Proceedings of the 2001 IEEE
International Conference on Robotics & Automation
Seoul, Korea · May 21-26, 2001

A 2-D visual servoing for underwater vehicle station keeping

J.-F. Lots, D. M. Lane, E. Trucco *
F. Chaumette †

Abstract

This paper introduces a 2-D visual servoing technique for the station keeping of an unmanned underwater vehicle (UUV) with respect to planar targets on the sea bed. The underwater vehicle is subject to sea current disturbances which make it drift from its desired position. Feature points from unmarked objects are extracted and tracked with a sparse feature tracker developed in-house [1]. The scene depth is estimated from a planar homography. To validate our approach, we emulate the dynamics of the surge and the sway degrees-of-freedom (d.o.f.) of an UUV model with a planar Cartesian robot in our water test tank. Successful station keeping experiments obtained with a range of sea current disturbances are presented.

1 Introduction

So far, in the underwater robotics field few attempts have been made to use vision sensors for control [2, 3, 4, 5, 6]. However, vision shows some interesting features compared to classical positioning sensors to perform station keeping tasks. For example, magnetic compasses suffer from a slow update rate and cannot be used in the vicinity of man-made metallic structures. More importantly, with the exception of depth sensors which are both accurate and fast, on-board translational motion sensors (for surge and sway) are integrating sensors (i.e. accelerometers, Doppler velocity logs) hence subject to drift, and therefore unsuitable for station keeping. A standard camera, however, is not subject to magnetic influences, and has an update rate of 25 Hz. It can also be used as a local absolute positioning sensor: indeed, comparing a current image to the first one yields a position measurement whose drift can be controlled.

Despite its short range (typically 3–10 metres) and the need for heavy computing power, visual control (also called visual servoing) allows very diverse tasks such as for example station keeping [2, 4] or pipe-following [3] to be carried out. Indeed, in [3], Rives and Borelly used the task function approach [7] to perform pipe-following with the Ifremer Vortex ROV in their swimming pool. Marks *et al.* [2] used a stereo camera and Laplacian of Gaussian filtering with dedicated hardware to solve the station keeping problem. Negahdaripour *et al.* [4] recovered 3-D motion from an optical flow based approach.

In a previous paper [6], we tackled the station keeping problem with a modification of the 2 1/2 D visual servoing method from Malis and Chaumette [8]. We proved in simulation its adequacy to underwater vehicle station keeping for its 6 d.o.f. of which 4 were controllable. We also showed preliminary experimental results on a planar Cartesian robot emulating the dynamics of the horizontal d.o.f. of an underwater ROV (ANGUS 003). The hydrodynamic and propulsion forces of this vehicle were experimentally identified with a planar motion mechanism in a towing tank of the Admiralty Marine Technology Establishment, where small-scaled models of ships and submarines are normally tested [9]. In this paper, we set ourselves to control an underwater vehicle in a horizontal plane; the complexity of our previous approach was therefore not necessary. Indeed, restricted to two degrees of freedom, this approach is equivalent to a 2-D visual servoing technique with depth estimation from planar homographies. We applied the proposed 2-D visual servoing technique to the aforementioned Cartesian robot in our water tank, emulating the dynamic behaviour of the surge and sway axis of ANGUS 003. Successful station keeping experiments were carried under various sea current disturbances.

The contributions of this paper lie (1) in the application of visual servoing to underwater vehicle station keeping, and (2) in the use of unmarked natural targets. The latter was made possible by the use of a sparse feature tracker developed in-house [1].

2 Theory

2.1 Underwater vehicle dynamic model

ANGUS 003 is a work-class ROV previously built and characterised in-house. Although no longer used for experiments, a 6 d.o.f nonlinear dynamic model has been iden-

*J.-F. Lots, D. M. Lane, and E. Trucco are with the Ocean Systems Laboratory, Department of Computing & Electrical Engineering Heriot-Watt University, Edinburgh, EH14-4AS, UK. http://www.cee.hw.ac.uk/oceans. Email:{ceejpl1, dml, mtc}@cee.hw.ac.uk

†F. Chaumette is with IRISA/INRIA, Campus de Beaulieu, 35042 Rennes, France. Email: Francois.Chaumette@irisa.fr

0-7803-6475-9/01/$10.00© 2001 IEEE

tified in a test tank with a planar motion mechanism [9]. For the purpose of this paper, we concerned ourselves with the dynamics of its surge and sway axis. Let the body-fixed vehicle's velocity vector be defined as $\nu = [u, v]^T$, and the Earth-fixed velocity vector as $\dot{\eta} = [\dot{x}, \dot{y}]^T$. We assume that the vehicle's heading angle ψ can remain constant and close to zero (using an independent control loop for example [6]). In addition, since ANGUS was designed to be stable in roll (θ) and pitch (ϕ), these angles remain also very small. As a result, the 2×2 Jacobian matrix relating the body-fixed velocity vector ν to the Earth-fixed velocity vector $\dot{\eta}$ can be assumed to be the 2×2 identity matrix. In the following, we will therefore no longer make the distinction between Earth-fixed position and velocity, and body-fixed ones. The motion equations for the ROV can thus be written as:

$$\begin{aligned} M_{11}\dot{u} &= B_{11}(u-u_c)(|u-v_c|+D_u)+a_1\beta+a_2|\beta| \\ M_{22}\dot{v} &= B_{22}(v-v_c)(|v-v_c|+D_v)+a_3\gamma \end{aligned}$$

where M_{ii} ($i = 1, 2$) are the mass matrix coefficients, the B_{ii} ($i = 1, 2$), D_u and D_v the hydrodynamic drag coefficients, u_c and v_c are the velocity of the sea current expressed in the body-fixed frame. The normalised control input of the two back thrusters is $\beta \in [-1, 1]$, $\gamma \in [-1, 1]$ is the normalised control input of the sway thrusters, and a_i ($i = 1, 2, 3$) are the thrusters' efficiency coefficients. Note that the back thrusters are less efficient when operated in reverse. The numerical values of the parameters are gathered in Table 1.

To assess the dynamic performances of the proposed visual servoing technique, we compared it to a "ground truth" PID controller assuming perfect position and velocity measurements at the same sampling rate as the visual servoing experiments, i.e. $T_s = 200$ ms. We have also included the time delay of one sampling period caused by the visual processing. The PID control law is:

$$\begin{bmatrix}\beta \\ \gamma\end{bmatrix} = \mathbf{K}_p \begin{bmatrix}x \\ y\end{bmatrix} + \mathbf{K}_d \begin{bmatrix}\dot{x} \\ \dot{y}\end{bmatrix} + \mathbf{K}_i \begin{bmatrix}\int_0^t x dt \\ \int_0^t y dt\end{bmatrix} \quad (1)$$

where $\mathbf{K}_p = diag(1.0, 4.0)$, $\mathbf{K}_d = diag(0.4, 0.4)$, and $\mathbf{K}_i = diag(0.01, 0.02)$.

Since surge and sway are decoupled, we tuned the parameters of each axis independently. We tried to obtain a time response similar to a second order system critically damped. In Figure 1, one can note the slow dynamic response of the ROV subject to a sea current velocity step input disturbance of $(u_c, v_c) = (-0.2, -0.5)$ m/s. The settling times, corresponding to a positioning error of less than one centimeter, of the surge and sway axis are $t_{surge} = 80$ s and $t_{sway} = 95$ s. Figure 2 shows the corresponding thrusters'

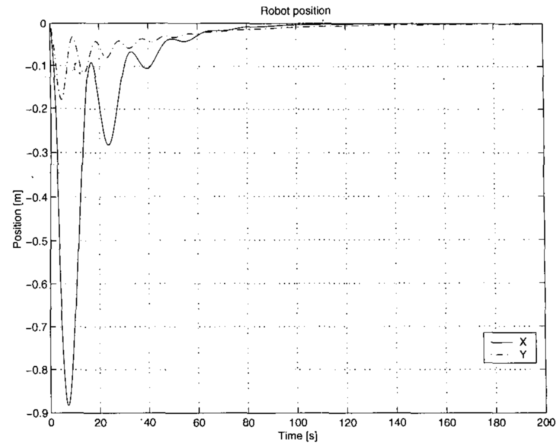

Figure 1: Positioning error of the 2 d.o.f. model of ANGUS subject to a sea current disturbance $(u_c, v_c) = (-0.2, -0.5)$ m/s with a PID controller.

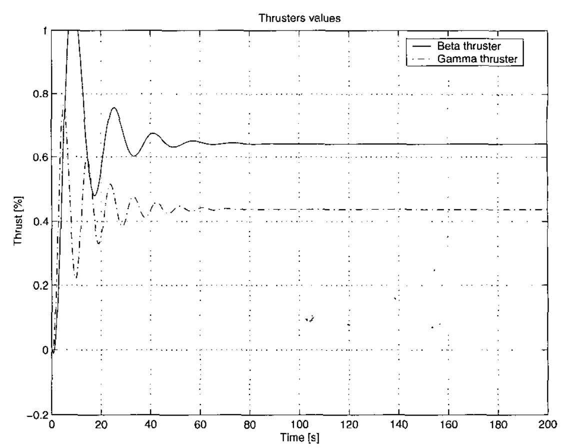

Figure 2: Thrusters values of the 2 d.o.f. model of ANGUS subject to a sea current disturbance $(u_c, v_c) = (-0.2, -0.5)$ m/s with a PID controller.

values. These numbers will serve as a comparison basis for the visual servoing experiments.

In the following sections, we present an image-based visual servoing approach to solve the station keeping problem. It is a simple 2-D visual servoing scheme [10] using the centre of gravity of the set of extracted features in the images as the visual feature. Since the camera is rigidly mounted onto the robot, the frame transformation between the robot frame and the camera frame is therefore constant. To simplify the expression of the following equations, we assumed that the camera reference frame and the robot frame coincided. In other words, controlling the robot is equivalent to controlling the camera whose motion dynamics are those of an underwater vehicle.

M_{11}	M_{22}	B_{11}	B_{22}	D_u
1800	2200	-350	-680	0.6
D_v	a_1	a_2	a_3	
0.6	500	200	250	

Table 1: Dynamic model parameters values

2.2 Depth estimation with a homography

Two views of a plane Π, taken from a single projective camera are related by a *projective homography* ([11]). A homography can be represented by a 3×3 matrix defined up to a scale factor. Given four correspondent points belonging to Π (provided that no three of them are collinear), it is possible to estimate the homography $\mathbf{H}$ by solving a linear system. For more than four points, a least-squares solution can be computed. An interesting property of the homography matrix is that it can be decomposed into a rotation matrix $\mathbf{R}$ and the product of the translation vector $\mathbf{t}/d^*$ between the two views and the vector normal to the plane $\mathbf{n}^*$ (see eq. 2). d^* is the distance from the camera to the plane in the desired position,

$$\mathbf{H} = \mathbf{R} + \frac{\mathbf{t}\,\mathbf{n}^{*T}}{d^*} \tag{2}$$

From the matrix $\mathbf{R}$, we can compute the rotation angle Θ and its vector $\mathbf{u}$. In our case, the normal to the target plane being parallel to the optical axis of the camera, the ratio of the Z-coordinates of a 3-D point expressed in the current frame (Z) and in the desired frame (Z^*) is $r = Z/Z^* = det(\mathbf{H})$ (for a more general description, see [8]). Therefore, the estimated depth of a 3-D point is given by:

$$Z = Z^*\, det(\mathbf{H}). \tag{3}$$

The estimated depth given by eq. (3) will be used in the design of the visual control law of the following section.

2.3 2-D visual servoing

A visual servoing positioning task can be expressed as the regulation to zero of a task function $\mathbf{e}(\mathbf{s}, t)$, where s is a visual feature (in our case, points in the image), and t is the time [7]. If $\mathbf{T}_c$ is the camera velocity screw (6-vector), and $\mathbf{L}(\mathbf{s}, t)$ is an appropriate matrix, called *interaction* matrix, then the camera velocity screw is related to the task function $\mathbf{e}$ by:

$$\dot{\mathbf{e}} = \mathbf{L}\,\mathbf{T}_c \tag{4}$$

In the case of 2-D visual servoing, if we consider a feature point $\mathbf{s}_i = [x_i, y_i]^T$ expressed in the image (Z_i being the depth of the corresponding 3-D point projected on the image point $\mathbf{s}_i$), $\mathbf{L}$ is made up by stacking the following lines (two lines for each feature point):

$$\begin{bmatrix} -1/Z_i & 0 & x_i/Z_i & x_i y_i & -(1+x_i^2) & y_i \\ 0 & -1/Z_i & y_i/Z_i & 1+y_i^2 & -x_i y_i & -x_i \end{bmatrix} \tag{5}$$

An exponential decrease of the task function is obtained by imposing $\dot{\mathbf{e}} = -\lambda\,\mathbf{e}$ (where λ is a positive scalar) so that the corresponding control law would be

$$\mathbf{T}_c = -\lambda \mathbf{L}^{-1}\mathbf{e} \tag{6}$$

Now, since we want to constrain the two translational d.o.f. of the camera parallel to the image plane, the information of one feature point is enough. For robustness purposes, we selected the centre of gravity $[x_g, y_g]^T$ of the set of extracted features. In that case, the interaction matrix has the simple form:

$$\mathbf{L} = \begin{bmatrix} -1/Z_g & 0 \\ 0 & -1/Z_g \end{bmatrix} \tag{7}$$

If we note $\mathbf{s}_g^* = [x_g^*, y_g^*]^T$ the position of the centre of gravity of the features set in the initial (*desired*) position, and $\mathbf{s}_g = [x_g, y_g]^T$ the *current* one, we can define $\mathbf{e}$ as $\mathbf{e} = \mathbf{s}_g - \mathbf{s}_g^*$ so that when $\mathbf{e}$ is zero, the camera is in its desired position. A proportional control law which regulates the camera at its desired position reads as eq. (6). To be able to reject constant sea current disturbances, since the closed loop system is of type 0, we needed an integration term in the control law. We chose to implement a vectorial PID controller to obtain a better stability than a simple PI controller. The proposed control law is then:

$$\begin{bmatrix} \beta \\ \gamma \end{bmatrix} = \begin{bmatrix} -Z_g & 0 \\ 0 & -Z_g \end{bmatrix} (\mathbf{K}_p\,\mathbf{e} + \mathbf{K}_d\,\dot{\mathbf{e}} + \mathbf{K}_i \int_0^t \mathbf{e}\,dt) \tag{8}$$

where $\mathbf{K}_p$, $\mathbf{K}_d$, and $\mathbf{K}_i$ are 2 diagonal matrices with the proportional, derivative and integration control gains of the PID, and Z_g is the depth of the centre of gravity in the current image, estimated from eq. (3) (see section 2.2).

3 Results

3.1 Experimental setup

To experimentally validate our method, we used a planar Cartesian robot in our $4\,\mathrm{m} \times 3\,\mathrm{m} \times 2\,\mathrm{m}$ test tank. This robot was servoed by a DMC-1360 motion control card from Galil Motion Control Inc. hosted on a VME crate. The VME crate was composed of a Unix Host (Motorola MVME-167 board) and three real-time target boards (Motorola MVME-162-22) running Motorola PSOS real-time operating system. We used one of these MVME-162-22 boards to drive the motion control card and ran the dynamic model of ANGUS.

An underwater black and white camera was rigidly mounted on the Cartesian robot. The camera was pointing downward about one metre from the bottom of the tank, imaging a surface of about one metre squared. Its analogue signal was digitised by a BrookTree Bt254 frame grabber hosted on a AD164R parallel processing board from Alpha Data Ltd. The latter was itself hosted by a Linux PC communicating via TCP/IP to the MVME-162-22 board. The visual servoing algorithms (feature extraction and tracking included) ran on the AD164R; the new thrusters values were sent via TCP/IP to the real-time target board. With this setup, the servoing ran at 5 Hz.

3.2 Visual station keeping experiments

Each station keeping experiments followed the same procedure. The robot was initially immobile, and the desired set of features was extracted from the first image taken from the initial position. We then applied a constant sea current disturbance (step input) of velocity (u_c, v_c). As the robot started drifting away from its desired position, the set of features was tracked in the current image. The depth was then estimated from the homography between the initial set of features and the current set of features as described in section 2.2. Similarly, the task function $\mathbf{e}$ was calculated with the centre of gravity of the current feature set. If a feature was lost during tracking, its corresponding feature in the initial set was removed. As long as at least four features remained tracked, the servoing could continue (four points, of which no three are collinear, are needed for the estimation of the homography).

For the whole set of experiments, the control gains of the PID remained constant with the following values: $\mathbf{K}_p = diag(1.0, 4.0)$, $\mathbf{K}_d = diag(2.0, 2.0)$, and $\mathbf{K}_i = diag(0.01, 0.01)$. The estimated distance to the target Z^* was roughly measured and set to 1.1 m.

The final positioning errors of the robot are collected in table 2. In each case, the station keeping was successfully performed to within a good positioning accuracy (the maximum error observed was of a few centimeters).

Sea current velocity (u_c, v_c) [m/s]	Position errors ΔX [mm]	ΔY [mm]
$(-0.1, -0.1)$	$+0.95$	$+5.60$
$(-0.2, -0.1)$	$+0.00$	$+8.80$
$(-0.3, -0.1)$	-0.70	$+7.80$
$(-0.4, -0.1)$	-1.20	$+6.55$
$(-0.5, -0.1)$	-10.30	$+3.65$
$(-0.1, -0.2)$	$+2.95$	-2.95
$(-0.2, -0.2)$	-8.65	$+0.30$
$(-0.3, -0.2)$	-1.70	-1.25
$(-0.4, -0.2)$	-4.90	-2.25
$(-0.5, -0.2)$	-33.65	-14.55

Table 2: Final positioning errors of the robot subject to a range of sea current disturbances.

Figure 3 shows the positioning error of the robot during a station keeping experiment with sea current velocity $(u_c, v_c) = (-0.5, -0.2)\,\mathrm{m/s}$. The robot successfully stabilizes itself within 1 cm accuracy with a settling time of 130 s in surge and 290 s for sway. These times are greater than for the ideal controller of section 2.1. This is due to the added dynamics of the Cartesian robot onto ANGUS', which is more pronounced at low speeds. However, the trajectory followed by the robot is similar to those obtained with our ideal controller.

The thrusters' values time history is shown in figure 4. Here again, we note a behaviour similar to the one of the ideal controller. Noise on the thrusters values is caused by mechanical vibrations of the robot metallic frame, as well as noise on the feature tracking caused by illumination changes.

Lastly, a snapshot of the initial underwater image of this experiment, as well as the final image are shown in figure 5. The trajectory of the tracked features are the white tracks. For display purposes, the images have a coarser resolution than the original one (which was 512×512 with 8 bits), and the aspect ratio is not respected.

4 Discussion

These results demonstrate clearly that visual station keeping can be performed successfully using our approach. These tests showed that the servoing is successful for a range of sea currents magnitudes and directions ($u_c = -0.5\,m/s$

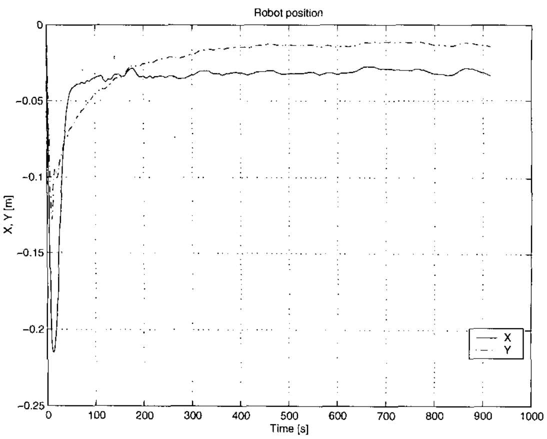

Figure 3: Visual station keeping experiment with sea current $(u_c, v_c) = (-0.5, -0.1)\,\mathrm{m/s}$: Cartesian robot position.

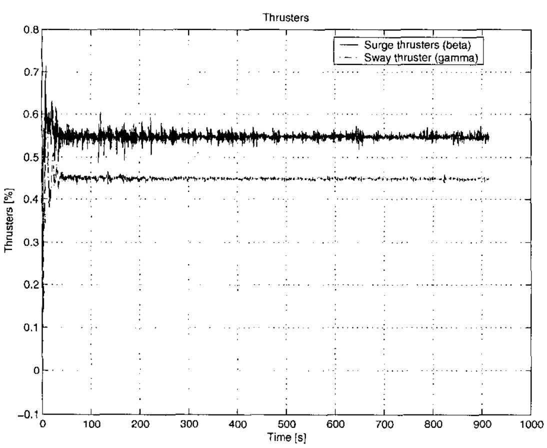

Figure 4: Visual station keeping experiment with sea current $(u_c, v_c) = (-0.5, -0.1)\,\mathrm{m/s}$: thrusters applied to the robot model.

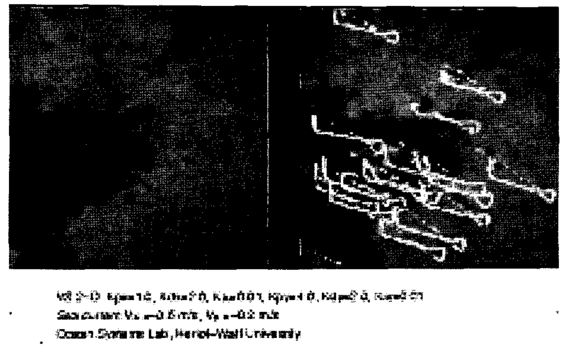

Figure 5: Desired image (left), and final image with pixel tracks (right) obtained with the visual station keeping experiment: sea current $(u_c, v_c) = (-0.5, -0.1)\,\mathrm{m/s}$.

and $v_c = -0.2\,m/s$ being the maximum admissible values).

The limitations in sea current velocity magnitude are due to several factors. One is the limited thrusters' power available for ANGUS which sets a maximum sea current velocity that the robot can counteract. The slow dynamics of the vehicle are also an issue. Indeed, too strong a sea current will cause the robot to drift so far away that no feature from the desired set would remain in the image. A solution to that problem would be to re-extract features during the servoing. However, in that case, the robot would drift unless the initial features are maintained in memory and recognised by data association.

A more serious limitation in our implementation is the 200 ms time delay incurred by the slow TCP/IP communication and the feature tracker running time (150 ms on 512×512 images) which decrease the control stability margins. Tests in simulation without time delay clearly demonstrated this downside effect.

Lastly, some of the performances limitations are caused by the feature tracking. The feature tracker relies on the assumption of small pixel motions in the image. Sampling the video signal at 5 Hz clearly reduces the working conditions. We carried out some tests to assess quantitatively the maximum admissible pixel displacement between two frames and it proved to be 5 pixels. However, 512×512 pixels images represents an important amount of data for tracking purposes, and explain the relatively slow tracking rate. Image sub-sampling would certainly increase the tracking speed to the cost of a reduced positioning accuracy. Abrupt changes of motion direction and changes in lighting, may also cause the tracker to lose features . Further experiments will be needed to assess quantitatively the effect of illumination changes on the tracking behaviour. However, we found that, with sufficient and constant illumination (4 lux), tracked features were successfully positioned with a repeatability of 0.5 pixels standard deviation.

5 Conclusion and Future work

This work showed the validity of 2-D visual servoing for underwater vehicle station keeping. This approach was robust to various sea current disturbances and allows good positioning precisions. The use of a linear PID controller permitted slow, but very stable hover capabilities. Besides, we were able to perform these experiments on a number of unmarked planar targets without any changes in the feature tracking parameters, which demonstrated the robustness of the visual processing part.

Future work will concentrate on transferring this visual servoing algorithm on one of our small ROVs, namely RAUVER built in-house. In addition, more characterisation will be done under various conditions of visibility and lighting since these are major issues underwater. Nonlinear controllers, making use of the vehicle's dynamics would also improve the disturbance rejection, and will be looked into.

6 Acknowledgements

This work was partially funded by EU-MAST project MAS3-CT97-0083 ARAMIS ("Advanced ROV Package for Automatic Investigation of Sediments").

References

[1] E. Trucco, Y. Petillot, I. Tena Ruiz, C. Plakas, and D. M. Lane, "Feature tracking in video and sonar subsea sequences with applications," *Computer Vision and Image Understanding*, vol. 79, no. 1, pp. 92–122, July 2000.

[2] R. L. Marks, H. H. Wang, M. J. Lee, and S. M. Rock, "Automatic visual station keeping of an underwater robot," in *MTS/IEEE OCEANS*, 1994, vol. 2, pp. 137–142.

[3] Patrick Rives and Jean-Jacques Borrelly, "Visual servoing techniques applied to an underwater vehicle," in *IEEE Int. Conference on Robotics and Automation*, Albuquerque, NM, USA, April 1997, vol. 3, pp. 1851–1856.

[4] S. Negahdaripour, X. Xu, and L. Jin, "Direct estimation of motion from sea floor images for automatic station-keeping of submersible platforms," *IEEE Journal of Oceanic Engineering*, vol. 24, no. 3, pp. 370–382, July 1999.

[5] D. M. Lane, E. Trucco, Y. Petillot, I. Tena Ruiz, K. Lebart, J.-F. Lots, and C. Plakas, "Embedded sonar and video processing for AUV applications," in *Proceedings Offshore Technology Conference*, May 2000.

[6] J.-F. Lots, D. M. Lane, and E. Trucco, "Application of 2 1/2 D visual servoing to underwater vehicle station-keeping," in *MTS/IEEE OCEANS*, 2000.

[7] C. Samson, B. Espiau, and M. Le Borgne, *Robot Control: the Task Function Approach*, Oxford University Press, 1990.

[8] Ezio Malis, François Chaumette, and Sylvie Boudet, "2 1/2 D visual servoing," *IEEE Trans. on Robotics and Automation*, vol. 15, no. 2, pp. 238–250, April 1999.

[9] P. Bellec, "Simulation of the six-degrees-of-freedom motion of a remotely controlled unmanned submersible," M.S. thesis, Department of Electrical and Electronic Engineering, Heriot-Watt University, Edinburgh, UK, 1983.

[10] Bernard Espiau, François Chaumette, and Patrick Rives, "A new approach to visual servoing in robotics," *IEEE Transactions on Robotics and Automation*, vol. 8, no. 3, pp. 313–325, 1992.

[11] R. Y. Tsai, T. S. Huang, and W. L. Zhu, "Estimating three-dimensional motion parameters of a rigid planar patch, II: Singular Value Decomposition," *IEEE Trans. On Acoustics, Speech, Signal Processing*, vol. 30, pp. 525–534, August 1982.