ELSEVIER

Vision-based ROV horizontal motion control: Near-seafloor experimental results

M. Caccia*

CNR-ISSIA Sez. di Genova, Via De Marini 6, 16149 Genova, Italy

Abstract

In this paper the problem of high-precision motion control of remotely operated vehicles (ROVs) in the proximity of the seabed through vision-based motion estimation is addressed. The proposed approach consists of the integration of a cheap monocular vision system for the estimate of the vehicle's linear motion with a conventional dual-loop hierarchical architecture for kinematics and dynamics control. Results obtained by operating at sea the Romeo ROV are presented, demonstrating the system capability in performing station-keeping in the presence of external disturbance and relatively high accuracy in horizontal motion control.
© 2006 Elsevier Ltd. All rights reserved.

Keywords: Underwater vision; Motion estimation; Guidance and control; Mobile robots

1. Introduction

The problem of accurate motion control of remotely operated vehicles (ROVs) in the proximity of the seabed is crucial in many service, scientific and archeological applications. A number of control methodologies have been proposed in the literature to handle the high degree of uncertainty which characterizes underwater operations in terms of external disturbance, system dynamics, actuator forces, sensor measurements and environmental structure. Indeed, only few of these techniques are reliable and precise for 3-D position and velocity sensing with an update rate compatible with fast and precise closed-loop feedback control in all degrees of freedom, and "*rare are experimental results for X–Y control of vehicles in the horizontal plane*" (Kinsey & Whitcomb, 2004).

In the following, a simple and cheap solution to this problem is proposed and demonstrated to be reliable at sea, where a conventional dual-loop guidance and control architecture is integrated with a single camera laser-triangulation optical-correlation sensor for the estimate of the horizontal motion of a ROV. The resulting system can satisfactorily perform basic tasks, such as station-keeping or moving back to a user-selected point, that, when executed automatically, are of extraordinary help to the ROV pilot during, for instance, scientific surveys.

In order to handle external disturbances and parametric model uncertainty, robust and adaptive control techniques have been proposed from the pioneer paper by Yoerger and Slotine (1985), introducing the use of sliding-mode control, and the first experimental validation of adaptive sliding control on a tethered underwater vehicle, performed on the RPV, a testbed vehicle for the development of the Jason ROV (Yoerger & Slotine, 1991). Later examples of the use of sliding-mode techniques are, for instance, the depth, altitude, heading and cross-track error controllers of the NPS ARIES AUV (Marco & Healey, 2001), and the heading and depth controllers of the autonomous underwater shuttle SIRENE (Silvestre, Aguiar, Oliveira, & Pascoal, 1998), which proved their capability in handling an uncertainty of the order of 50% in the estimation of the vehicle's hydrodynamic parameters. Adaptive control schemes were proposed in Fossen and Sagatun (1991) and Yuh (1990), and experimentally demonstrated, for instance, with the ODIN AUV in Antonelli, Chiaverini, Sarkar, and West (2001). Satisfactorily performances in horizontal motion control of the Romeo ROV, in the case position measurements were provided by a couple of echo-sounders tracking the walls of a testing pool, have been

*Tel.: +39 106475612; fax: +39 106475600.
E-mail address: max@ian.ge.cnr.it.

obtained adopting a dual-loop hierarchical guidance and control scheme, based on a set of Lyapunov-based guidance task functions and a PI gain scheduling controller able to reduce the robot dynamics to a nominal characteristic equation (Caccia & Veruggio, 2000). State-of-the-art results in underwater robotic vehicles dynamic positioning with an extended experimental comparison of trajectory-tracking controllers for low-speed maneuvering (PD, exactly linearizing and nonlinear controllers, and their adaptive extensions) can be found in Smalwood and Whitcomb (2004).

Accurate and reliable solutions to the problem of estimating the horizontal motion of an ROV are provided by acoustics. The combination of high-frequency long base-line (LBL, 300 kHz), Doppler velocimeter (1.2 MHz), and ring-laser gyro can guarantee accurate motion control on the horizontal plane as in the case of archaeological applications of the JASON ROV (Whitcomb, Yoerger, Singh, & Howland, 1999). Anyway, the very complex logistics in terms of careful placement of transponders and the very limited maximum range of high-frequency LBL suggest the use of Doppler-based navigation systems (Kinsey & Whitcomb, 2003) or cheap, standalone optical vision devices for horizontal motion estimation. In particular, in the last years, the exponential rise in computing performance and the availability of high-resolution digital cameras boosted the research in underwater visual navigation starting from the basic station-keeping techniques, founded on obtaining the robot position by tracking texture features using image filtering and correlation, proposed in Marks, Wang, Lee, and Rock (1994). These results were transferred to the MBARI Ventana ROV (Leabourne, Rock, Fleischer, & Burton, 1997), demonstrating a precision of the order of 10 cm when operating at sea at an altitude from the seabed of 1 m in the hypothesis of no yaw rotations.

In order to handle the induction of optical flow in the scene by the motion of the light sources mounted on the ROV and moving together with the camera(s), a revised definition of optical flow as "*the perceived transformation of brightness patterns in an image sequence*" was introduced in Negahdaripour (1998), and used for direct estimation of linear and yaw motion from seafloor images in Negahdaripour, Xu, and Jin (1999), where accurate station-keeping is demonstrated in experiments with a three-thruster floating vehicle in a water tank. At sea tests with a Phantom XTL ROV pointed out the strong coupling between the constraints on robust motion sensing from images and the vehicle control (Xu & Negahdaripour, 1999). These techniques were improved and integrated in a mosaic-based concurrent mapping and localization scheme in Negahdaripour and Xu (2002), where, in spite of the high degree of robustness of the gradient-based motion estimation and mosaicing methods, the inability of the control system of the Phantom XTL ROV to execute corrective actions promptly for maintaining station was confirmed. Quite interesting results in combined vision-based motion estimation and mosaicing were demonstrated with a Phantom 500SP ROV in Gracias, Van der Zwaan, Bernardino, and Santos-Victor (2003).

In this context, a monocular video device was designed, developed and integrated on the Romeo ROV in order to exploit at sea the dual-loop guidance and control architecture satisfactorily tested in pool (Caccia & Veruggio, 2000). The resulting laser-triangulation optical-correlation sensor, presented in Caccia (2003b), solves the problem of estimating the image depth by directly measuring it through a laser beam triangulation system, locally structuring the environment with a set of laser spots of known orientation with respect to the camera axis (Clark, Wallace, & Pronzato, 1998; Marques, Nunes, & de Almeida, 1998). An example of the use of colour CCD cameras and parallel lasers for estimating the motion of an AUV with respect to man-made structures is reported in Kondo and Ura (2004). The automatic selection of suitable areas of interest allows the tracking of clearly distinguishable parts of the image and, in any case, a measurement of the reliability of the motion estimate. Although preliminary results were presented in Caccia (2005), a complete overview of the integration of the vision-based horizontal motion estimation system and the dual-loop guidance and control architecture is given in this paper, focusing on the discussion of extended experimental results demonstrating satisfactory performance when executing basic tasks to support the ROV pilot during scientific survey missions.

The paper is organized as follows. A general view of the navigation, guidance and control (NGC) architecture is given in Section 2, together with a short description of the Romeo ROV and a view of the optical device, while the adopted nomenclature and vehicle, sensor and environment models are presented in Section 3. Section 4 presents the laser-triangulation optical-correlation sensor, discussing the image processing algorithms and the effects of pitch and roll small oscillations of the ROV surge and sway. Guidance and control algorithms are presented in Section 5, while experimental results are reported in Section 6 pointing out the precision of the vision-based device and the effectiveness of the proposed control algorithms in counteracting current and tether disturbance occurring during typical benthic operations of the Romeo ROV.

2. System architecture

The system consists of the Romeo ROV (Caccia, Bono, Bruzzone, & Veruggio, 2000) integrated with a laser-triangulation optical-correlation sensor (see Fig. 1), which, mounted downward-looking below the vehicle, measures its horizontal position and speed, and range from the sea bottom. The ROV, shown in Fig. 2, was designed with particular attention to the geometry of the propulsion system in order to obtain a vehicle able, at least in line of principle, to maneuver with high precision both in the vertical and horizontal plane minimizing the interactions with the sea floor. Thus, the four vertical propellers were
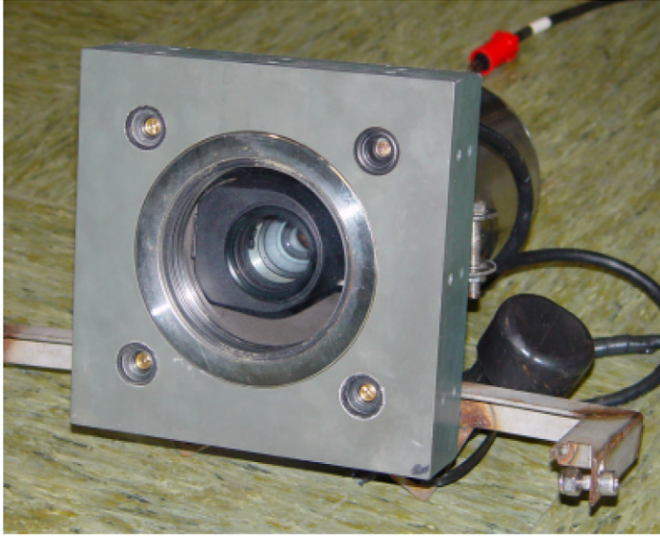
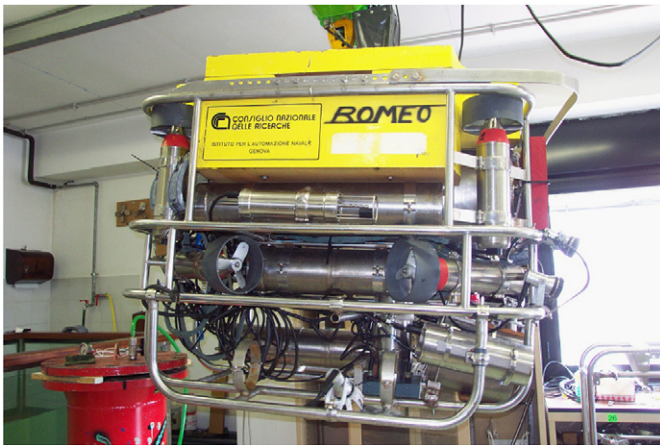Fig. 1. Laser-triangulation optical-correlation sensor.

Fig. 2. Romeo benthic survey configuration.

positioned on the top vehicle corners, and the four horizontal thrusters were aligned to the horizontal diagonals in the mid corners of the frame. The symmetry in thruster location allowed a smooth distribution of the control actions over the actuators, and the redundancy of the actuation system enabled the vehicle to handle faults in the propulsion system without sensibly altering its motion control performances.

The core vehicle, weighting about 350 kg in air, is composed of a frame ($130 \times 90 \times 66$ cm (lwh)), equipped with a number of titanium cylindrical canisters for electronics (100×32 cm ld), batteries (80×15 cm ld), DC/DC converters (80×15 cm ld), and compass, gyro, and inclinometers (60×15 cm ld). The standard toolsled, which measures $130 \times 90 \times 30 =$ cm (lwh), brings additional batteries.

As far as the NGC system is concerned, as shown in Fig. 3, it consists of a dual-loop hierarchical guidance and control architecture of the type presented in Caccia and
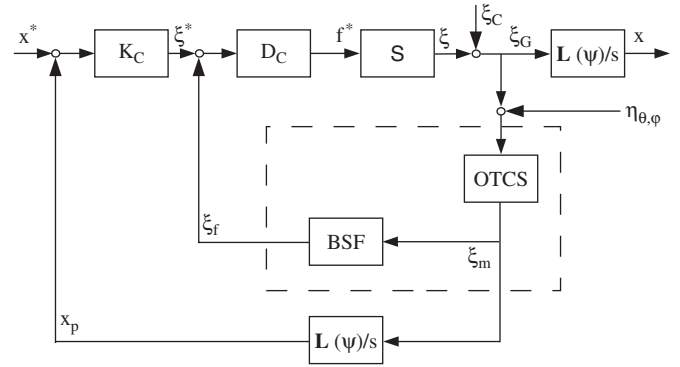

Fig. 3. NGC dual-loop architecture.

Veruggio (2000), constituted by a dynamics controller D_c controlling the vehicle linear speed ξ, i.e. surge and sway, and a kinematics controller K_c handling position control. Since, as usual, see Section 3.1 for details, the vehicle dynamics S with respect to the water is represented in a body-fixed reference frame, neglecting pitch and roll, the robot position in an earth-fixed frame is obtained by integrating the body-fixed linear speed including the contribution of the sea current multiplied by a rotation matrix L, which is a function of the vehicle heading ψ. The linear speed with respect to the ground in a vehicle-fixed frame is measured by a laser-triangulation optical-correlation sensor (OTCS) of the type described in Caccia (2003b). The well-known effect of indistinguishability between small surge and sway displacements and pitch and roll rotations when a monocamera video device for motion estimation is mounted downward-looking below a ROV (see, for instance, Marks, Rock, & Lee, 1995) can be modelled by adding a quasi-sinusoidal disturbance $\eta_{\theta,\varphi}$ and tackled by band-stop filtering (BSF) the measured speed as discussed in Caccia (2003a). The vehicle horizontal position with respect to an earth-fixed frame is, at this stage, simply predicted by integrating the measured surge and sway multiplied by a rotation matrix $L(\psi)$.

3. Modelling and nomenclature

3.1. Vehicle model

As discussed in Fossen (1994), the motion of marine vehicles is usually described with respect to an earth-fixed inertial reference frame $\langle e \rangle$ and a moving body-fixed reference frame $\langle v \rangle$, whose origin coincides with the center of gravity of the vehicle. Thus, position and orientation of the vehicle are described relative to the inertial reference frame, while linear and angular speeds are expressed relative to the body-fixed reference frame.

The vehicle kinematics nomenclature follows (see Fig. 4):
$\underline{x} = [x \ y \ z]^T$: ROV position relative to the earth-fixed reference frame;

$[\varphi \ \theta \ \psi]^T$: ROV roll, pitch and yaw angles relative to the earth-fixed reference frame;
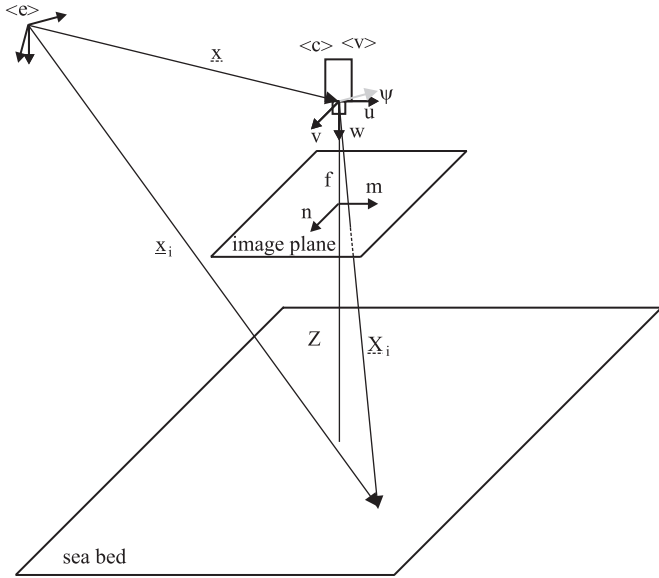
Fig. 4. Camera mounted downward-looking below the ROV: nomenclature.

$[u\ v\ w]^\mathrm{T}$: ROV linear speed (surge, sway, heave) relative to the vehicle-fixed reference frame;

$[p\ q\ r]^\mathrm{T}$: ROV angular speed (roll, pitch and yaw rates) relative to the vehicle-fixed reference frame.

Since the vehicle is a rigid body floating in the water, it is necessary to distinguish between its velocity with respect to the water, i.e. $\underline{\xi} = [u\ v]^\mathrm{T}$ in the horizontal plane, and its ground speed, i.e. $\underline{\xi}_G = \underline{\xi} + \underline{\xi}_C$ including the sea current, both expressed with respect to the vehicle-fixed reference frame.

The vehicle position $\underline{x}$ in the earth-fixed reference frame is related to the vehicle speed $\underline{\xi}_G = [u_G\ v_G]^\mathrm{T}$ with respect to the ground in the body-fixed frame by

$$\underline{\dot{x}} = L(\psi)\underline{\xi}_G, \tag{1}$$

where

$$L(\psi) = \begin{bmatrix} \cos\psi & -\sin\psi \\ \sin\psi & \cos\psi \end{bmatrix}.$$

3.2. Camera-laser sensor model

The developed video system for measuring range from surfaces is shown in Fig. 1. It consists of a video camera and four parallel red laser diodes. A camera-fixed reference frame $\langle c \rangle$ is defined with the z-axis directed towards the scene.

The camera and image basic nomenclature follows (see Fig. 4):

f: focal length;

$[m\ n]^\mathrm{T}$: image point in the image plane;

$[\dot{m}\ \dot{n}]$: image motion field in the image plane;

$\underline{X} = [X\ Y\ Z]$: coordinates of the generic point in the 3-D space (referred to the camera frame).

Point coordinates in the 3-D space and in the image plane are related by the camera perspective model

$$\begin{bmatrix} m \\ n \end{bmatrix} = \frac{f}{Z} \begin{bmatrix} X \\ Y \end{bmatrix}. \tag{2}$$

In the case the camera is mounted downward-looking below the vehicle, the frames $\langle c \rangle$ and $\langle v \rangle$ are assumed to coincide. Denoting with $\underline{x}_i$ the coordinates of the generic ith point in the 3-D space (referred to the earth-fixed frame), then $\underline{X}_i = \underline{x}_i - \underline{x}$ and $\underline{\dot{X}}_i = -\underline{\dot{x}}$.

The four laser diodes are rigidly connected to the camera-fixed frame and laser spots are given by the intersection between the laser rays and the seabed: if the vehicle pitch and roll are zero, the corresponding Z-axes are vertical and the laser spot Z coordinates in the camera frame represent the altitude of the vehicle from the surface.

3.3. Seabed model

At a generic instant, the seabed can be locally represented in the camera (vehicle)-fixed reference frame by the plane π_L

$$X \sin\alpha\cos\gamma + Y \sin\alpha\sin\gamma + (Z - h)\cos\alpha = 0, \tag{3}$$

where $\alpha \in [0, \pi/2]$ and $\gamma \in (-\pi, \pi]$ are the seabed maximum slope and its orientation and h is the vehicle altitude.

Since the small area covered by the image, the seabed could be locally assumed to be horizontal, and Eq. (3) reduces to

$$Z = h = c_\pi. \tag{4}$$

3.4. Rigid-body motion: motion field of a stationary scene point

Considering a vehicle-fixed camera moving at linear and angular speed $[u\ v\ w]^\mathrm{T}$ and $[p\ q\ r]^\mathrm{T}$, respectively, the motion field of a generic 3-D point in the camera frame is (Marks et al., 1995)

$$\begin{bmatrix} \dot{m} \\ \dot{n} \end{bmatrix} = -\frac{f}{Z}\begin{bmatrix} u \\ v \end{bmatrix} + \frac{w}{Z}\begin{bmatrix} m \\ n \end{bmatrix} + r\begin{bmatrix} n \\ -m \end{bmatrix}$$
$$+ f\begin{bmatrix} -q \\ p \end{bmatrix} + \frac{pn - qm}{f}\begin{bmatrix} m \\ n \end{bmatrix}. \tag{5}$$

4. Vision-based motion estimation

4.1. Image processing algorithms

In the following a brief summary of the image processing system and algorithms used for measuring the camera (vehicle) speed is given. For details the reader can refer to
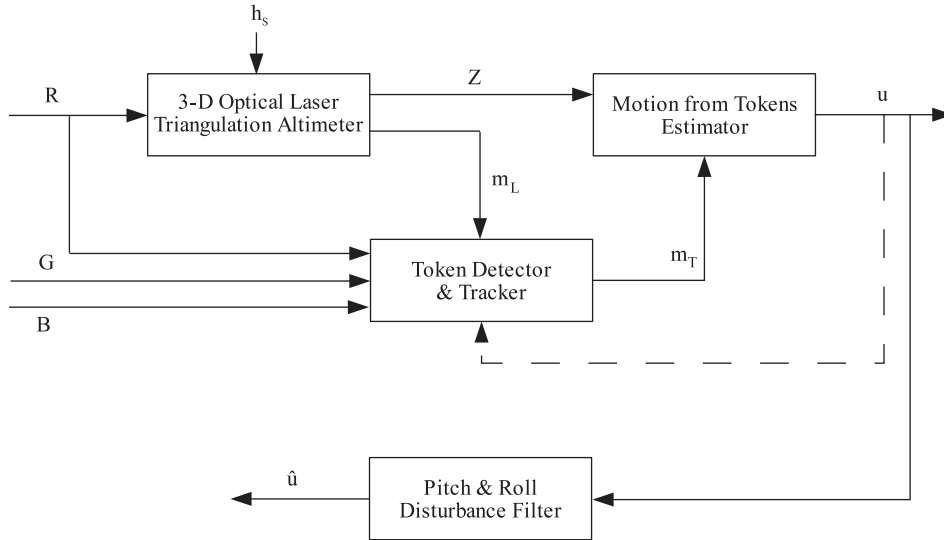
Fig. 5. OTCS architecture.

Caccia (2003b). The system basically consists of three modules (see Fig. 5):

- 3-D optical laser triangulation altimeter: detects and tracks the laser spots in the image coordinates, estimates their spatial coordinates in the camera (vehicle)-fixed frame, and finally computes the seabed range. Since laser diodes emit only a red spot, the R component of the image is processed. The vision-based estimate of the scene depth can be integrated with altitude measurements supplied by acoustic altimeters mounted on the vehicle to increase system reliability and field of work.
- Token detector and tracker: automatically detects and tracks areas of interest in the image, characterized by high-local variance, which can also be seen as a measurement of the observability of an image template. After a 2-D band-pass filtering to enhance specific spatial wavelengths, local variances are computed to evaluate contrast, and high-local variance areas are extracted as templates. As discussed in Misu, Hashimoto, and Ninomiya (1999), in order to reduce computation, band-pass filtering is performed executing averaging as low-pass filtering and sub-sampling and Laplacian filtering as high-pass filtering. Since in the hypothesis of constant heading operations close to the seabed small rotations and inter-frame variations in scene depth occur, template tracking is performed through the computation of the highest correlation displacement in a neighborhood of the previous location, computed according to the estimated motion. Token tracking fails when the correlation gets lower than a suitable threshold.
- Motion from tokens estimator: computes the vehicle motion in the camera-fixed reference frame from token displacements in consecutive images assuming that the image depth is supplied by the 3-D optical

laser triangulation altimeter. In the case yaw motion is considered, neglecting pitch and roll, Eq. (5) reduces to

$$\begin{bmatrix} \dot{m} \\ \dot{n} \end{bmatrix} = -\frac{f}{Z}\begin{bmatrix} u \\ v \end{bmatrix} + \frac{w}{Z}\begin{bmatrix} m \\ n \end{bmatrix} + r\begin{bmatrix} n \\ -m \end{bmatrix}. \qquad (6)$$

In the case the image depth is assumed to be constant (this hypothesis is reasonable given the small area covered by the image), defining the normalized speed $\tilde{u} = u/Z$, $\tilde{v} = v/Z$ and $\tilde{w} = w/Z$, the following over-constrained system can be obtained given N tracked image templates and solved with a LS algorithm:

$$\begin{bmatrix} -f & 0 & m_1 & n_1 \\ 0 & -f & n_1 & -m_1 \\ \vdots & \vdots & \vdots & \vdots \\ -f & 0 & m_N & n_N \\ 0 & -f & n_N & -m_N \end{bmatrix}\begin{bmatrix} \tilde{u} \\ \tilde{v} \\ \tilde{w} \\ r \end{bmatrix} = \begin{bmatrix} \dot{m}_1 \\ \dot{n}_1 \\ \vdots \\ \dot{m}_N \\ \dot{n}_N \end{bmatrix}. \qquad (7)$$

4.2. Pitch and roll induced noise: band-stop filter

Small oscillations in uncontrolled pitch and roll induce quasi-sinusoidal disturbance on the measured surge and sway according to the *first-order* relationship

$$\begin{bmatrix} \dot{m} \\ \dot{n} \end{bmatrix} \approx -\frac{f}{Z}\begin{bmatrix} u \\ v \end{bmatrix} + \frac{w}{Z}\begin{bmatrix} m \\ n \end{bmatrix} + f\begin{bmatrix} -q \\ p \end{bmatrix}. \qquad (8)$$

As discussed in Caccia (2003a), where experimental results are presented, the power spectral density of these oscillations is independent from the maneuvers executed by the ROV, and their amplitude is not negligible at very low speed. Indeed, for typical ROV benthic operations at an altitude of about 1 m, an angular rate of 1 deg/s corresponds to a disturbance of about 1.75 cm/s on the
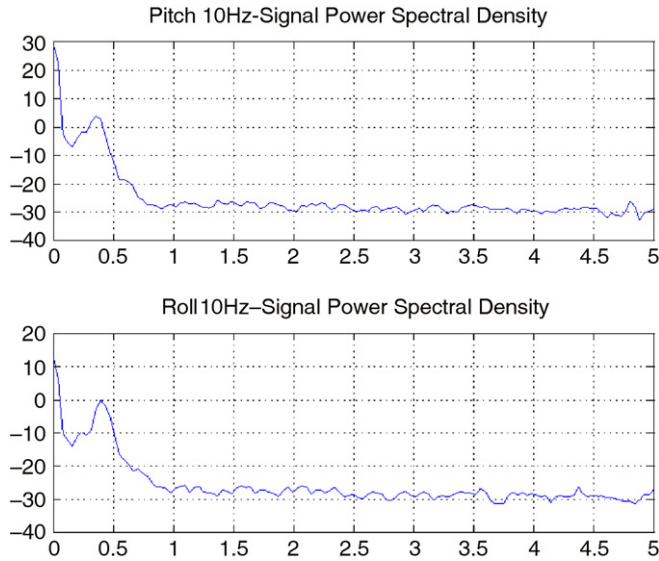
Fig. 6. Power spectral density of pitch and roll measured by the Romeo ROV inclinometers.

estimated linear speed in the case one pixel corresponds to 1 mm at a range of 1 m.

As shown by Fig. 6, where an example of the power spectral density of the measured pitch and roll of the Romeo ROV is plotted, this disturbance can be rejected by suitable BSF introducing some delay. In particular, a couple of Butterworth bandstop digital filters, that lose no more than 1.0 dB in the passband and with at least 5.0 dB of attenuation in the stopband, have been designed. The passband and stopband edge frequencies are [0.15 0.55] Hz and [0.25 0.45] Hz for the surge/pitch filter, and [0.20 0.60] Hz and [0.30 0.50] Hz for the sway/roll filter, respectively.

5. Guidance and control

5.1. Dynamics control (speed control)

Surge and sway controllers are based on the practical 1-DOF uncoupled model of vehicle dynamics (Caccia, Indiveri, & Veruggio, 2000):

$$m_\xi \dot{\xi} = -k_\xi \xi - k_{\xi|\xi|} \xi |\xi| + f_\xi, \tag{9}$$

where ξ, k_ξ, $k_{\xi|\xi|}$, m_ξ and f_ξ represent the linear speed with respect to the water, linear and quadratic drag coefficients, inertia included added mass, and applied force, respectively.

Linearization of Eq. (9) about the operating point $\xi = \xi^*$ and $f = f^*(\xi^*) : \dot{\xi}(\xi^*, f^*) = 0$, i.e. $f^*(\xi^*) = k_\xi \xi^* + k_{\xi|\xi|} \xi^* |\xi^*|$, results in the family of parameterized linear models

$$\dot{\xi}_\delta = -\frac{k_\xi + 2k_{\xi|\xi|}|\xi^*|}{m_\xi} \xi_\delta + \frac{1}{m_\xi} f_\delta, \tag{10}$$

where $\xi_\delta = \xi - \xi^*$ and $f_\delta = f - f^*$.

Thus, according to the gain-scheduling technique presented in Khalil (1996), at each constant operating point ξ^* the controller assumes the form

$$f = f^* + f_\delta, \tag{11}$$

where

$$f_\delta = k_P e + k_I \gamma, \quad \dot{\gamma} = e = \xi - \xi^* = \xi_\delta \tag{12}$$

implements a gain-scheduling PI controller in order to obtain a desired characteristic equation for the closed-loop linearized system of the form

$$s^2 + 2\sigma s + \sigma^2 + \omega_n^2 = 0. \tag{13}$$

After some calculations the gains $k_P = k_\xi + 2k_{\xi|\xi|}|\xi^*| - 2m_\xi \sigma$ and $k_I = -m_\xi(\sigma^2 + \omega_n^2)$ are obtained.

In operating conditions an anti-windup mechanism is implemented such that $|\gamma| \leqslant \eta_{MAX}$.

5.2. Kinematics control (position control)

Defined a hovering task function of PI-type, $\underline{e} = (\underline{x} - \underline{x}^*) + \mu \int_0^t (\underline{x} - \underline{x}^*) d\tau$, the kinematics controller assumes the form

$$\xi^* = -g_P L^{-1}(\underline{x} - \underline{x}^*) - g_I L^{-1} \int_0^t (\underline{x} - \underline{x}^*) d\tau \tag{14}$$

with $g_P = \lambda + \mu$ and $g_I = \lambda\mu$, $\lambda > 0$ and $\mu \geqslant 0$.

In order to minimize wind-up effects, the integrator is enabled/disabled with an hysteresis mechanism when the range from the target $r = \sqrt{(\underline{x} - \underline{x}^*)^T (\underline{x} - \underline{x}^*)}$ gets lower/higher than I_e^{ON}/I_e^{OFF}, respectively. In addition, the proportional and integral control actions are saturated so that $|g_P L^{-1}(\underline{x} - \underline{x}^*)| \leqslant \xi_P^{MAX}$ and $|g_I L^{-1} \int_0^t (\underline{x} - \underline{x}^*) d\tau| \leqslant \xi_I^{MAX}$, respectively.

6. Experimental results

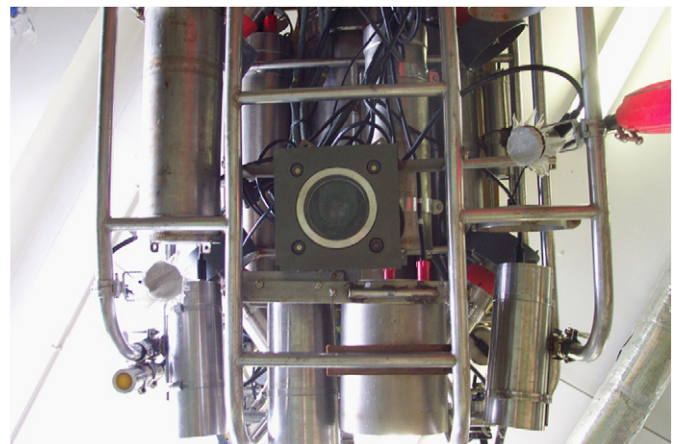Experiments have been performed with the Romeo ROV equipped with the laser-triangulation optical-correlation

Fig. 7. Laser-triangulation optical-correlation sensor mounted below the ROV.
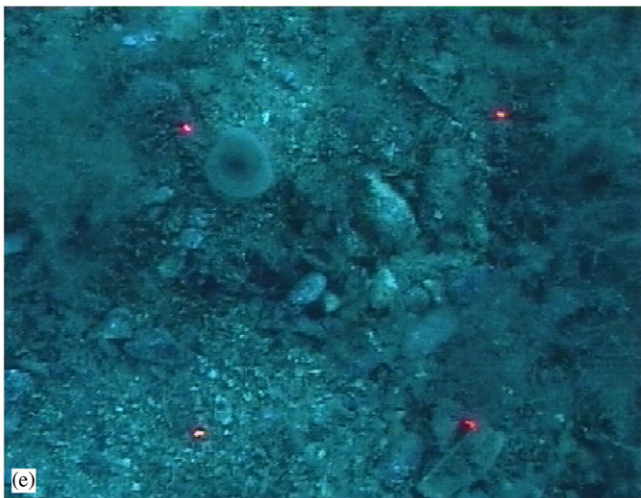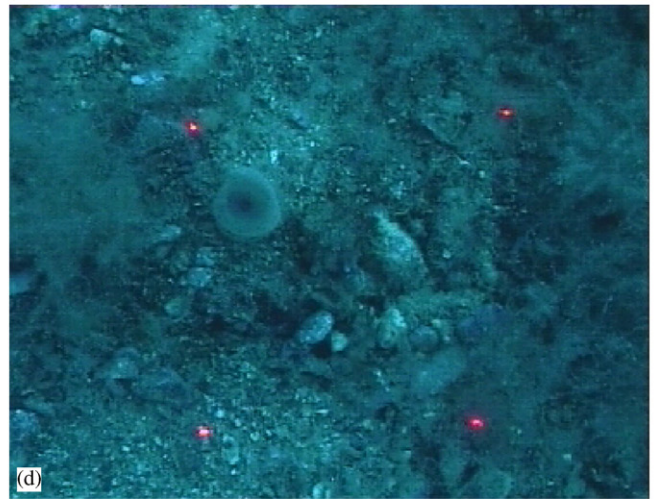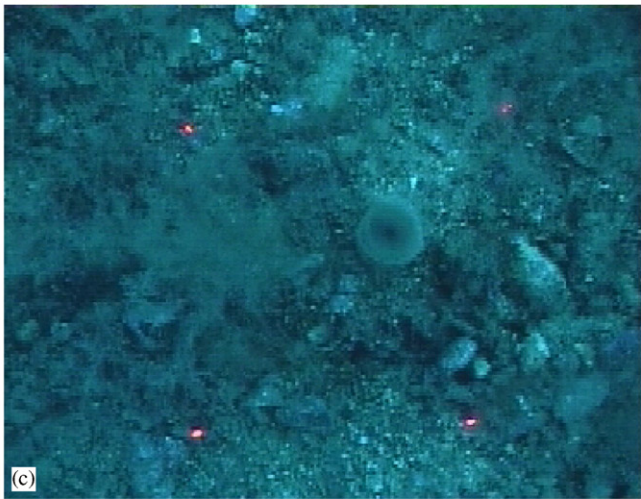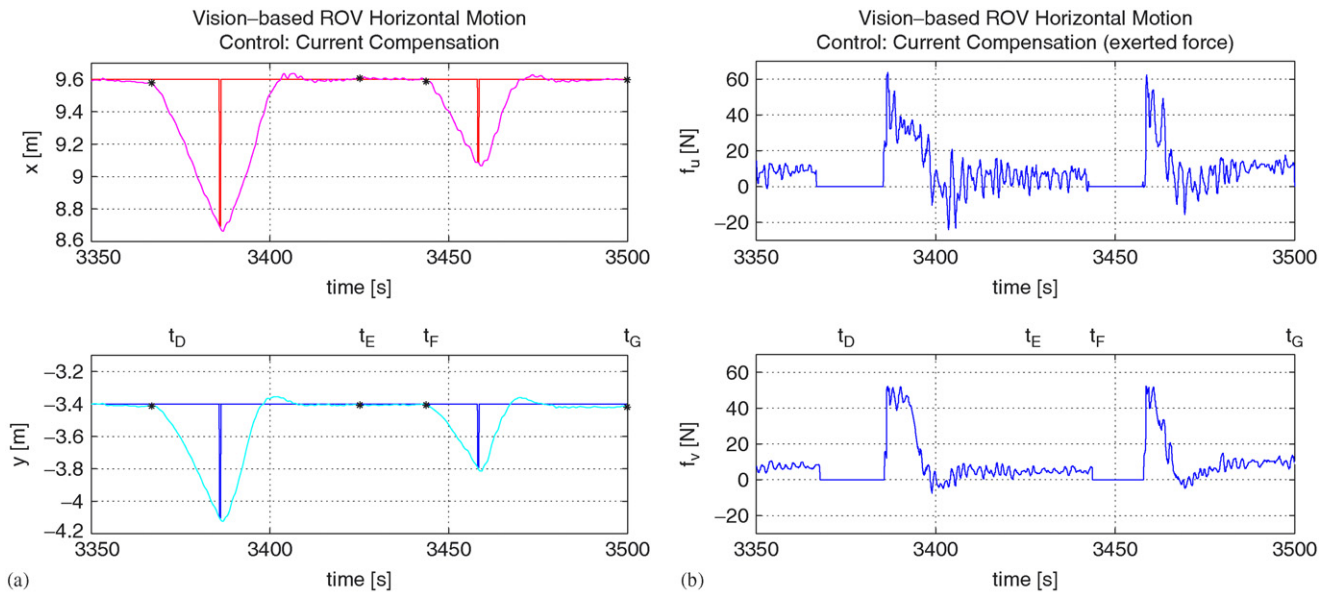
Fig. 8. ROV reference and estimated *x–y* coordinates and exerted control surge and sway forces: ROV camera views at time t_D, t_E, t_F and t_G are shown clockwise from the mid line: (a) Reference and estimated *x–y*; (b) Surge and sway force.

Table 1
Speed integral vs. camera displacement from couple of images: external disturbance compensation test

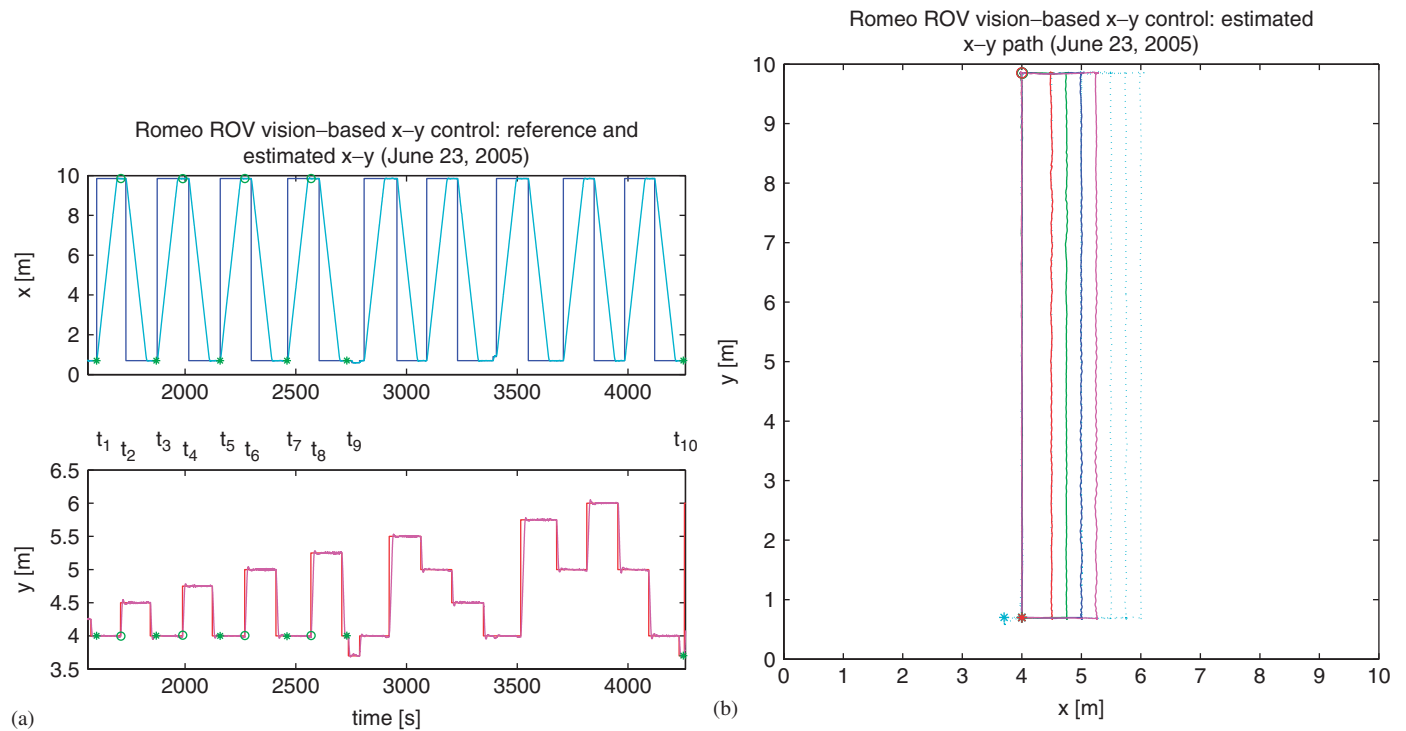| Δt	Z	Δx	Δy	Δx_{GT}	Δy_{GT}	$e_{\Delta x}$	$e_{\Delta y}$
$t_E - t_D$	0.80	0.03	0.00	−0.01	0.06	0.04	−0.06
$t_F - t_D$	0.80	0.01	0.00	−0.03	0.07	0.04	−0.07
$t_G - t_D$	0.81	0.02	0.00	−0.04	0.10	0.06	−0.10

Fig. 9. Catamaran yaw rate vs. rudder angle and propeller revolution rate.

sensor in the Ligurian Sea, Portofino Park area, in July 2003 and June 2005.

The video camera has been positioned inside a suitable steel canister, while the four red laser diodes are rigidly connected in the corners of a 13 cm side square, with their rays perpendicular to the image plane. The selected camera is the high-sensitivity (0.8 lux F1.2, 50 IRE; 0.4 lux F1.2, 30 IRE) Sony SSC-DC330 $\frac{1}{3}$in high-resolution Exwave HAD CCD Colour Camera. Images were acquired and processed in real time at 5 frames per second at the resolution of 360×272 RGB pixels by a PC equipped with a Leutron PicPort-Colour frame-grabber and a Pentium III CPU at 800 MHz running MS Windows2000. It is worth noting that the proposed image processing algorithm does not require a strong computational effort, and that most of the time was wasted in a rough display of the processed image to allow the human operator to monitor the system behavior in real time. Anyway, the relatively slow image processing frequency of 5 Hz is sufficient with respect to the vehicle dynamics. Calibration determined a focal length f

of about 1063.9 pixel (i.e. 1 pixel corresponds to about 1 mm with a scene depth of 1 m). The application, written in C++, using Intel Integrated Performance Primitives v2.0 for image processing and signal filtering, received the ROV telemetry, including acoustic altimeter data, via datagram sockets from the vehicle control system. The resulting optical device has been mounted downward-looking below the Romeo ROV (see Fig. 7). A special illumination system at diffuse light[1] was built in order to minimize the effects of the motion of the light source together the vehicle (Negahdaripour, 1998). The camera scene was illuminated by two 50 W halogen lamps covered by suitable diffusers and, working in the proximity of the seabed, i.e. at a range between 0.60 and 1.5 m, was not affected by the lamps mounted in front of the ROV for pilot/scientist video and photo cameras (see Fig. 2). The ROV operator could control the camera iris in order to optimize the scene illumination.

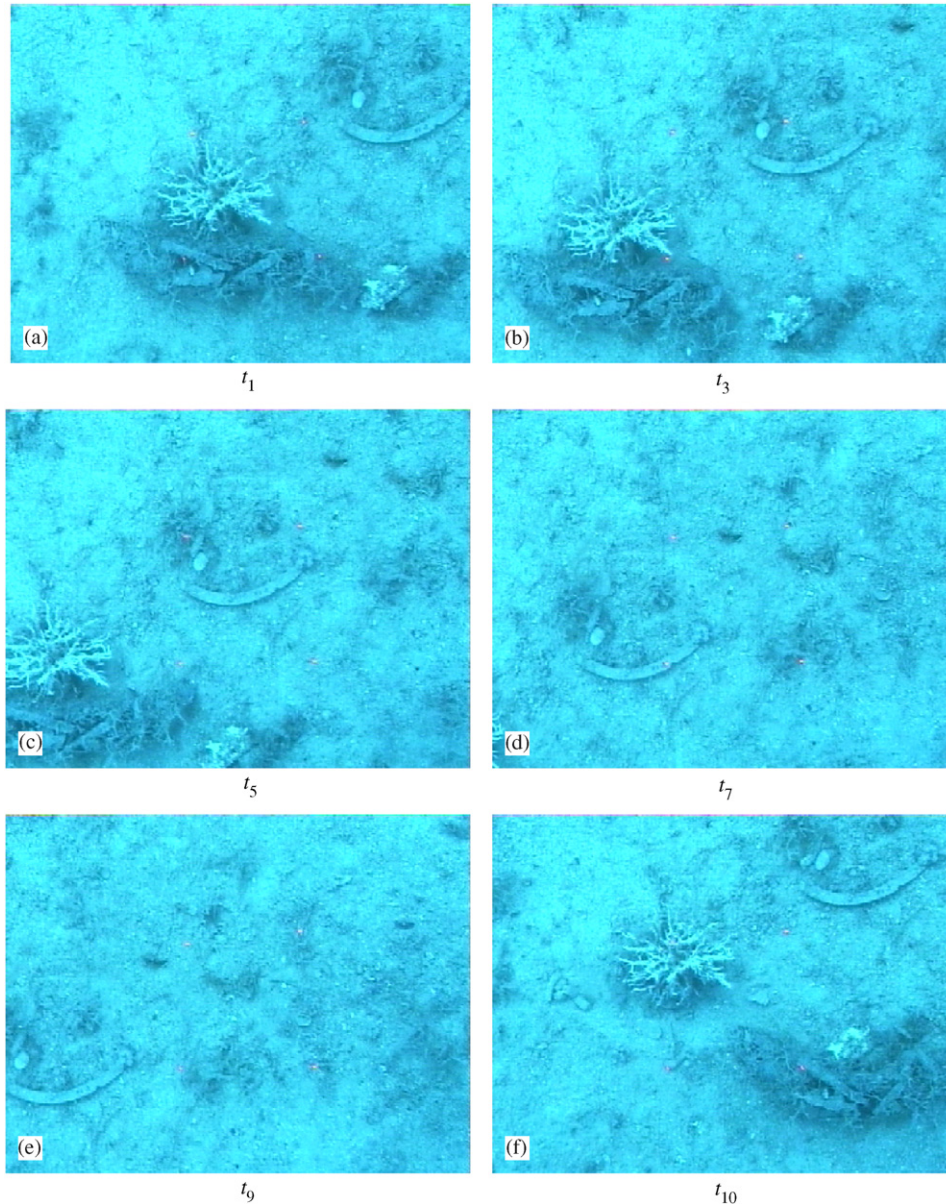[1] The system was designed and built by Giorgio Bruzzone.

Fig. 10. From top-left to bottom-right: ROV camera views at time t_1, t_3, t_5, t_7, t_9 and t_{10}.

During the experiments the ROV worked in auto-altitude using the image depth computed by the optical laser spot triangulation system as altitude measurement.

As discussed in Caccia and Veruggio (2000), the natural frequency $\omega_0 = \sqrt{\sigma^2 + \omega_n^2}$ of the dominating poles of the closed-loop linearized second-order system (13) was chosen so that $\omega_0 \Delta t \simeq 0.1 - 0.5$, where Δt is the sampling interval, according to the hints given in Åstrom and Wittenmark (1989), and the ratio ω_n/σ has been fixed to 0.1. Since the Butterworth band-stop filter of the estimated speed, whose coefficients are reported in Caccia (2003a), introduces a maximum delay of about 1.6 s, it seemed reasonable to assume this value as Δt. As a consequence, the parameters of the dynamics controller D_C were $\sigma = 0.3$, $\omega_n = 0.03$ and $\eta_{MAX} = 0.2$. The kinematics controller was parameterized,

on the basis of empirical considerations, by $\lambda = 0.2$, $\mu = 0.12$, $\xi_P^{MAX} = 0.1$, $\xi_I^{MAX} = 0.05$, $I_e^{ON} = 0.12$ and $I_e^{OFF} = 0.25$. During the tests the ROV worked in auto-heading with fixed orientation. The vehicle trajectory $\underline{x}$ was computed by integrating the vision-measured surge and sway (before stopband filtering) with the compass heading measurements according to Eq. (1).

In the following two basic experiments are discussed. The former test demonstrates the system capability of counteracting external disturbance and remaining over an interesting point, while the latter one shows the precision of the guidance and control system as well as of the vision-based motion estimator. In particular, the difference, $e_{\Delta x}$ and $e_{\Delta y}$, between the estimated displacements, Δx and Δy, obtained by integrating the vision-based surge and sway

measurements and the compass heading, and the ground-truth camera displacements, Δx_{GT} and Δy_{GT}, directly computed off-line from the motion of tokens tracked in reference images according to a discrete version of Eq. (7), was considered.

6.1. Disturbance rejection experiment

In the experiment discussed in the following, showing the system capability in counteracting external disturbances, the vehicle worked at an altitude, i.e. image depth, of about 80 cm, which corresponds to a field of view of about 21 × 28 cm in the images shown below. Surge and sway forces were zeroed at time t_D and t_F, while the ROV was hovering a target area, in order to show the effects of environmental disturbance, i.e. sea current and tether tension, on the vehicle. As shown in Fig. 8(a), the ROV drifted at a speed of about 5.5 cm/s, but, when the controller was switched on again, it was able to drive the vehicle again over the operating point. Fig. 8(b) shows that the control action was nonzero to compensate external disturbance during station-keeping.

Results, obtained by comparing the online estimated and ground-truth displacements, are shown in Table 1, where the error in the estimated horizontal position is lower than 0.18 m. The images taken by the ROV camera at time t_D, t_E, t_F and t_G (see Fig. 8) demonstrate that the precision of the vision-based dead-reckoning position estimate was satisfying for the ROV pilot. In particular, the displace-ment in the image sequence of the circular feature positioned about in the center of the picture taken at time t_D reveals a backward (negative Δx_{GT}) rightward (positive Δy_{GT}) motion of the vehicle, confirmed by the ground-truth camera motion, Δx_{GT} and Δy_{GT}, directly computed off-line from the displacements of tokens tracked in reference images according to a discrete version of Eq. (7).

6.2. System precision evaluation

In this trials the ROV moved along a rectangular grid of way-points (see Fig. 9(b)) periodically going back to a start point located in the left-bottom corner. The reference and estimated x and y positions are plotted in Fig. 9(a), pointing out the high precision of the position controller with respect to the estimated signal. A quantitative evaluation of the overall system precision, which is determined by the vision-based motion estimator accuracy, has been performed on the basis of multiple passages over the same location, which allow an off-line computation of ground-truth camera displacements. In particular, the passages on the starting location S and on the point P have been examined. The trial took about 40 min. Since images are taken from an altitude of about 1.3 m, they cover an area of about 0.35 m by 0.46 m. On the basis of the displacement of the white bush and the arc-like feature, denoting the point S, located in the center and in the top-right corner of the image taken at time t_1 (see Fig. 10), and of the white stone, denoting the point P, located towards
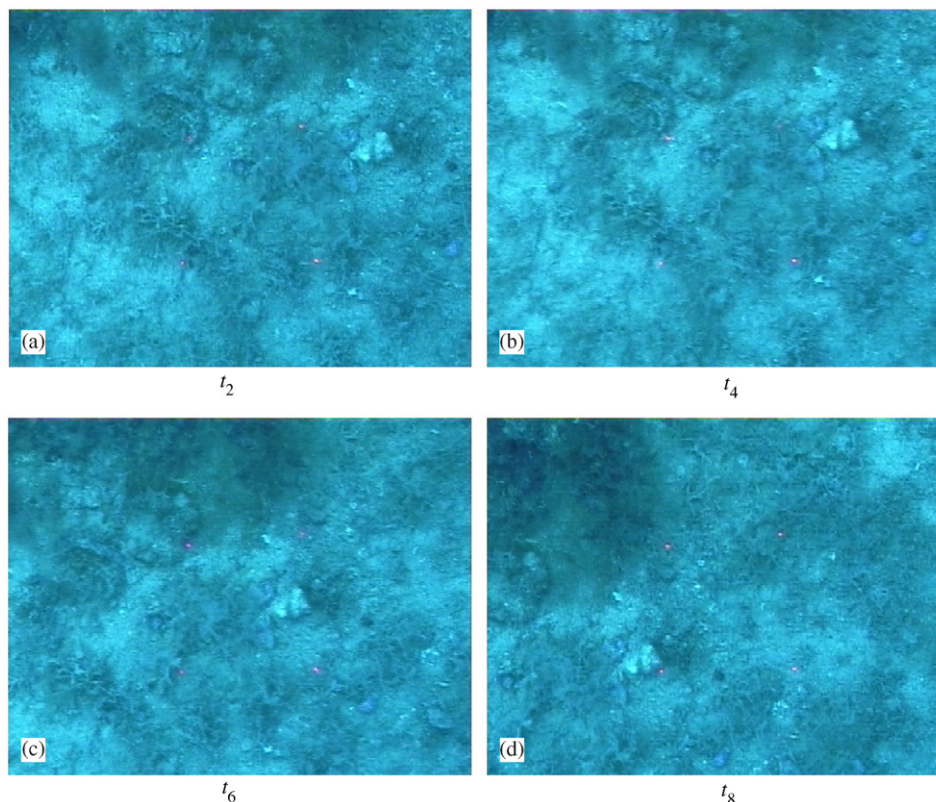

Fig. 11. From top-left to bottom-right: ROV camera views at time t_2, t_4, t_6 and t_8.

Table 2
Speed integral vs. ground-truth camera displacement from couple of images

Interval	Δt	Z	Δx	Δy	Δx_{GT}	Δy_{GT}	$e_{\Delta x}$	$e_{\Delta y}$	l_{path}	$\frac{|e|}{l_{path}}$ (%)
$t_3 - t_1$	270.0	1.30	0.00	0.00	0.03	0.08	-0.03	-0.08	19.30	0.442
$t_5 - t_1$	558.1	1.31	0.00	0.00	0.06	0.15	-0.06	-0.15	39.10	0.413
$t_7 - t_1$	860.1	1.31	0.00	0.00	0.13	0.24	-0.13	-0.24	59.40	0.459
$t_9 - t_1$	1130.1	1.30	0.00	0.00	0.15	0.33	-0.15	-0.33	80.20	0.451
$t_{10} - t_1$	2650.1	1.29	0.00	-0.30	-0.05	0.02	0.05	-0.34	181.1	0.065
$t_4 - t_2$	279.3	1.31	-0.00	0.01	-0.01	0.01	0.01	0.00	19.30	0.051
$t_6 - t_2$	560.1	1.28	-0.00	0.01	0.05	0.08	-0.05	-0.07	39.10	0.220
$t_8 - t_2$	859.3	1.30	-0.00	0.01	0.10	0.21	-0.10	-0.20	59.40	0.376

the top-right corner of the image taken at time t_2 (see Fig. 11), the reader can roughly evaluate the camera motion. In particular, referring to point S, a progressive forward (positive Δx_{GT}) rightward (positive Δy_{GT}) motion of the camera from t_1 to t_9, although the online estimated displacement was about zero. The camera is again over point S at time t_{10}, when the online estimated displacement is -0.30 m along the y-axis, with an error with respect to the ground-truth values of about 0.35 m (see Table 2 for numerical results). Anyway, as confirmed by the results relative to point P, the position error is always lower than the 0.6% of the covered path l_{path}.

7. Conclusion

Experimental results in ROV horizontal motion control using a vision-based motion estimator integrated with a dual-loop hierarchical guidance and control architecture has been presented in this paper, showing satisfying results in at sea trials carried out with the Romeo ROV. Tests were executed with the ROV operating at constant heading and altitude, allowing a direct comparison with results obtained by controlling the Romeo motion using the same guidance and control structure in a pool, where position was estimated from acoustic range measurements from environmental features (Caccia & Veruggio, 2000), neglecting more advanced gain-scheduling and integral control techniques presented in Khalil (1996). Anyway, promising results in vision-based estimate of yaw motion (Caccia, 2003b) should lead to the extension of the proposed approach to full, i.e. linear and angular, motion estimation and control on the horizontal plane.

Acknowledgments

References

Antonelli, G., Chiaverini, S., Sarkar, N., & West, M. (2001). Adaptive control of an autonomous underwater vehicle: Experimental results on ODIN. *IEEE Transactions on Control Systems Technology*, *9*(5), 756–765.

Åstrom, K. J., & Wittenmark, B. W. (1989). *Adaptive control.* USA: Addison-Wesley Publishing Company.

Caccia, M. (2003a). Pitch and roll disturbance rejection in vision-based linear speed estimation for UUVs. In *Proceedings of MCMC 2003* (pp. 313–318).

Caccia, M. (2003b). Vision-based linear motion estimation for unmanned underwater vehicles. In *Proceedings of ICRA 2003* (Vol. 1, pp. 977–982).

Caccia, M. (2005). Experiments in low cost high precision motion control for ROVs. In *Proceedings of ICRA 2005*.

Caccia, M., Bono, R., Bruzzone, G., & Veruggio, G. (2000). Unmanned underwater vehicles for scientific applications and robotics research: The ROMEO project. *Marine Technology Society Journal*, *24*(2), 3–17.

Caccia, M., Indiveri, G., & Veruggio, G. (2000). Modelling and identification of open-frame variable configuration unmanned underwater vehicles. *IEEE Journal of Oceanic Engineering*, *25*(2), 227–240.

Caccia, M., & Veruggio, G. (2000). Guidance and control of a reconfigurable unmanned underwater vehicle. *Control Engineering Practice*, *8*(1), 21–37.

Clark, J., Wallace, A. K., & Pronzato, G. L. (1998). Measuring range using a triangulation sensor with variable geometry. *IEEE Transactions on Robotics and Automation*, *14*(1), 60–68.

Fossen, T. I. (1994). *Guidance and control of ocean vehicles.* England: Wiley.

Fossen, T. I., & Sagatun, S. I. (1991). Adaptive control of nonlinear underwater robotic systems. In *Proceedings of IEEE conference on robotics and automation* (pp. 1687–1695).

Gracias, N. R., Van der Zwaan, S., Bernardino, A., & Santos-Victor, J. (2003). Mosaic-based navigation for autonomous underwater vehicles. *IEEE Journal of Oceanic Engineering*, *28*(4), 609–624.

Khalil, H. K. (1996). *Nonlinear systems.* Englewood Cliffs, NJ: Prentice-Hall.

Kinsey, J. C., & Whitcomb, L. L. (2003). Preliminary field experience with the DVLNAV integrated navigation system for manned and unmanned submersibles. In *Proceedings of first IFAC workshop on guidance and control of underwater vehicles*, Newport, UK (pp. 83–88).

Kinsey, J. C., & Whitcomb, L. L. (2004). Preliminary field experience with the DVLNAV integrated navigation system for oceanographic submersibles. *Control Engineering Practice*, *12*(12), 1541–1549.

Kondo, H., & Ura, T. (2004). Navigation of an AUV for investigation of underwater structures. *Control Engineering Practice*, *12*(12), 1551–1559.

Leabourne, K. N., Rock, S. M., Fleischer, S. D., & Burton, R. (1997). Station keeping of an ROV using vision technology. In *Proceedings of oceans'97* (Vol. 1, pp. 634–640).

Marco, D. B., & Healey, A. J. (2001). Command, control, and navigation experimental results with the NPS ARIES AUV. *IEEE Journal of Oceanic Engineering*, 26(4), 466–476.

Marks, R. L., Rock, S. M., & Lee, M. J. (1995). Real-time video mosaicking of the ocean floor. *IEEE Journal of Oceanic Engineering*, 20(3), 229–241.

Marks, R. L., Wang, H. H., Lee, M. J., & Rock, S. M. (1994). Automatic visual station keeping of an underwater robot. In *Proceedings of OCEANS '94* (Vol. 2, pp. 137–142).

Marques, L., Nunes, U., & de Almeida, A. T. (1998). A new 3d optical triangulation sensor for robotics. In *Proceedings of fifth international workshop on advanced motion control* (pp. 512–517).

Misu, T., Hashimoto, T., & Ninomiya, K. (1999). Optical guidance for autonomous landing of spacecraft. *IEEE Transactions on Aerospace and Electronic Systems*, 35(2), 459–473.

Negahdaripour, S. (1998). Revised definition of optical flow: Integration of radiometric and geometric cues for dynamic scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(9), 961–979.

Negahdaripour, S., & Xu, X. (2002). Mosaic-based positioning and improved motion-estimation methods for automatic navigation of submersible vehicles. *IEEE Journal of Oceanic Engineering*, 27(1), 79–99.

Negahdaripour, S., Xu, X., & Jin, L. (1999). Direct estimation of motion from seafloor images for automatic station keeping of submersible platforms. *IEEE Journal of Oceanic Engineering*, 24(3), 370–382.

Silvestre, C., Aguiar, A., Oliveira, P., & Pascoal, A. (1998). Control of the SIRENE underwater shuttle: System design and tests at sea. In *Proceedings of the 17th international conference on offshore mechanics and arctic engineering*, Lisbon, Portugal.

Smalwood, D. A., & Whitcomb, L. L. (2004). Model-based dynamic positioning of underwater robotic vehicles: Theory and experiment. *IEEE Journal of Oceanic Engineering*, 29(1), 169–186.

Whitcomb, L., Yoerger, D., Singh, H., & Howland, J. (1999). Advances in underwater robot vehicles for deep ocean exploration: Navigation, control, and survey operations. In *Proceedings of the ninth international symposium of robotics research*, Snowbird, USA.

Xu, X., & Negahdaripour, S. (1999). Automatic optical station keeping and navigation of an ROV; sea trial experiments. In *Proceedings of OCEANS '99* (Vol. 1, pp. 71–76).

Yoerger, D. R., & Slotine, J. E. (1985). Robust trajectory control of underwater vehicles. *IEEE Journal of Oceanic Engineering*, 10(4), 462–470.

Yoerger, D. R., & Slotine, J. E. (1991). Adaptive sliding control of an experimental underwater vehicle. In *Proceedings of IEEE international conference on robotics and automation* (pp. 2746–2751).

Yuh, J. (1990). Modelling and control of underwater robotic vehicles. *IEEE Transactions on Systems, Man and Cybernetics*, 20(6), 1475–1483.