted from the Hough plane which will indicate the position of the cable in the image plane, (4).

$$r = x\cos(\theta) + y\sin(\theta) \tag{4}$$

The 2D information of the cable is then transformed into the 3D space (AUV coordinates), using the transformation matrix of the camera mounted on the vehicle. The depth of the cable with respect to the AUV is achieved by the downward looking acoustic sensor. By the fusion of these data, the 3D position of the cable with respect to the AUV is calculated.

C. Vision Estimator

In the above section it is assumed that there are no well defined geometric shapes underwater. However, in a real environment situation there is a possibility of having objects other than the cable with similar features. To avoid misinterpretations, an estimator is developed using the pin hole camera model and the dynamics of the AUV to estimate the where abouts of the cable in the following image frame.

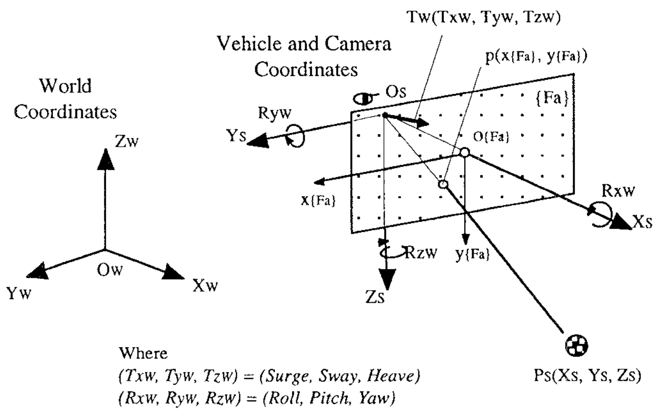


Fig. 4: Vehicle and Camera Coordinates

The motion parameters of the object of interest which is to be followed can be derived with the pin-hole camera model shown in Fig. 4. A point P_s projects into a point p in the image plane with image coordinates (x,y) given by;

$$x = X_s/Z_s \quad and \quad y = Y_s/Z_s \tag{5}$$

The perspective projection and the focal length is assumed to be unity. If the vehicle moves in a static environment with a translational velocity $T_w(T_{xw}, T_{yw}, T_{zw})$ and with an angular velocity $R_w(R_{xw}, R_{yw}, R_{zw})$ with respect to the vehicle frame, then the velocity of point P_s can be derived as;

$$\frac{dP_s}{dt} = -T_w - R_w \times P_s \tag{6}$$

By taking the time derivatives of the expressions for x and y and using (5) and (6), we obtain (7) which is used as the motion estimator in the proposed tracking algorithms.

$$\dot{x} = \left[x\frac{T_{xw}}{X_s} - \frac{T_{yw}}{X_s} \right] + R_w \begin{pmatrix} x \\ xy \\ -(1+x^2) \end{pmatrix}$$
$$\dot{y} = \left[y\frac{T_{xw}}{X_s} - \frac{T_{yw}}{X_s} \right] + R_w \begin{pmatrix} -x \\ 1+y^2 \\ -xy \end{pmatrix} \tag{7}$$

With the use of the proposed estimator, the cable is searched only in the estimated region and as a result the image processing time is reduced and avoids the misinterpretation of any other object as the cable. Delay time due to image processing can be compensated using this estimator.

III. Twin-Burger Operation for AUCT

The software of the Twin-Burger for the AUCT is divided into three main control strategies; High, Middle and Low level controllers.

A. High level controller

This level of controller is called the steering mode and is divided further into four modes of operations;

1) Waiting Mode: At the begining, the AUV is in this mode until it is triggered by the transponder to start the mission. Once the AUV is triggered it will switch into the descending mode.

2) Descending Mode: The AUV dives until the height from the cable is 1.5m and switches into the tracking mode.

3) Tracking Mode: In this mode the AUV tracks the underwater cable and if the distance of tracking is covered then the AUV switches to the ascending mode. If the cable is lost then the searching mode is activated.

4) Searching Mode: In this mode the Twin-Burger searches for the cable by starting from the direction by which the cable was last seen in the image. In this mode the vehicle stops surging and searches the cable by changing the heading directions.

5) Ascending Mode: In this mode, AUV ascends to the water surface and switches into the waiting mode.

The transition diagram of the steering mode is shown in Fig. 5.

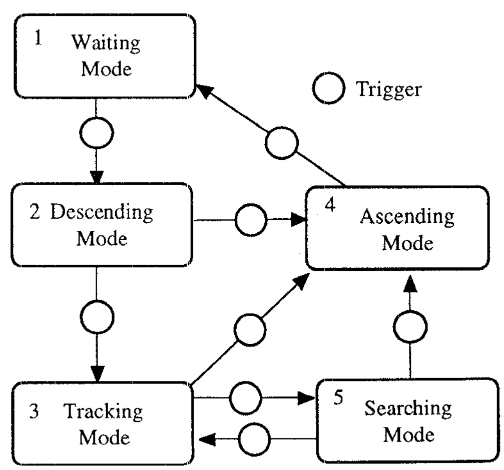


Fig. 5: Transition Diagram of AUCT in the Steering Mode

B. Middle level controller

In the middle level controller, the necessary targets for the control of the Twin-Burger are generated as shown in Fig. 6. Here, the data from the CCD camera are processed and the results with other sensing data are sent to the high and lower level controllers to activate their modes of operations.

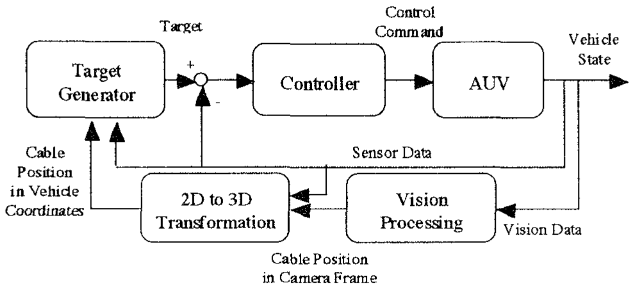


Fig. 6: Schematic Diagram of the Control Architecture of the AUCT system

C. Low level controller

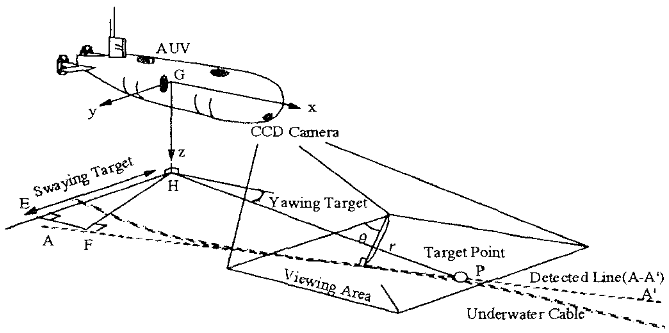


Fig. 7: Generation of the Swaying and Yawing Targets

The line AA' shown in Fig. 7 indicates the 3D line detected in section II. The far most point P on the cable seen in the image is selected as the point for control. The controller of the AUV looks at the point P and steers the vehicle so as to reach P. Steering is done as surging, swaying and yawing. During tracking operation the surging target is kept constant while the other two targets are generated as shown in Fig. 7. The yawing target angle is selected as the angle between the x direction of the robot and the line HP. The sway target is the distance HE along the y direction. HE is the projection of the perpendicular HF dropped to the line AA'. At lake Biwa mission the low level control was done using the Passive Switched Action (PSA) controller.

IV. Results

The Twin-Burger 2 as shown in Fig. 8 was used for the lake Biwa mission. The vehicle consists of three pressure hulls, two on the top and one below. Main computer system placed in the top hulls consist of fourteen T-800 transputers. A T-800 based INMOS B429 image processing board and two CCD cameras are used for vision processing. The CCD cameras are mounted

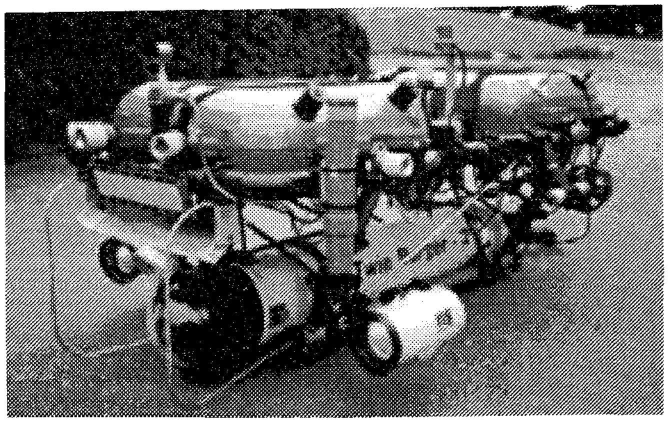


Fig. 8: Twin-Burger 2

in the front of the vehicle. As mentioned in section II, the image area to be processed is selected in each frame using an estimator and the speed of processing is achieved at 10frames/sec.. Five 40W thrusters are placed in such a way that two thrusters for forward movement, two for vertical movement and one for side way movement. A 50kHz acoustic communication link is available to start/stop the mission.

A yellow colored hose pipe was laid along the lake bed as shown in Fig. 9.

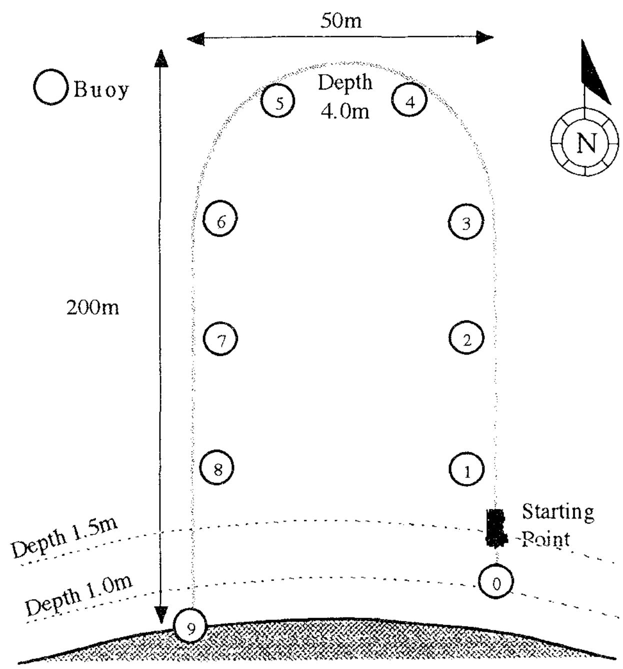
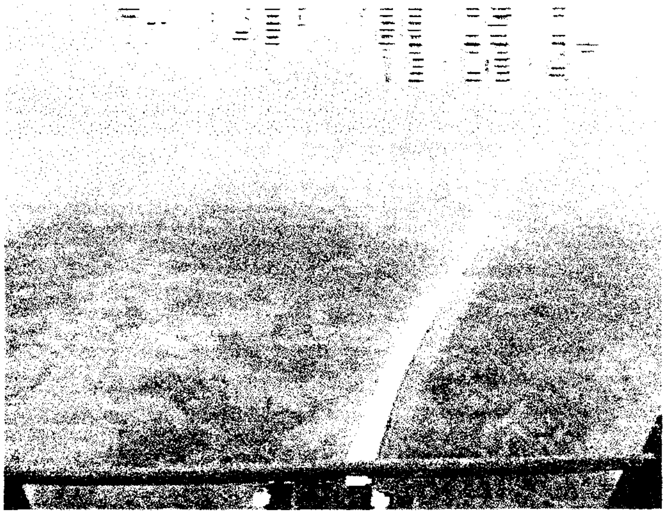


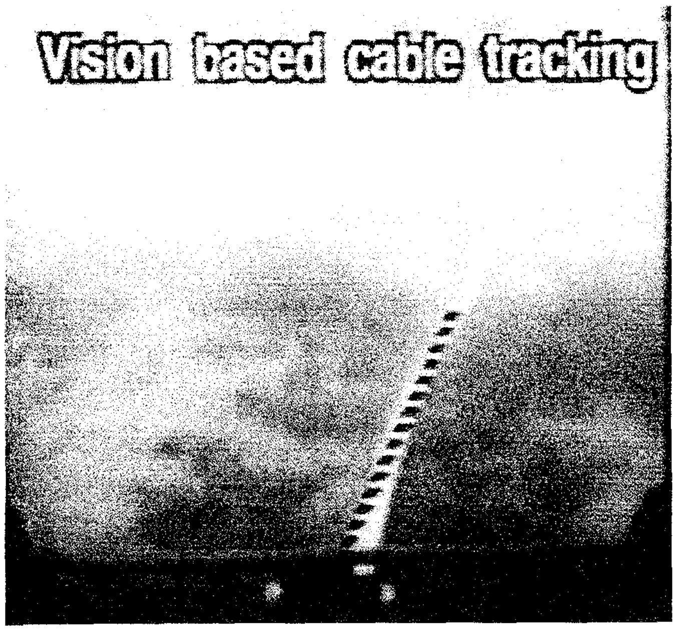
Fig. 9: Underwater Cable Setting

During the tracking operation the altitude target was 1.25m based on the acoustic height sensor. The surge velocity target was at 0.2m/s during the tracking operation and 0m/s for searching. Surging feedback was achieved by the forward direction flow sensor. Heading sensor and visual data were used in the feedback loop of the yawing controller while only visual data was used in the feedback of swaying controller. The sampling time of the controller was 10Hz.

The image of the underwater cable is shown in Fig. 10(a) and Fig. 10(b) shows the the extraction of the cable by extracting the direction and the position of the maximum pixel concentration in the image by Hough transformation.



(a) Original Image



(b) Position of Cable Imposed

Fig. 10: Extraction of the Cable Location in the Image by Hough Transformation

The information of the cable in the image and the operation modes of the Twin-Burger during the AUCT operation at lake Biwa is shown in Fig. 11. The target location in the image to which the vehicle has to be steered is r = 96 and q = 180. It can be seen from the results given in Fig. 11, that from 200sec to 520sec there has been a continuous tracking operation of the cable and in else where it had switched into the search mode very often due to the fact that the cable was covered by waterweeds and was not visible to the Twin-Burger. As a result the targets are

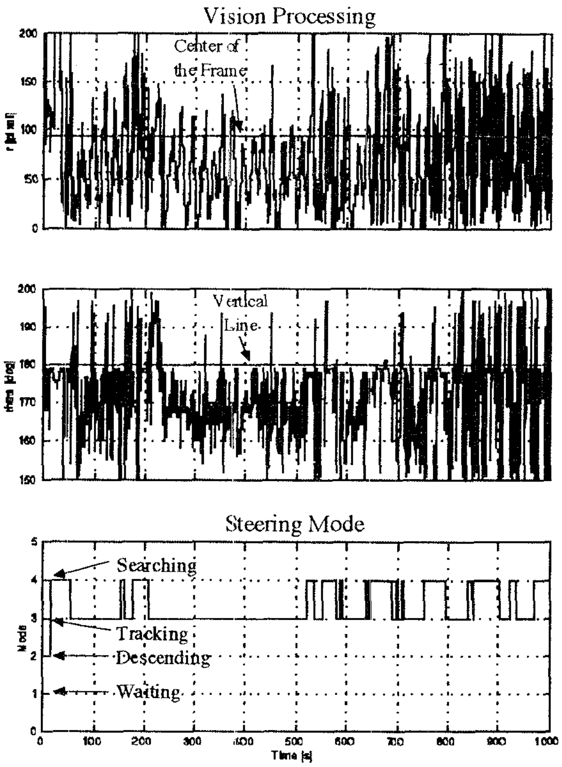


Fig. 11: Vision Processing and the Operation Modes

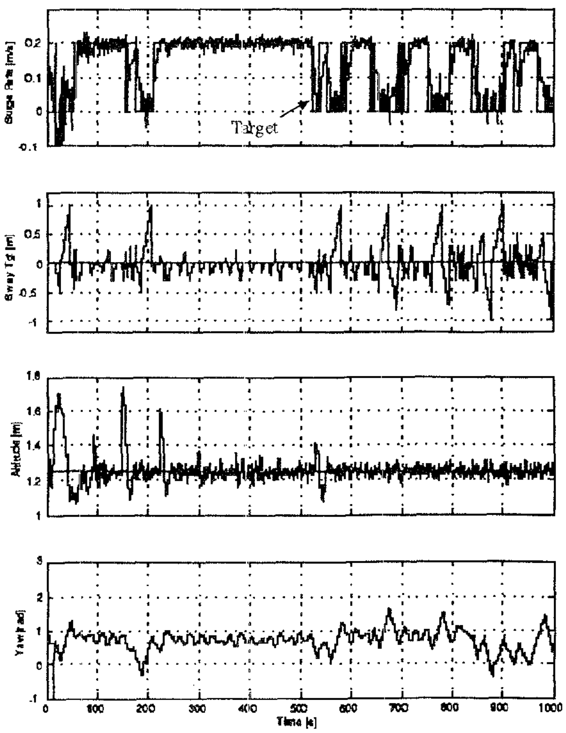


Fig. 12: The Control Targets Generated by the Middle Level Controller

generated as shown in Fig. 12. During the waiting and descending modes the cable was not clearly visible to the Twin-Burger and therefore the image processing data was not used to generate the targets. From the control targets shown in Fig. 12, it can be seen that during waiting and descending modes the sway and surge targets are at zero and only the altitude and yaw targets are generated. The smooth tracking operation was achieved for a distance of 60m. Once the Twin-Burger covered 200m the cable was totally invisible and Twin-Burger ascended to the surface.

V. Conclusions

This paper discussed the navigation of an AUV with the use of visual data fused with other sensors. The performance of the proposed system is demonstrated by implementing Twin-Burger 2 to navigate autonomously at the lake Biwa by following an underwater cable. This mission demonstrates the importance of visual data in the process of navigation of AUVs. The trade off between the processing delay to information is managed by fusing other sensing modalities. Indicating the desired path by the use of an underwater cable, solves the problem of underwater positioning. Therefore it can be concluded that the proposed system is suitable for carrying out routine maintenance operations in predetermined areas.

Acknowledgement

The authors would like to thank the Biwako Research Institute for their support given during the lake Biwa cable tracking experiments.

References

[1] 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, December 1996, pp. 663-668.

[2] R. Blidberg, "Autonomous underwater vehicles, Current activities and research opportunities", Int. Rob. Aut. Sys., vol. 7., no. 2-3, pp. 139-150, August 1991.

[3] T. Fujii, T. Ura, Y. Kuroda, H. Chiba, Y. Nose and K. Aramaki: "Development of a Versatile Test-Bed "Twin-Burger" toward Realization of intelligent Behaviours of Autonomous Underwater Vehicles", Proc. of Oceans'93, Victoria, Canada, September 1993, pp. 1.186-1.192.

[4] A. Huertas and G. Medioni, "Detection of Intensity Changes with Subpixel Accuracy Using Laplacian-Gaussian Masks", IEEE Trans. on PAMI, vol. 8, no. 5, September 1986, pp. 651-664.

[5] Jules S. Jaffe, "Computer Modeling and the Design of Optimal Underwater Imaging Systems", IEEE J. of Oceanic Eng., vol. 15, no. 2, April 1990, pp. 101-111.

[6] N.P. Papanikolopoulos, P.K. Khosla, "Adaptive Robotic Visual Tracking: Theory and Experiments", IEEE Trans. on Automatic Control, vol. 38, no. 3, pp. 429-444, March 1993.