Visual Inspection of Sea Bottom Structures by an Autonomous Underwater Vehicle

Gian Luca Foresti, *Senior Member, IEEE*

Abstract—This paper describes a vision-based system for inspections of underwater structures, e.g., pipelines, cables, etc., by an autonomous underwater vehicle (AUV). Usually, underwater inspections are performed by remote operated vehicles (ROVs) driven by human operators placed in a support vessel. However, this task is often challenging, especially in conditions of poor visibility or in presence of strong currents. The system proposed allows the AUV to accomplish the task in autonomy. Moreover, the use of a three-dimensional (3-D) model of the environment and of an extended Kalman filter (EKF) allows the guidance and the control of the vehicle in real time. Experiments done on real underwater images have demonstrated the validity of the proposed method and its efficiency in the case of critical and complex situations.

Index Terms—Autonomous vehicles, extended Kalman filter (EKF), feature extraction, object tracking, underwater imaging.

I. INTRODUCTION

AUTONOMOUS underwater vehicles (AUVs) are generally designed for inspection of underwater environments, especially submersed man-made structures (pipelines, cables, columns of offshore platforms, etc.) [1]–[7]. Such inspections must be done periodically in order to control the state of maintenance of the above structures. Until now, remote operated vehicles (ROVs), driven by human operators placed in a support vessel, have been utilized. The main task for a human operator is the so-called *stationkeeping* task [8], which consists of keeping the vehicle in the desired position and orientation, possibly in the presence of various types of disturbances such as undersea currents. During long term missions, this operation often becomes hard and challenging, especially in underwater environments characterized by dark waters or by presence of strong currents. To give an accurate measurement of the vehicle's shifting through the interpretation of data furnished by multiple sensors can allow autonomous stationkeeping and, consequently, can help the navigation of the AUV. Furthermore, the complexity of the AUV navigation task can be reduced if the vehicle operates close to the sea bottom. In fact, the textures on the bottom can provide important visual information for movement assessment. Image processing methods can be applied to extract visual information from the underwater environment. For example, the detection of bottom landmarks, whose position and shape are *a-priori* known, can be used to help the driving system of the AUV [9].

Over the last decade, there have been major technological advances in the development of visual sensing capability as a mechanism for terrestrial robotic systems to extract information about the environment. Several image-processing methods have been developed to estimate the motion of a vehicle with respect to a stationary environment and to reconstruct three–dimensional (3-D) scenes from image sequences for robot navigation purposes. Generally, these methods compute the displacement on the image plane of specific scene features [9], the displacement of brightness patterns over the whole image (*optical flow*) [2], [10], or the correlation between features extracted from consecutive pairs of images [11]. The tracking of remarkable scene features, as, for instance, single points, edges, or optical flow, etc., gives significant results only in laboratory test sites [12] or in environments where lights are controlled [13] and limited results in a restricted set of outdoor applications [14], [15]. Unfortunately, most machine vision methods developed for these systems cannot be employed directly underwater due to particular characteristics of the environment. For example, nonuniform lighting, restricted visibility (due to illumination attenuation), forward/back-scattering in the medium (due to the presence of some material in suspension), etc., make the feature extraction and matching process unstable over consecutive frames and the consequent image interpretation task very complex.

This paper describes an artificial vision system for sea bottom inspection of man-made structures, i.e., pipelines, with an AUV. The navigation of the AUV is based on the detection of representative and robust features, e.g., geometric elements like pipeline borders, into the underwater environment. The AUV utilized is torpedo-like, and composed of the fore, the vessel body, and the aft [16]. The vessel body is the only part watersealed and contains the electronic equipment. The fore and aft are water filled and contain the thrusters and some sensors, i.e., some inertial sensors, a sonar, a depth sensor, a charge coupled device (CCD) camera, and an image acquisition system able to process about 25 frames/s. The AUV has been equipped with two processors and two industrial versa module eurocard (VME) buses. In particular, one processor is assigned to the vision-based subsystem whereas the other one is assigned to the remaining subsystems (AUV motion control, data communication, etc.).

The problem relative to inspections of underwater pipelines has been tackled by different researchers. Zingaretti *et al.* [17] proposed a hierarchical vision-based system, which first detect on the image the lines corresponding to the borders of the pipeline, and then it uses a Kalman filter to estimate the line parameters. Grau [7] proposed the use of a specific architecture, planned for texture extraction and able to operate at a frequency

of 50 frames per second to detect pipelines in underwater images. Branca [3] proposed a vision system for tracking a pipeline based on the correspondence between features, extracted from consecutive image pairs. The main limitation of these methods is that the detection of the pipeline is performed only on the two-dimensional (2-D) image. Moreover, they can fail when, due to low illumination conditions or to the presence of sand covering the pipeline, the whole pipeline or some of its parts disappear from the image for some consecutive frames.

More complex techniques have been proposed to track underwater cables. For instance, Kojina *et al.* [4] considered a cable with a low voltage current, and tracked it by analyzing the magnetic field by two three-axial magnetometers. Balasuriya and Ura [5], [6] proposed a system able to integrate the visual parameters gathered by the appearance of the cable in the image with the ones obtained by on-board inertial sensors to obtain a kind of linguistics controls (to draw near/far, to move to the right or left, etc.) which would be then utilized for the navigation of an AUV.

The main innovative aspect of the system proposed in this paper is its ability to infer 3-D information about the AUV pose and motion in the scene starting from 2-D observations of the pipeline borders in a long underwater image sequence. A 3-D geometric model of the scene is used to add to the system the *a-priori* information (*geometric constraints*) needed to solve the ill-posed 2-D/3-D transformation problem. An extended Kalman filter (EKF) [18] is applied to estimate the relative position and orientation between the video camera, placed in a fixed position on the AUV, and the inspected pipeline. The estimation is done by integrating the data given by the inertial sensors and the position in the image of the two lines corresponding to the pipeline borders. In particular, a voting-based method [19] for line–feature extraction in multisensor images (particularly suited for underwater images) has been considered. It is quite important to note that the proposed method allows the estimation of the AUV pose and motion parameters also in critical situations such as low visibility or a pipeline partially or totally covered by sand, etc.

The paper is organized as follows. Section II describes the 3-D model utilized for the scene reconstruction, while Section III examines the proposed vision system. Section IV describes the feature extraction phase where the lines corresponding to the pipeline borders are searched. Section V describes the EKF applied to estimate the AUV pose starting from the observation of the pipeline edges on the image plane; in particular, the system model, the measure model, and the constraints, which allow to compute the equations of the EKF are described. Finally, Section VI presents several experimental results obtained by testing the system on real underwater image sequences.

II. THREE-DIMENSIONAL SCENE MODEL

The laws of perspective transformations allow mapping at 3-D points onto 2-D image points and also inference of partial information about 3-D scene points from 2-D image points. 2-D into 3-D transformations are completely defined whenever a
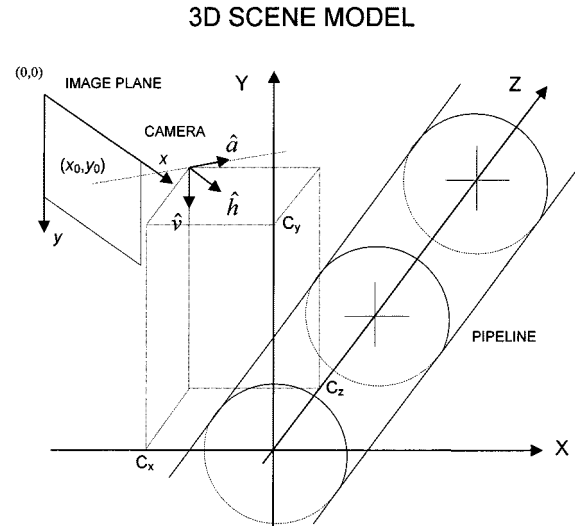
3D SCENE MODEL

Fig. 1.   Three-dimensional scene model and *pin-hole* camera model.

3-D scene model is established and some geometric constraints are considered.

The general model of the 3-D scene (Fig. 1) used in this paper is based on the *pinhole* model [20] of the camera. The underwater pipeline is considered as a cylinder of diameter d. The (X, Y, Z) general reference system (*world reference system*) is positioned as to have the Z-axis coincident with the pipeline axis and the X-axis parallel to the sea surface. The position of the camera with respect to the (X, Y, Z) reference system is given by the coordinates C_x, C_y, and C_z, while the orientation of the camera is given by the three versors $\hat{a}$, $\hat{v}$, and $\hat{h}$. The axes of the image reference system (x, y) are parallel to the versors $\hat{h}$ and $\hat{v}$, respectively. The y-axis turns underneath and the origin of the image reference system is placed on the top-left corner of the image. The versor $\hat{a}$ has the same direction of the optical axis. As the camera is fixed and placed in a given point of the AUV, it is possible to define a transformation between the camera reference system and the AUV reference system, which is centered in its center of mass.

The camera model together with the perspective transformation laws allow to define the relation between a point $P \equiv (P_x, P_y, P_z)$ in the 3-D reference system and the corresponding point $p \equiv (i, j)$ on the 2-D image plane

$$\begin{bmatrix} i' \\ j' \\ w \end{bmatrix} = \begin{bmatrix} T_0 & T_1 & T_2 & T_3 \\ T_4 & T_5 & T_6 & T_7 \\ T_8 & T_9 & T_{10} & T_{11} \end{bmatrix} \begin{bmatrix} P_x \\ P_y \\ P_z \\ 1 \end{bmatrix} = [T] \begin{bmatrix} P_x \\ P_y \\ P_z \\ 1 \end{bmatrix}. \quad (1)$$

Solving the equation system in (1) with respect to i and j we obtain

$$i = \frac{i'}{w} = \frac{T_0 \cdot P_x + T_1 \cdot P_y + T_2 \cdot P_z + T_3}{T_8 \cdot P_x + T_9 \cdot P_y + T_{10} \cdot P_z + T_{11}} \quad (2a)$$

$$j = \frac{j'}{w} = \frac{T_4 \cdot P_x + T_5 \cdot P_y + T_6 \cdot P_z + T_7}{T_8 \cdot P_x + T_9 \cdot P_y + T_{10} \cdot P_z + T_{11}} \quad (2b)$$

where the component w allows to eliminate the nonlinearity introduced by the perspective equations ($i = f(X/Z)$ and $j = Y/Z$) [21], and the coefficients T_i ($i = 0, \ldots, 11$) of the $[T]$

perspective transformation matrix (called also calibration matrix or *MIR* matrix) can be computed as follows:

$$T_0 = (f/\delta x)h_x + i_0 a_x \tag{3a}$$

$$T_1 = (f/\delta x)h_y + i_0 a_y \tag{3b}$$

$$T_2 = (f/\delta x)h_z + i_0 a_z \tag{3c}$$

$$T_3 = -(f/\delta x)C \cdot \hat{h} + i_0 C \cdot \hat{a} \tag{3d}$$

$$T_4 = (f/\delta y)v_x + j_0 a_x \tag{3e}$$

$$T_5 = (f/\delta y)v_y + j_0 a_y \tag{3f}$$

$$T_6 = (f/\delta y)v_z + j_0 a_z \tag{3g}$$

$$T_7 = -(f/\delta y)C \cdot \hat{v} + j_0 C \cdot \hat{a} \tag{3h}$$

$$T_8 = a_x \tag{3i}$$

$$T_9 = a_y \tag{3j}$$

$$T_{10} = a_z \tag{3k}$$

$$T_{11} = -C \cdot \hat{a} \tag{3l}$$

where

(i_0, j_0) coordinates of the center of the image;

f focal length of the camera;

δx and δy horizontal and vertical sampling ranges, respectively.

If the physical dimensions of the image plane are $K \times H$, then the relations between the (x, y) physical coordinates and the row and column indexes (i, j) of the $N \times N$ image are given by

$$x = \left(j - \frac{N-1}{2}\right)\delta x \tag{4a}$$

$$y = \left(-i + \frac{N-1}{2}\right)\delta y \tag{4b}$$

where $\delta x = K/N$ and $\delta y = H/N$.

To carry out the 2-D into 3-D transformation (one to many transformation), which consists of determination of the 3-D point $P \equiv (P_x, P_y, P_z)$ corresponding to an identified point $p \equiv (i, j)$ on the image plane, some geometric constraints should be considered, e.g., at least one among the coordinates P_x, P_y, and P_z must be known. For instance, if the coordinate P_z is known, the other two coordinates P_y and P_x can be obtained by solving the system formed by (2a) and (2b).

Definition 1: The camera model is completely defined when the matrix $[T]$ [i.e., the coefficients T_i ($i = 0, \ldots, 11$)] is known for each position of the camera.

The coefficients of the matrix $[T]$ depend on the intrinsic and extrinsic parameters of the camera and can be calculated using a *calibration algorithm* [22]. The intrinsic parameters are the geometrical and optical parameters of the camera (focal length f, sampling intervals δx and δy), while the extrinsic ones are those concerning the position (C_x, C_y, C_z) and the orientation $(\hat{a}, \hat{v}, \hat{h})$ of the camera with respect to the general reference system (X, Y, Z).

Two different methods, i.e., the direct and the indirect method, can be applied to obtain the camera calibration. The first, and more immediate, consists of two steps: a) measuring the parameters of the camera physically, and then b) calculating
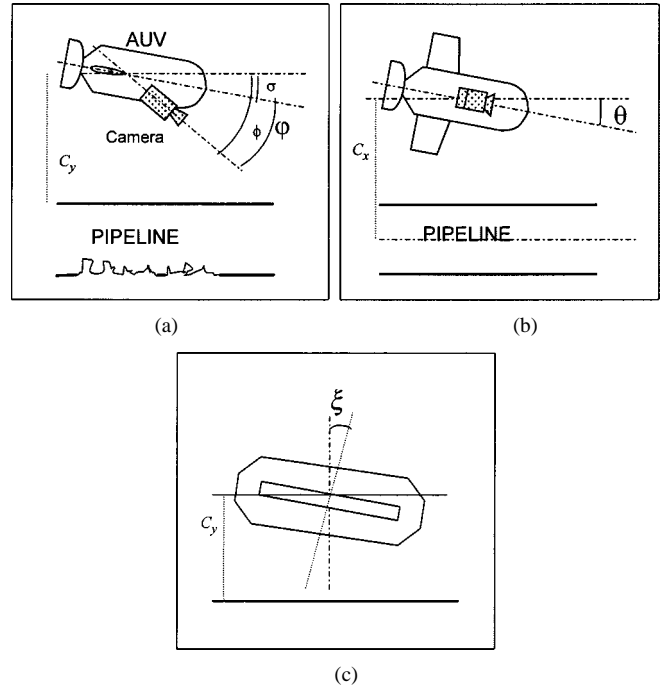

Fig. 2. (a) Tilt, (b) pan, and (c) roll camera orientation angles displayed according to the vehicle and the pipeline positions.

the T_i coefficients by solving the equations (3a)–(3l). The second approach considers the matrix $[T]$ as an autonomous entity, which can be determined if at least six noncoplanar points in the 3-D general reference system and their corresponding points in the image plane are known [22].

In the proposed system, the direct method has been chosen as the intrinsic parameters given by the camera manufacturer are constant. The extrinsic parameters are estimated every time a new image is acquired by the camera by means of an EKF.

Let φ be the angle between the optical axis of the camera and the XZ plane, i.e., *tilt* angle [Fig. 2(a)], and let θ be the angle between the optical axis of the camera and the YZ plane, i.e., *pan* angle [Fig. 2(b)]. Let ξ be the angle between the optical axis of the camera and the XY plane, i.e., *roll* angle [Fig. 2(c)]. Consequently, the three versors $\hat{a}$, $\hat{v}$, and $\hat{h}$ are the following:

$$\hat{h} = \begin{pmatrix} \cos\theta \cdot \cos\xi \\ -\sin\xi \cdot \cos\theta \\ -\sin\theta \end{pmatrix}$$

$$\hat{v} = \begin{pmatrix} -\sin\varphi \cdot \sin\theta \cdot \cos\xi - \cos\varphi \cdot \sin\xi \\ -\cos\varphi \cdot \cos\xi + \sin\varphi \cdot \sin\theta \cdot \sin\xi \\ -\sin\varphi \cdot \cos\theta \end{pmatrix}$$

$$\hat{a} = \begin{pmatrix} \cos\varphi \cdot \sin\theta \cdot \cos\xi - \sin\varphi \cdot \sin\xi \\ -\sin\varphi \cdot \cos\xi + \sin\theta \cdot \sin\xi \cdot \cos\varphi \\ \cos\varphi \cdot \cos\theta \end{pmatrix}.$$

This constraint implies that the still unknown variables, whose determination allows the calculation of the T_i coefficients, are C_x, C_y, C_z, θ, φ, and ξ. In the next Section, the development of the proposed AUV guidance system will be explained on the basis of the 3-D scene model and the assumptions on the camera position and orientation already given.
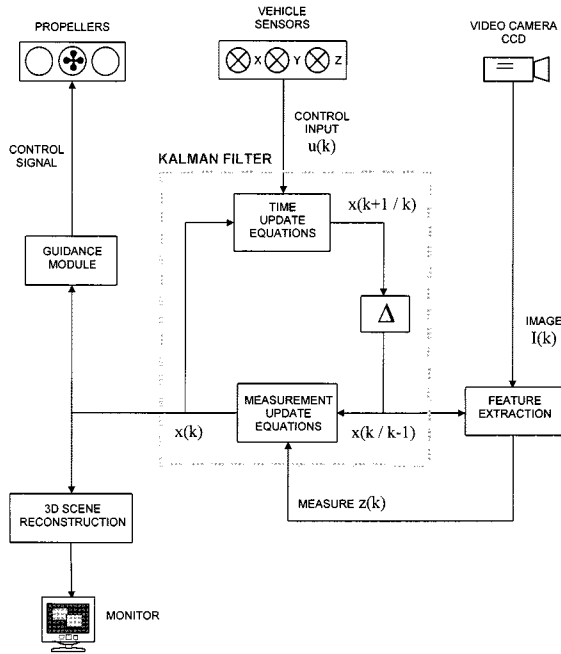
AUV GUIDANCE SYSTEM

Fig. 3. Functional scheme of the AUV guidance system.

Fig. 4. Points, i_1, i_2, i_3, and i_4, used as 2-D measures of the EKF; these points are obtained as intersection between the image borders and the detected lines associated with the pipeline edges.

III. SYSTEM DESCRIPTION

Fig. 3 shows the general functional scheme of the proposed AUV guidance system. The system is mainly based on the use of an EKF [18]. The EKF is characterized by two sets of equations, i.e., the *measurement updating* equations (*updating step*) and the *time updating* equations (*prediction step*), which allow to estimate the state of a system on the basis of a set of measures that can be affected by noise and/or characterized by errors.

In the proposed system, the EKF is used to estimate the orientation and the position of the camera (and, consequently, of the AUV) with respect to the position of the underwater pipeline in the general reference system. Let us define the vector $\boldsymbol{x} = (C_x, C_y, C_z, \theta, \varphi, \xi)$ as the *system state*. The estimation of the vector $\boldsymbol{x}$ is based on three elements:

1) the position of the two lines corresponding to the edges of the pipeline in the image (Fig. 4) (only the abscissa of the intersection points of the two lines with the image borders are considered);
2) the vehicle motion parameters, i.e., the lateral, vertical, and frontal speeds, s_x, s_y, and s_z, and the lateral, vertical, and frontal accelerations, $\boldsymbol{\kappa}_x$, $\boldsymbol{\kappa}_y$, and $\boldsymbol{\kappa}_z$ given by the inertial sensors;
3) the 3-D scene model.

Let us define the vector $\boldsymbol{z} = (p_1, p_2, p_3, p_4)$ as the *system measure*.

The detection of the straight lines corresponding to the borders of the pipeline in each image is given by the *feature extraction* module, which will be described in the next section. This module is essentially composed of two parts: 1) a line extraction submodule, which is based on an improvement of the Hough
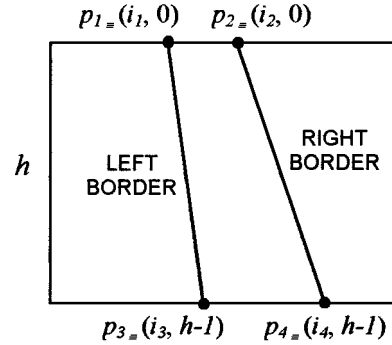
Transform method [19] and 2) a *grouping* submodule, which locates, in a set of candidate lines, the two that most likely belong to the pipeline borders. The grouping operation is performed on the basis of a previous estimate of the system state and on the geometrical characteristics of the scene.

After having estimated the relative position between the pipeline and the camera, and consequently, the distance and the orientation of the vehicle with respect to the pipeline (as the camera is fixed on the AUV), a suitable control signal can be generated to guide the vehicle in the desired position. Usually, the AUV runs at a constant distance (2–3 m) over the pipeline and with a motion direction parallel to the main pipeline axis. Finally, a user friendly man–machine interface (Fig. 5) is used to display in real-time to the remote operator the input images, the detected pipeline borders, the estimated AUV pose and motion parameters (AUV status), and the 3-D reconstruction of the pipeline on the basis of the AUV status and the detected borders.

IV. FEATURE EXTRACTION

The flow diagram of the feature extraction module, for the determination of the two straight lines corresponding to the external edges of the pipeline in the observed image, is presented in Fig. 6. Three submodules compose it: edge extraction, straight-line extraction, and straight-line grouping. After having extracted the edges from the image, a voting based approach is utilized to detect straight lines [19]. Among the extracted lines, the two lines corresponding to the pipeline borders are then searched for by the grouping module.

The Canny filter [23] is applied to extract edges from the input image $I(x, y)$. It furnishes also information about the phase $F(k)$ and the modulus $M(k)$ of the detected edges. An improvement of the Hough transform algorithm has been used to extract straight lines from underwater images [19]. The main problem to be solved in this phase is the extreme irregular alignment of the edges corresponding to the pipeline borders. Seaweed, fouling, and other disturbances typical of the underwater environment make extremely difficult in these conditions to distinguish between the instances of a line effectively induced by the pipeline edge, and those depending on the random alignment of the many other extracted edges. The changes we brought in
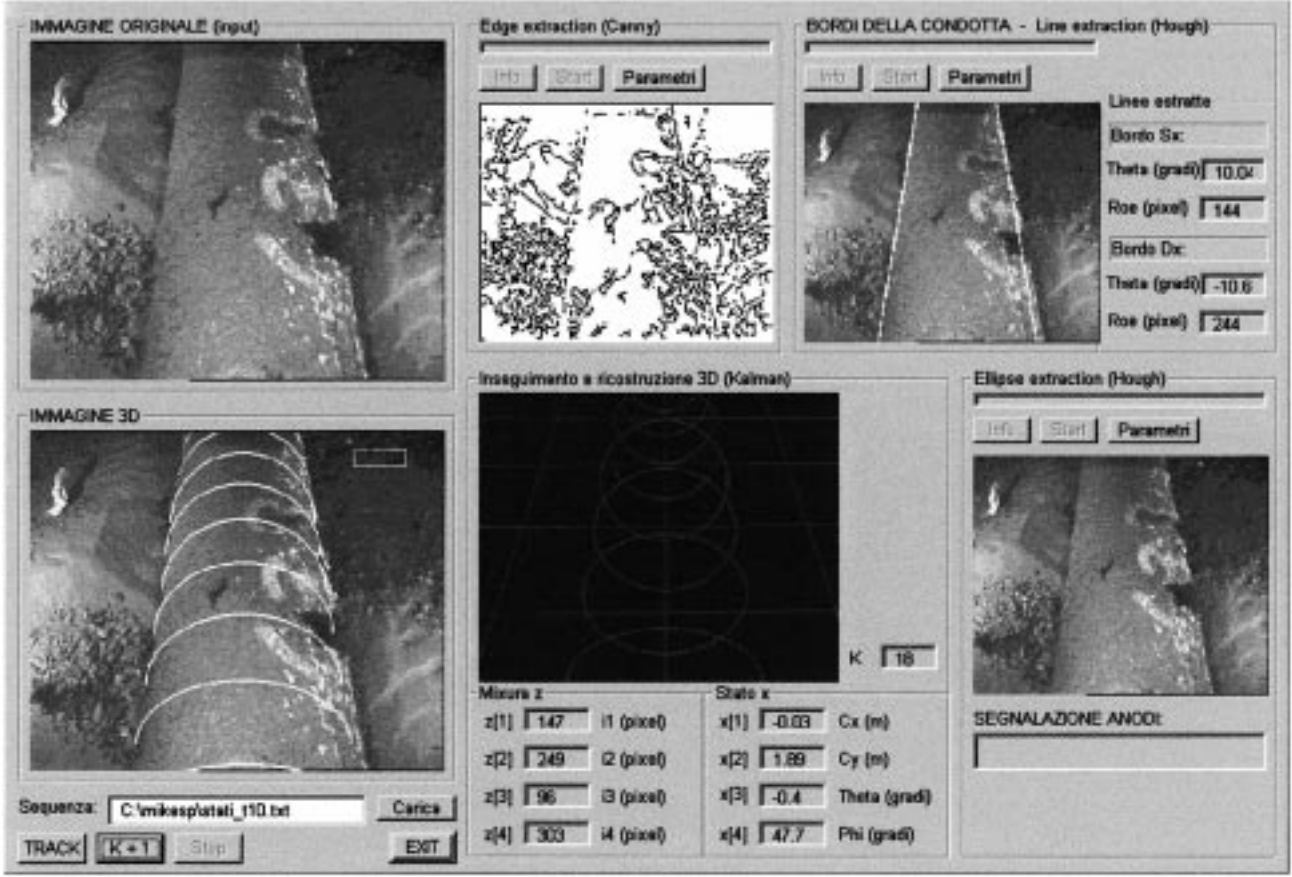
Fig. 5.   Man–machine interface of the AUV guidance system.

the standard version of the Hough transform technique allows to extract, for each instance of the detected straight line, the following information:

1) the α and ρ parameters, which locate it (*normal parametrization*) [Fig. 7(a)];
2) the average amplitude (module) M of the edges that vote for it;
3) a monodimensional array, called *domain* (dom[i]), which allows, in the case that the pipeline borders are partially occluded, storage of information about the end points of collinear segments on the pipeline border [Fig. 7(b)].

After having extracted all straight lines from the input image, it is possible to reinforce the desired line instances by means of "restorations" of the domains as filling up of small holes, removal of some isolated points or segments, and so on. Then, a reliability parameter Λ is computed for each line by considering its average amplitude, its length and its robustness.

Let $r = (\alpha_r, \rho_r, M_r, \text{dom}_r)$ be a generic straight line. Let d_1 be the index of the first bit equal to 1 of dom_r and d_2 the index of the last bit equal to 1 of dom_r, i.e.,

$$d_1 = \min\{i \in [0, h) \,|\, \text{dom}_r[i] = 1\} \qquad (5a)$$
$$d_2 = \max\{i \in [0, h) \,|\, \text{dom}_r[i] = 1\}. \qquad (5b)$$

Let n_r be the number of bits equal to 1 in the domain of the straight line r. Let $l_r = (d_2 - d_1) \cdot \tan(\alpha_r)$ be the length of the straight line r and $L_r = h \cdot \tan(\theta_r)$ its maximum length, where h represents the high of the image (expressed in pixels)

[Fig. 7(b)]. We define the ratio between the parameters l_r and L_r as a reliability parameter $PL_r \in [0, 1]$ of the length of the straight-line r, i.e.,

$$PL_r = \frac{l_r}{L_r} = \frac{(d_2 - d_1)}{h}. \qquad (6a)$$

Then, we define the ratio between n_r and the difference $(d_2 - d_1)$ as a robustness parameter $PR_r \in [0, 1]$ of the straight-line r, i.e.,

$$PR_r = \frac{n_r}{(d_2 - d_1)}. \qquad (6b)$$

Moreover, we define the parameter $PF_r \in [0, 1]$ as a reliability parameter of the strength of the straight-line r, i.e.,

$$PF_r = \frac{M_r}{\max_{r \in R}(M_r)} \qquad (6c)$$

where R is the set of all candidate straight lines.

The straight lines characterized by low PL and PR values are eliminated from the set R. In Fig. 8(a) and (b), two examples of straight lines extracted from underwater images are given. The detected straight lines with $PL > 0.5$ are shown in Fig. 8(c) and (d), respectively. In Fig. 8(e) and (f), the detected straight lines with $PL > 0.5$ and $PR > 0.65$ are shown.

Finally, we define the reliability parameter $\Lambda_r \in [0, 1]$ of the straight line r the following quantity:

$$\Lambda_r = \gamma_{PL} \cdot PL_r + \gamma_{PR} \cdot PR_r + \gamma_{PF} \cdot PF_r \qquad (6d)$$
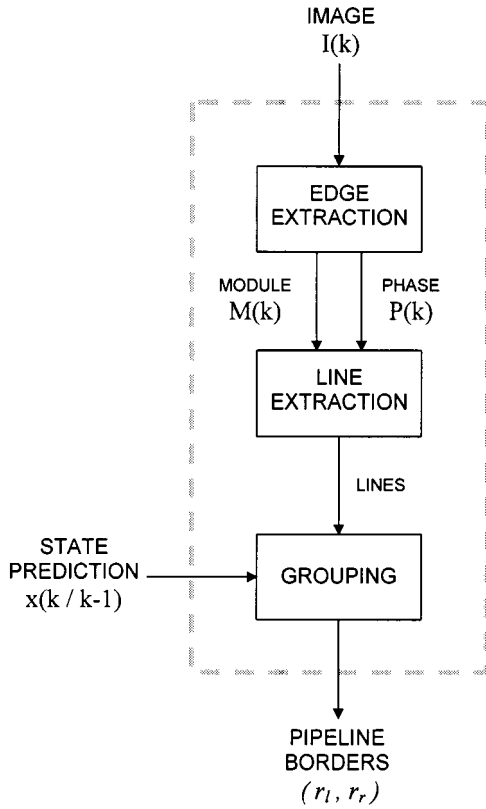
FEATURE EXTRACTION MODULE

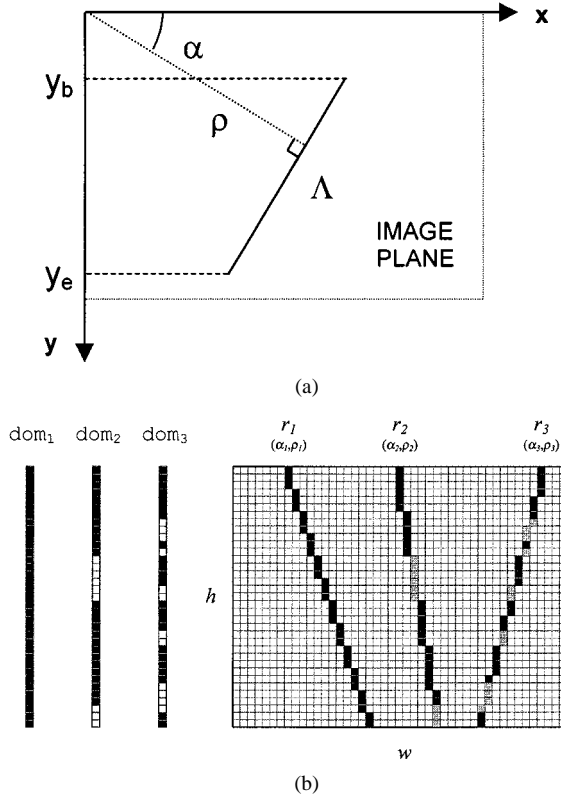Fig. 6. Functional scheme of the feature extraction module.

(a)

(b)

Fig. 7. (a) Normal parametrization of a straight line on the image plane. (b) Representation of the domains associated with three different straight lines.

where γ_{PL}, γ_{PR} and γ_{PF} are weights used to normalize the different parameters and to enforce the most significant contributions. The following relation holds: $\gamma_{PL} + \gamma_{PR} + \gamma_{PF} = 1$.

So, at the end of this phase, every line is characterized by the following parameters: $(\alpha, \theta, y_b, y_e, \Lambda)$ where y_b and y_e are the y coordinates of the line endpoints, and Λ is the associated reliability parameter. Tables I and II show the parameters obtained for the straight lines detected in Fig. 8(e) and (f), respectively.

During the grouping phase, only pairs of lines which are compatible with the geometrical characteristic of the scene and the perspective laws are considered. For example, it is not possible that two candidate lines representing the pipeline borders intersect or diverge. Then, for each compatible pair of lines, a *grouping reliability* parameter (ΛR), which takes into account also the geometrical characteristics (symmetry, convergence, etc.) of the lines, is computed. Moreover, the prediction of the position of the searched lines (when it is available by the EKF, is also used to choose the best pair of lines among the compatible pairs characterized by a high reliability parameter. When the position of the sought lines cannot be predicted, the compatible pair of lines with the highest grouping reliability parameter, is selected. Fig. 8(g) and (h) show the compatible pairs of lines with the highest grouping reliability parameter among the lines extracted respectively from the images in Fig. 8(a) and (b). The parameters of the chosen lines are pointed out in Tables I and II, respectively.

However, it can happen that no lines could be extracted or that starting from the extracted ones there are no compatible pairs. For instance, the following occurrences can be underlined:

1) pipeline partially or totally covered by sand;
2) presence of structures, like trestles [16], which limit the camera field of view;
3) low illumination conditions or low contrast in the image;
4) large quantity of material in suspension, etc.

In these cases, the feature extraction module cannot provide measurements to the EKF, and so the evaluation of the state variables is based only on the data given by the inertial sensors.

V. AUV POSE AND MOTION ESTIMATION

A. System Model

An EKF is applied to estimate the camera position and orientation, i.e., the parameters, C_x, C_y, C_z, θ, φ, ξ, which allow computation of the position and heading of the AUV. The model used to predict the evolution of the selected variables is called the *system model*.

It is worth noting that the C_z coordinate of the camera (or of the center of mass of the AUV) can be easily obtained by means of a depth or a sonar sensor. Moreover, as the diameter d of the pipeline can be considered constant for long sections, the position on the image plane of the lines corresponding to the pipeline borders does not depend on the value of the C_z coordinate. In this way, it is possible to simplify the system model by excluding the C_z coordinate from the quantities of interest and consequently from the system state vector x. Indeed, it is likely to consider the camera as always lying on a fixed Z plane.

The movement of the camera along the X, Y, and Z-axes strictly depends on the motion of the AUV: the lateral, vertical,
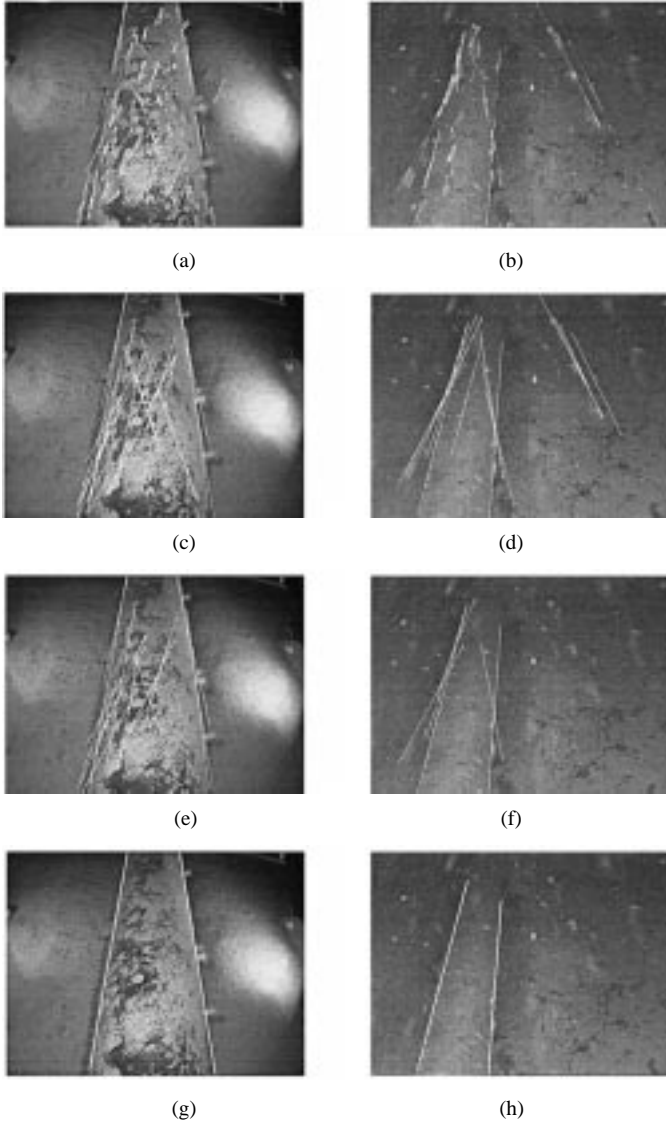
Fig. 8. (a), (b) Straight lines extracted from two different underwater images of a pipeline. (c), (d) Straight lines with $PL > 0.5$. (e), (f) Straight lines with $PL > 0.5$ and $PR > 0.65$. (g), (h) The pair of straight lines with the highest grouping reliability parameter ΛR.

TABLE I
VALUES OF THE PL, PR, PF, AND Λ PARAMETERS COMPUTED ON THE STRAIGHT LINES DETECTED ON THE IMAGE IN FIG. 8(a). THE Λ PARAMETER HAS BEEN COMPUTED WITH $\gamma_{PL} = 0.5$, $\gamma_{PR} = 0.35$, AND $\gamma_{PF} = 0.15$

N°	α	ρ	PL	PR	PF	Λ
1	9.8	153	0.96	0.70	0.51	0.80
2	12.6	153	0.49	1.00	0.5	0.67
3	170.8	-219	0.99	0.74	0.93	0.89
4	28.7	219	0.48	1.00	0.69	0.70
5	14.7	171	0.80	0.81	0.56	0.76
6	10.5	171	0.75	0.87	0.65	0.78
7	17.5	171	0.55	0.82	0.49	0.64
8	23.8	195	0.68	0.85	0.61	0.73
9	172.9	-225	0.78	0.79	1.00	0.82
10	11.2	189	0.61	0.68	0.91	0.68
11	26.6	231	0.62	0.72	0.89	0.70
12	25.2	231	0.62	0.73	0.87	0.70
13	9.1	183	0.52	0.77	0.83	0.65
14	167.3	-213	0.66	0.80	1.00	0.76
15	161.0	-123	0.47	0.67	0.73	0.58
16	28.7	195	0.49	0.69	0.61	0.58
17	27.3	201	0.54	0.66	0.58	0.59
18	17.5	225	0.49	0.76	0.64	0.61
19	16.1	225	0.51	0.79	0.62	0.62

TABLE II
VALUES OF THE PL, PR, PF, AND Λ PARAMETERS COMPUTED ON THE STRAIGHT LINES DETECTED ON THE IMAGE IN FIG. 8(b). THE Λ PARAMETER HAS BEEN COMPUTED WITH $\gamma_{PL} = 0.5$, $\gamma_{PR} = 0.35$, AND $\gamma_{PF} = 0.15$

N°	α	ρ	PL	PR	PF	Λ
1	14.7	129	0.85	0.87	0.82	0.85
2	16.8	129	0.60	0.71	0.64	0.65
3	28.7	147	0.54	0.66	0.61	0.59
4	21.7	135	0.41	0.69	0.63	0.54
5	169.4	-123	0.52	0.85	0.75	0.67
6	4.2	171	0.78	0.78	1.00	0.81

and frontal speeds, s_x, s_y, and s_z, and the lateral, vertical, and frontal accelerations, κ_x, κ_y, and κ_z are measured by the inertial sensors. In the proposed approach, the evolution of the variables C_x and C_y is modeled by approximating the motion of the AUV in the interval τ between two successive image acquisitions with uniform rectilinear motion. By considering null the acceleration vector, i.e., $\kappa_x = \kappa_y = \kappa_z = 0$, we obtain

$$\boldsymbol{C}_{k+1} = \boldsymbol{C}_k + [\Psi] \begin{pmatrix} s_{xk} \\ s_{yk} \\ s_{zk} \end{pmatrix} \cdot \tau \qquad (7)$$

where $[\Psi]$ is shown in (12) at the bottom of the next page and $\sigma_k = \varphi_k - \phi$. Solving the system for the interesting variables C_x, C_y, we obtain

$$C_{xk+1} = f_1(C_{xk}, C_{yk}, \varphi_k, \theta_k, \xi_k)$$
$$= C_{xk} + [s_{xk} \cos\theta_k \cdot \cos\xi_k - s_{yk}(\sin\sigma_k \cdot \sin\theta_k$$
$$\cdot \cos\xi_k + \cos\sigma_k \cdot \sin\xi_k)$$

$$+ s_{zk}(\cos\sigma_k \cdot \sin\theta_k \cdot \cos\xi_k$$
$$- \sin\sigma_k \cdot \sin\xi_k)] \cdot \tau \qquad (8a)$$
$$C_{yk+1} = f_2(C_{xk}, C_{yk}, \varphi_k, \theta_k, \xi_k)$$
$$= C_{yk} + [-s_{xk}\cos\theta_k \cdot \sin\xi_k - s_{yk}(\sin\sigma_k \cdot \sin\theta_k$$
$$\cdot \sin\xi_k - \cos\sigma_k \cdot \cos\xi_k)$$
$$+ s_{zk}(\cos\sigma_k \cdot \sin\theta_k \cdot \sin\xi_k$$
$$- \sin\sigma_k \cdot \cos\xi_k)] \cdot \tau. \qquad (8b)$$

Then, by fixing as zero the vertical speed s_y, the following equations holds:

$$C_{xk+1} = f_1(C_{xk}, C_{yk}, \varphi_k, \theta_k, \xi_k)$$
$$= C_{xk} + [s_{xk} \cos\theta_k \cdot \cos\xi_k + s_{zk}(\cos\sigma_k \cdot \sin\theta_k$$
$$\cdot \cos\xi_k - \sin\sigma_k \cdot \sin\xi_k)] \cdot \tau \qquad (8c)$$
$$C_{yk+1} = f_2(C_{xk}, C_{yk}, \varphi_k, \theta_k, \xi_k)$$
$$= C_{yk} + [-s_{xk} \cos\theta_k \cdot \sin\xi_k + s_{zk}(\cos\sigma_k \cdot \sin\theta_k$$
$$\cdot \sin\xi_k - \sin\sigma_k \cdot \cos\xi_k)] \cdot \tau. \qquad (8d)$$

As the AUV normally navigates over the pipeline with its principal axis parallel to the seabed (i.e., XZ plane) and as this position is kept also during vertical movements, the pan and the
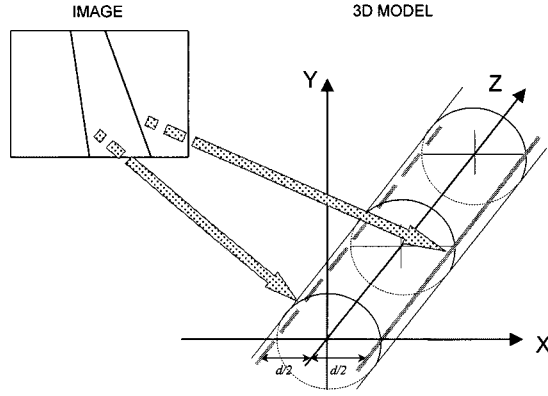
Fig. 9. Three-dimensional (3-D) model of the pipeline and the related projection on the 2-D image.

roll angles are generally close to 0, while the tilt angle is close to the ϕ value (i.e., $\sigma = 0$). Moreover, as the time interval τ is short (e.g., the system acquires several frames per second), the variation of the above angles on the time interval τ is very small. In this way, the evolution of the variables θ, φ, and ξ can be approximately considered constant at all times, i.e., $\dot{\theta} = 0$, $\dot{\varphi} = 0$, and $\dot{\xi} = 0$

$$\theta_{k+1} = f_3(C_{xk}, C_{yk}, \varphi_k, \theta_k, \xi_k) = \theta_k \qquad (9a)$$

$$\varphi_{k+1} = f_4(C_{xk}, C_{yk}, \varphi_k, \theta_k, \xi_k) = \varphi_k \qquad (9b)$$

$$\xi_{k+1} = f_5(C_{xk}, C_{yk}, \varphi_k, \theta_k, \xi_k) = \xi_k. \qquad (9c)$$

Finally, the system state can be summarized in this way:

$$\boldsymbol{x}_{k+1} = \boldsymbol{f}(\boldsymbol{x}_k, \boldsymbol{u}_k) \qquad (10)$$

where $\boldsymbol{f} = (f_1, f_2, f_3, f_4, f_5)^T$ is the state transition function and $\boldsymbol{u}_k = (s_x, s_z, 0, 0, 0)$.

B. Measurement Model

The system measurement requires as input the abscissa values of the points p_1, p_2, p_3, and p_4 selected as observations (see Fig. 4). In order to complete the design of the EKF, a relationship (*measurement model*) between the state vector $\boldsymbol{x}$ and the measurement vector $\boldsymbol{z}$ should be established

$$\boldsymbol{z_k} = g(\boldsymbol{x_k}) \qquad (11)$$

where $\boldsymbol{z} = (i_1, i_2, i_3, i_4)^T$ and $\boldsymbol{g} = (g_1, g_2, g_3, g_4)^T$.

The determination of the g_i functions results simple if the following approximation is considered:

The pair of lines on the image plane selected as features by the system corresponds to the pair of lines generated by the intersection of the pipeline with the XZ plane (Fig. 9).

This approximation allows to consider the 2-D points p_1, p_2, p_3, and p_4 as the projection into the image plane of the 3-D points P_1, P_2, P_3 and P_4 of the scene

$$p_1 \equiv (i_1, 0) \qquad \leftarrow \quad P_1 \equiv (-d/2, 0, P_{z1})$$

$$p_2 \equiv (i_2, 0) \qquad \leftarrow \quad P_2 \equiv (d/2, 0, P_{z2})$$

$$p_3 \equiv (i_3, h-1) \quad \leftarrow \quad P_3 \equiv (-d/2, 0, P_{z3})$$

$$p_4 \equiv (i_4, h-1) \quad \leftarrow \quad P_4 \equiv (d/2, 0, P_{z4}).$$

The functions g_i are achieved by solving (2b) for the P_{zi} values, and by substituting the obtained results in (2a). In the Appendix, the g_i functions are shown.

C. Filter Equations

In order to take into account the noise affecting both the state and the measurement systems, two random variables, $\boldsymbol{w}_k$ and $\boldsymbol{\eta}_k$, have been considered, i.e.,

$$\boldsymbol{x}_{k+1} = \boldsymbol{f}(\boldsymbol{x}_k, \boldsymbol{u}_k) + \boldsymbol{w}_k \qquad (13a)$$

$$\boldsymbol{z}_k = \boldsymbol{g}(\boldsymbol{x}_k) + \boldsymbol{\eta}_k. \qquad (13b)$$

According to the theory of the Kalman filter [24], the random variables $\boldsymbol{w}_k$ and $\boldsymbol{\eta}_k$ are assumed to be independent and with a Gaussian distribution with zero mean and covariance matrices, $\boldsymbol{Q}$ and $\boldsymbol{R}$, i.e.,

$$\boldsymbol{Q} = \begin{pmatrix} \sigma_{Cx}^2 & 0 & 0 & 0 & 0 \\ 0 & \sigma_{Cy}^2 & 0 & 0 & 0 \\ 0 & 0 & \sigma_{\theta}^2 & 0 & 0 \\ 0 & 0 & 0 & \sigma_{\varphi}^2 & 0 \\ 0 & 0 & 0 & 0 & \sigma_{\xi}^2 \end{pmatrix}$$

and

$$\boldsymbol{R} = \begin{pmatrix} \sigma_{i1}^2 & 0 & 0 & 0 \\ 0 & \sigma_{i2}^2 & 0 & 0 \\ 0 & 0 & \sigma_{i3}^2 & 0 \\ 0 & 0 & 0 & \sigma_{i4}^2 \end{pmatrix}.$$

The EKF model requires linearization of the system by approximating the g functions with Taylor's series expansions, retaining only the first-order terms

$$\boldsymbol{x}_{k+1} = \boldsymbol{F}_k \boldsymbol{x}_k + \boldsymbol{B} \boldsymbol{u}_k + \boldsymbol{w}_k \qquad (14a)$$

$$\boldsymbol{z}_k = \boldsymbol{G}_k \boldsymbol{x}_k + \eta_k \qquad (14b)$$

where the matrix $\boldsymbol{F}$ is a square matrix relating x_k and x_{k+1} in the absence of both noise and forcing function, $\boldsymbol{B}$ is a matrix relating the input control u_k and the system state x_k, and $\boldsymbol{G}$ is a matrix relating the system state and the system measure, i.e., the Jacobean of the function g, $\boldsymbol{G}_k(\boldsymbol{x}_k) = (\partial/\partial \boldsymbol{x}_k)g(\boldsymbol{x}_k)$. The matrix coefficients are calculated at every iteration k of the filter.

$$[\Psi] = \begin{pmatrix} \cos\theta_k \cdot \cos\xi_k & -\sin\sigma_k \cdot \sin\theta_k \cdot \cos\xi_k - \cos\sigma_k \cdot \sin\xi_k & \cos\sigma_k \cdot \sin\theta_k \cdot \cos\xi_k - \sin\sigma_k \cdot \sin\xi_k \\ -\sin\xi_k \cos\theta_k & -\cos\sigma_k \cdot \cos\xi_k + \sin\sigma_k \cdot \sin\theta_k \cdot \sin\xi_k & -\sin\sigma_k \cdot \cos\xi_k + \sin\theta_k \cdot \sin\xi_k \cdot \cos\sigma_k \\ -\sin\theta_k & -\sin\sigma_k \cdot \cos\theta_k & \cos\sigma_k \cdot \cos\theta_k \end{pmatrix} \qquad (12)$$
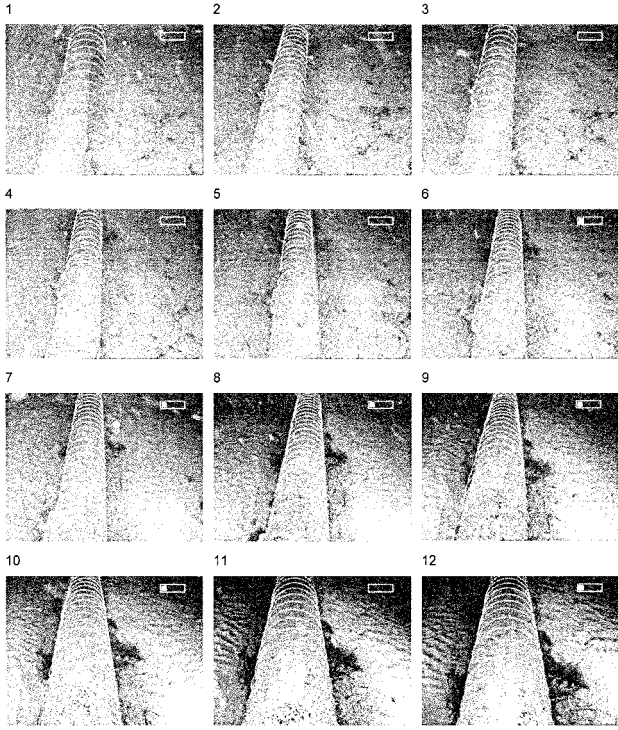
Fig. 10.   Sequence of 12 real underwater images. On each image, the profile of the pipeline has been drawn, reconstructed on the basis of the detected 2-D lines and the information coming from the EKF.

Fig. 11.   Temporal evolution of the state variables C_x, C_y, C_z, θ, and φ.

The prediction equations (*prediction phase*) of the EKF read as follows:

$$\hat{x}_{k+1}^- = F_k \hat{x}_k + B u_k \tag{15a}$$

$$P_{k+1}^- = F_k P_k F_k^T + Q_k \tag{15b}$$

where $\hat{x}_{k+1}^-$ is the status vector predicted by the Kalman filter and P_{k+1}^- is the covariance matrix of the error estimated on the system state before the updating step based on the measurement at the time instant t_k.

The updating equations (*updating phase*) result as follows:

$$K_k = P_k^- G_k^T \left(G_k P_k^- G_k^T + R_k \right)^{-1} \tag{16a}$$

$$\hat{x}_k = \hat{x}_k^- + K \left(z_k - g \left(\hat{x}_k^- \right) \right) \tag{16b}$$

$$P_k = (I - K_k G_k) P_k^- \tag{16c}$$

where K_k is the filter gain at the step k.

D. Absence of Measurements

In the real cases, it can sometimes happen that, due to occlusions or some errors in the process of data (image) acquisition and transmission, some measurements can be lost. In these cases, the EKF cannot reach the performance expected. To overcome this, the method suggested by Chen [24] has been adopted. By starting from the hypothesis of the absence of measurements at time k_0, the state estimate is computed by substituting the optimal prediction $\hat{x}_{k0}^-$ to the unknown optimal estimate $\hat{x}_{k0}$. Equations (15a) and (16b), used to determine $\hat{x}_{k_0+1}^-$ and $\hat{x}_{k0}$, cannot be utilized, as they depend on the unknown measurement

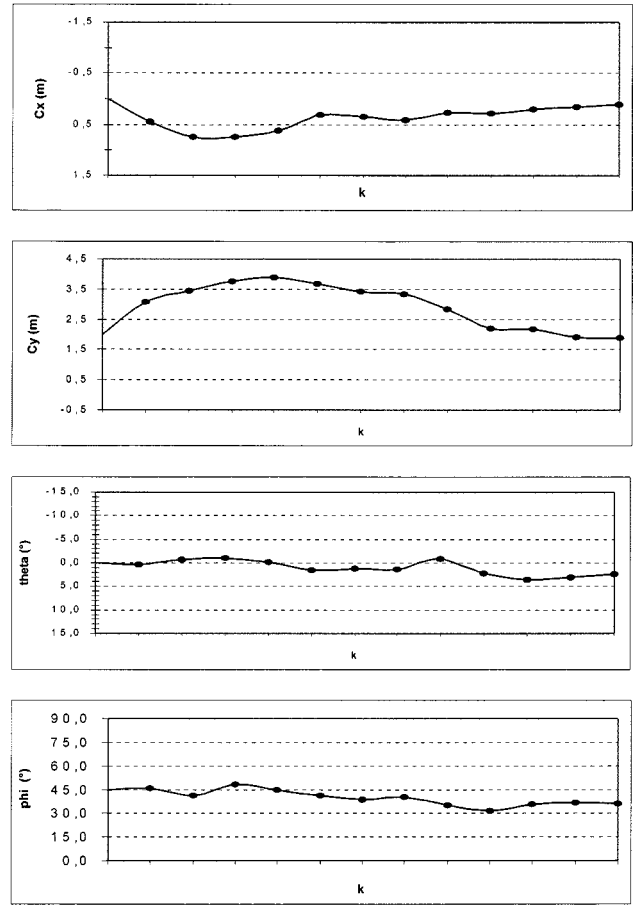z_{k0}. To this end, the following relations have been considered in the case of the absence of measurements:

$$\hat{x}_{k0} = \hat{x}_{k0}^- \tag{17a}$$

$$\hat{x}_{k0+1} = f \left(\hat{x}_{k0}, u_{k0} \right) = f \left(\hat{x}_{k0}^-, u_{k0} \right). \tag{17b}$$

This prediction is the linear estimate of the unknown state, and it is based on all the previous easurements $\{z_1, z_2, \ldots, z_{k0-1}\}$ that are available.

VI. Experimental Results

This section presents the results we obtained by testing the AUV guidance system on long sequences of real underwater images. Tests on AUV missions have been performed on two different environment conditions: 1) absence of down streams (normal conditions) and 2) presence of down streams. Then, some comparisons with other methods used for lane mark detection for autonomous land vehicles have been done.

A. Normal Conditions

Fig. 10 shows 12 images acquired during an AUV mission characterized by low down streams. In these conditions, the hypothesis that the roll angle ξ remains closed to 0 during the AUV mission is assumed. The diameter d of the pipeline is known and constant (i.e., about 1 m) and the intrinsic parameters of the camera (a 8100 COHU compact color charge coupled devices (CCD) camera with focal length varying from 3.9 mm to 85.8

TABLE III
INFORMATION ABOUT THE SYSTEM STATE VARIABLES, THE SYSTEM MEASUREMENTS, AND VALUES OF THE
COVARIANCE MATRIX $P(k)$ OBTAINED ON THE IMAGES IN FIG. 10

		z(k) measure				x*(k)				X*(k)- x'(k)				Covariance matrix P*(k)			
k	Y/N	i1	i2	i3	i4	Cx (m)	Cy (m)	θ (°)	φ (°)	ΔCx (m)	ΔCy (m)	Δθ (°)	Δφ (°)	P*[1][1] σ^2_{Cx}	P*[2][2] σ^2_{Cy}	P*[3][3] σ^2_{θ}	P*[4][4] σ^2_{φ}
0	-	0	0	0	0	0	2	0,0	45,0	0,00	0,00	0,0	0,0	1,00E+00	1,00E+00	3,00E-01	3,00E-01
1	Y	127	172	61	152	0,44	3,07	0,4	45,8	-0,44	-1,07	-0,4	-0,8	4,30E-04	6,16E-03	1,07E-04	1,41E-03
2	Y	137	181	59	158	0,74	3,43	-0,6	41,0	-0,30	-0,36	1,0	4,8	1,36E-03	1,87E-02	1,48E-04	2,19E-03
3	Y	127	189	68	166	0,74	3,75	-1,0	48,1	0,00	-0,32	0,4	-7,1	2,55E-03	2,88E-02	1,96E-04	2,32E-03
4	Y	143	188	84	181	0,62	3 89	-0,2	45,0	0,12	-0,14	-0,8	3,1	2,47E-03	3,38E-02	2,31E-04	3,13E-03
5	Y	150	188	102	209	0,31	3,68	1,6	41,1	0,31	0,21	-1,8	3,9	2,54E-03	4,02E-02	1,82E-04	3,17E-03
6	Y	151	188	95	206	0,35	3,41	1,2	38,5	-0,03	0,27	0,4	2,6	1,99E-03	4,02E-02	1,12E-04	2,77E-03
7	Y	143	188	87	198	0,41	3,34	1,3	40,3	-0,06	0,06	0,0	-1,8	1,88E-03	3,54E-02	1,09E-04	2,37E-03
8	Y	169	204	95	225	0,27	2,84	-0,8	35,3	0,14	0,50	2,0	5,0	1,89E-03	3,19E-02	1,25E-04	2,38E-03
9	Y	146	181	63	217	0,29	2,19	2,3	31,8	-0,02	0,65	-3,1	3,5	1,24E-03	2,29E-02	9,30E-05	1,72E-03
10	Y	127	182	63	225	0 20	2,17	3,5	35,7	0,10	0,01	-1,2	-4,0	9,30E-04	1,05E-02	8,80E-05	1,06E-03
11	Y	128	191	56	242	0,16	1,91	3,1	36,8	0,03	0,27	0,5	-1,1	7,80E-04	9,16E-03	8,80E-05	1,15E-03
12	Y	136	200	67	256	0,11	1,87	2,4	36,2	0,05	0,04	0,7	0,6	5,50E-04	5,80E-03	9,00E-05	9,88E-04

mm) are known: $f = 0.04$ m, $\delta x = 0.0001$ m, $\delta y = 0.0001$ m. The EKF has been initialized with the following covariance matrices R and Q:

$$R = \begin{pmatrix} 16 & 0 & 0 & 0 \\ 0 & 16 & 0 & 0 \\ 0 & 0 & 16 & 0 \\ 0 & 0 & 0 & 16 \end{pmatrix}$$

$$Q = \begin{pmatrix} 0.01 & 0 & 0 & 0 \\ 0 & 0.01 & 0 & 0 \\ 0 & 0 & 0.0012 & 0 \\ 0 & 0 & 0 & 0.087 \end{pmatrix}.$$

The matrix R has been obtained by assuming an initial uncertainty of about 4 pixels on each element of the measurement vector $z = (i_1, i_2, i_3, i_4)^T$, while the matrix Q has been obtained by assuming an uncertainty of about 0.1 m on both C_x and C_y, two degrees on the tilt angle θ and five degrees on the pan angle φ. The x_0 state vector (*initial state*) and to the P_0 matrix (*uncertain on the initial state*) has been initialized as follows:

$$x_0 = \begin{pmatrix} 0.0 \\ 2.0 \\ 0.0 \\ 0.78 \end{pmatrix} \qquad P_0 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0.3 & 0 \\ 0 & 0 & 0 & 0.3 \end{pmatrix}.$$

Starting from these initial values, the acquisition of every new frame, and consequently, the acquisition of new measurements, gives a contribution to the progressive correction of the state estimate until convergence of the EKF. After convergence of the filter, at each iteration, an accurate AUV pose (in terms of the parameters $C_x, C_y, \theta, \varphi$) is obtained. The profile of the pipeline, based on the above estimated parameters, has been drawn on the images of the test sequence (Fig. 10). As already visible from the third and fourth frame, the 3-D pipeline reconstruction matches with the original image; this fact confirms the convergence of the EKF. The temporal evolution of the state variables is shown in the graphs in Fig. 11. Table III shows all the significant data obtained during the experiment.

The graphs in Fig. 12 show the real (measured) and the estimated behavior of the tilt and pan angles of the AUV during
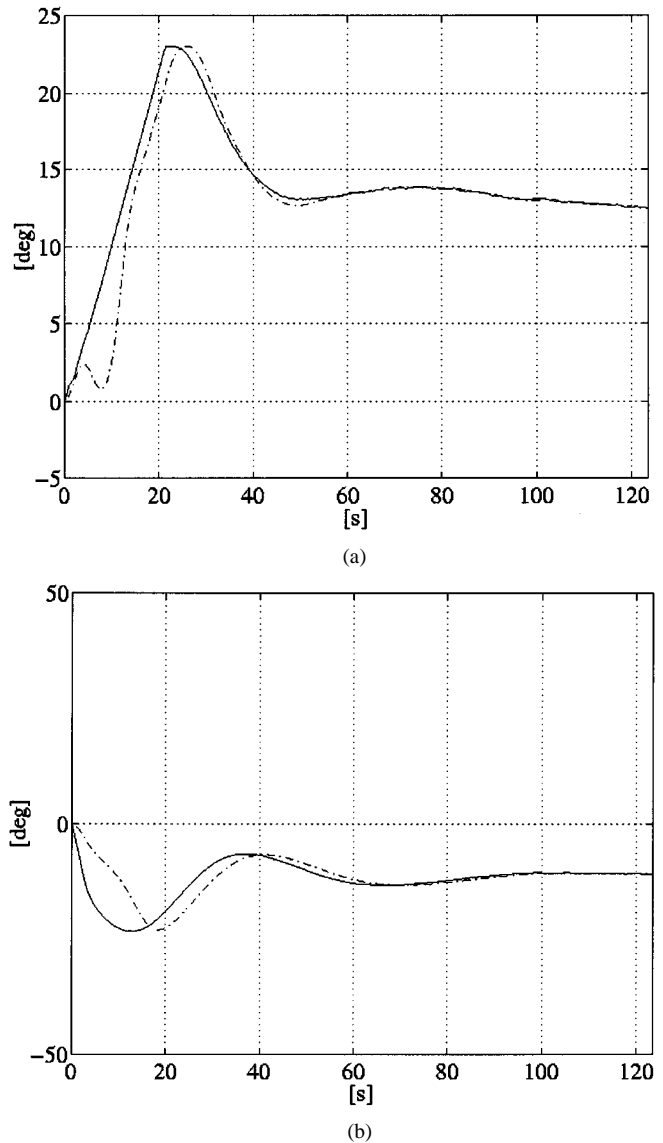

(a)

(b)

Fig. 12. Real (measured) and the estimated behavior of (a) the tilt and (b) the pan angles of the AUV during the first 120 m of the mission. The dashed dot line represents the real behavior and the continue line the estimated one.

ORIGINAL IMAGES   3D RECONSTRUCTION   3D PIPELINE PROFILE DRAWN ON THE ORIGINAL IMAGES

(a)   (b)   (c)

(d)   (e)   (f)

(g)   (h)   (i)

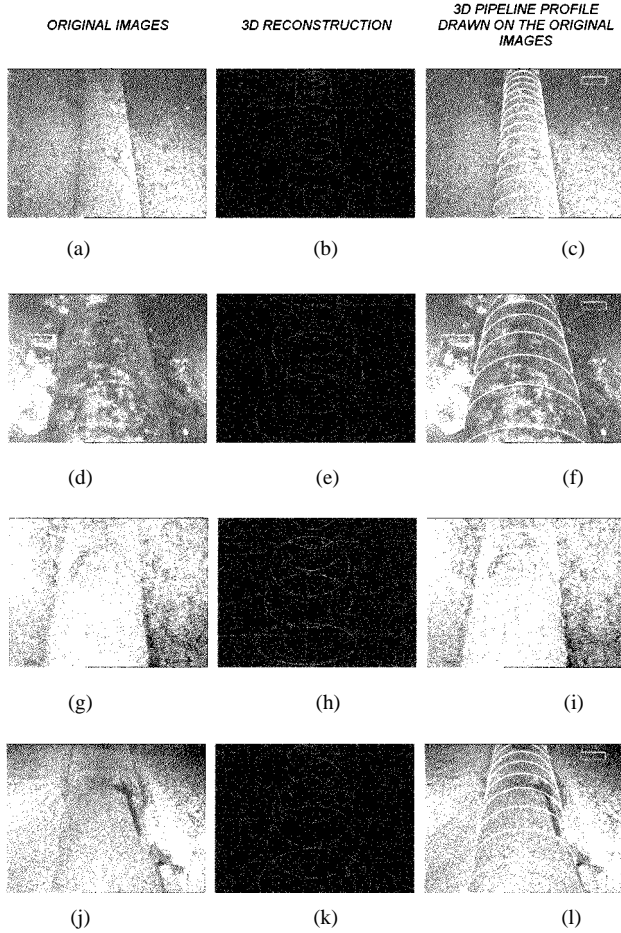(j)   (k)   (l)

Fig. 13. Some complex images taken from different AUV missions. (a), (d), (g), and (j) original images. (b), (e), (h), and (k) 3-D reconstruction of the pipeline. (c), (f), (i), and (l) 3-D profile of each pipeline, superimposed to the original images.
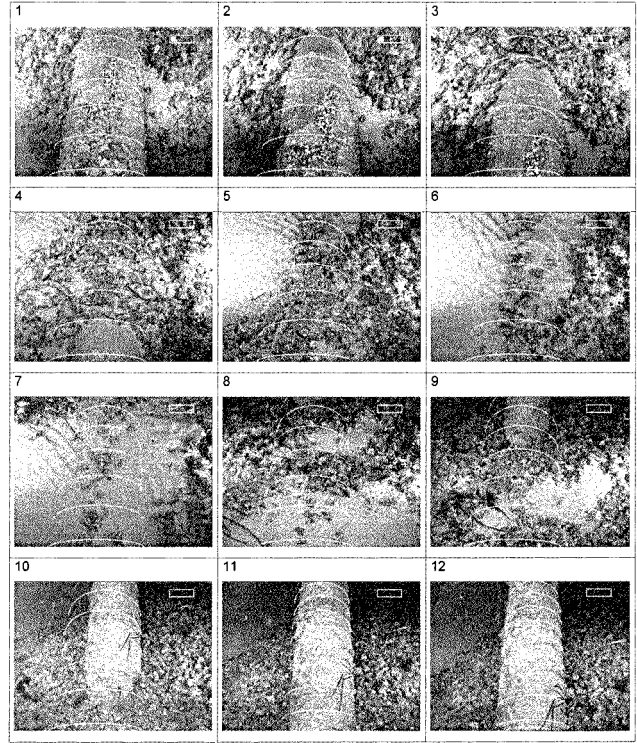

Fig. 14. Some examples of results obtained on a real sequence of underwater images where the pipeline is partially or totally covered sand. The system is able to predict the position of the pipeline also in absence of measurements (frames 4–9).

the first 120 m of the mission (the AUV moves at a constant speed of about 1 m/s). It is worth noting that, after the first 30 s of the mission (i.e., the AUV has covered about 60 m along the pipeline) the proposed system is able to correctly estimate the AUV parameters.

Fig. 13 (first column) shows some complex images taken from different AUV missions. These images have been selected as they represent different working real conditions:

1) different types of sea bottom (e.g., sand, rocky, etc.);
2) presence of unforeseen objects (e.g., garbage, etc.);
3) pipelines with different diameter;
4) pipelines with different types of encrustations and seaweeds.

In the central column, the 3-D reconstruction of the detected pipeline is drawn. The 3-D profile of each pipeline, superimposed to the original image, remarks the validity of the reconstruction (third column).

Real experiments demonstrate also that the absence of measurements for short periods (about five to six frames) cannot compromise the AUV navigation, which continues on the basis of data furnished by inertial sensors. If absence of measurements continues for a long period, the performance of the system depends on the reliability of the inertial sensors: the more accurate

are the data furnished by them, the longer the pursuit can continue until a new measurement is available.

Fig. 14 shows a complex image sequence where the pipeline is partially or totally covered by sand. The estimation of the AUV pose has been performed exclusively on the basis of the system model. This fact has generated an increase of the covariance error of the state variables (Fig. 15), and consequently a loss of accuracy in the AUV pose estimation. Table IV shows all the significant data obtained during the experiments.

B. Presence of Down Streams

Some tests have been performed in presence of a down stream with a speed of about $[0, 0.5, 0]^T$ m/s coming from the lateral (left) side of the AUV. The system parameters have been initialized as in the previous example. In this case, the roll angle ξ has been added to the state vector as a new parameter to be estimated, and it has been initialized by assuming an uncertainty of about $5°$.

Fig. 16(a) shows the real (measured) and the estimated trajectory of the AUV during a mission of 4500 m along a pipeline of diameter of about 1 m at a constant speed of about 1.5 m/s. The dashed dot line represents the real behavior, the continue line the estimated one. We can observe that the maximum error on the variable C_x is about 3 m, and it occurs in the sections where the radius of curvature of the pipeline is high. Fig. 16(b) shows the real (measured) and the estimated behavior of the roll angle of the AUV during the first 750 m of the mission. The dashed dot line represents the real behavior, the continue line the estimated one. It is worth noting that, after the first 50 s of the mission

TABLE IV
INFORMATION ABOUT THE SYSTEM STATE VARIABLES, THE SYSTEM MEASUREMENTS AND VALUES OF THE
COVARIANCE MATRIX $P(k)$ OBTAINED ON THE IMAGES IN FIG. 14

		z(k) measure				x*(k)				X*(k)- x(k)				Covariance matrix P*(k)			
k	Y/N	i1	i2	i3	i4	Cx (m)	Cy (m)	θ (°)	φ (°)	ΔCx (m)	ΔCy (m)	Δθ (°)	Δφ (°)	P*[1][1] σ^2_{Cx}	P*[2][2] σ^2_{Cy}	P*[3][3] σ^2_{θ}	P*[4][4] σ^2_{φ}
1	N	0	0	0	0	0,19	2,31	-2,1	62,7	0,00	0,00	0,0	0,0	4,03E-02	4,41E-02	7,76E-03	9,37E-03
2	N	0	0	0	0	0,19	2,31	-2,1	62,7	0,00	0,00	0,0	0,0	8,03E-02	8,41E-02	1,54E-02	1,70E-02
3	N	0	0	0	0	0,19	2,31	-2,1	62,7	0,00	0,00	0,0	0,0	1,20E-01	1,24E-01	2,30E-02	2,46E-02
4	N	0	0	0	0	0,19	2,31	-2,1	62,7	0,00	0,00	0,0	0,0	1,60E-01	1,64E-01	3,06E-02	3,22E-02
5	N	0	0	0	0	0,19	2,31	-2,1	62,7	0,00	0,00	0,0	0,0	2,00E-01	2,04E-01	3,82E-02	3,98E-02
6	N	0	0	0	0	0,19	2,31	-2,1	62,7	0,00	0,00	0,0	0,0	2,40E-01	2,44E-01	4,58E-02	4,74E-02
7	N	0	0	0	0	0,19	2,31	-2,1	62,7	0,00	0,00	0,0	0,0	2,80E-01	2,84E-01	5,35E-02	5,51E-02
8	N	0	0	0	0	0,19	2,31	-2,1	62,7	0,00	0,00	0,0	0,0	3,20E-01	3,24E-01	6,11E-02	6,27E-02
9	N	0	0	0	0	0,19	2,31	-2,1	62,7	0,00	0,00	0,0	0,0	3,60E-01	3,64E-01	6,87E-02	7,03E-02
10	N	0	0	0	0	0,19	2,31	-2,1	62,7	0,00	0,00	0,0	0,0	4,00E-01	4,04E-01	7,63E-02	7,79E-02
11	Y	156	245	128	260	0,05	2,93	-1,4	61,7	0,14	-0,63	-0,7	1,0	3,40E-04	6,62E-03	1,78E-04	3,18E-03
12	Y	156	246	126	266	-0,01	2,98	-1,3	59,5	0,06	-0,05	-0,1	2,2	5,40E-04	1,21E-02	1,51E-04	3,22E-03

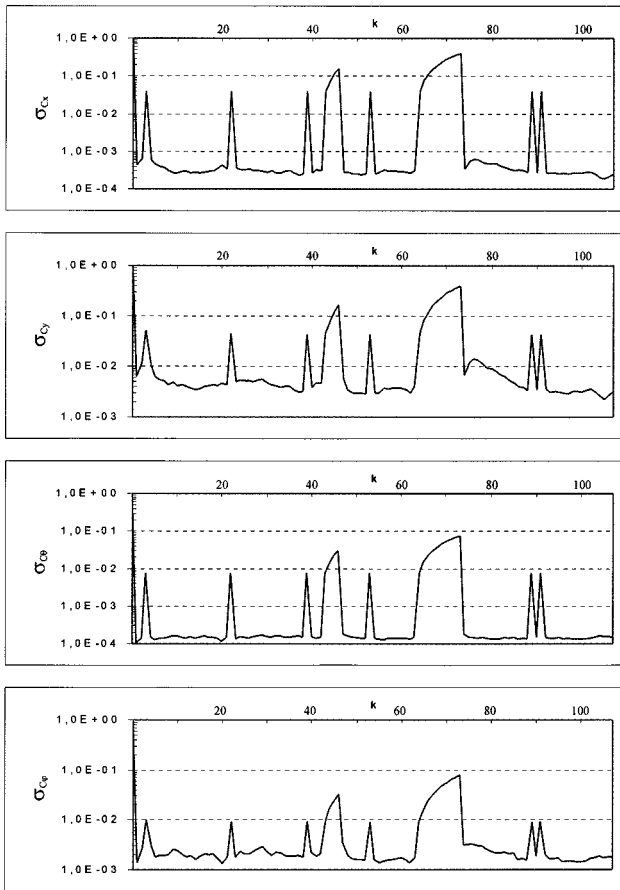Fig. 15. Temporal evolution of the covariance values of the four state variables C_x, C_y, θ, and φ of the system. The peaks correspond to the system states estimated in absence of measures.

(i.e., the AUV has covered about 75 m along the pipeline) the estimate of the roll parameter becomes very accurate.

C. Comparisons With Lane Mark Detection Methods

In order to demonstrate the validity of the proposed approach, a comparison with recent edge-based road lane detection methods for autonomous vehicles has been done.

In particular, the works proposed by 1) Kasprzak et al. [25], 2) Tang and Kasturi [26], and 3) Lai and Yung [27] have been considered. Kasprzak et al. [25] presented a road lane estimation method which classify the edges into three classes (background edges, road edges, and nonroad edges). This method uses the vanishing point to define the horizontal ground. Edges above the horizontal ground are eliminated, while edges below are further classified as road or nonroad edges. Tang and Kasturi [26] developed a runway detection method for aircraft. They used the Canny's edge detector to extract edges from the image. Then, edges are classified into straight lines, circular arcs, and various types of junctions. Recently, Lai and Yung [27] proposed a new lane detection method for visual traffic surveillance applications. First, the Sobel edge detector is applied to extract edges from the input image. Then, these edges are thinned and approximated with straight lines where each line has an associated orientation and length. Straight lines are first clustered according to their orientation using a K-means clustering technique. Camera parameters are used to perform the orientation clustering in 3-D in such a way that only parallel lines should be considered. Finally, from the orientation discrimination result, straight lines are linked and clustered according to their length using the K-means technique again.

Table V shows the obtained results in terms of correct, false, and missed line detection by applying the above methods and the proposed one to a set of about five representative underwater sequences (of about 2×10^3 images) of pipelines taken from different AUV missions. These images represent complex scenes characterized by different illumination conditions and water depth, shadows, presence of sand partially covering the pipeline and other types of objects (e.g., drums, garbage, trestles, etc.). False detection occurs when the system detects a pair of lines which are not really present on the scene or which do not belong to the pipeline borders. Missed detection occurs when the system is unable to detect any pair of lines. The performances of these systems have been computed by comparing the obtained results with ground truth images, i.e., the real images with pipeline borders manually detected. In order to evaluate if the pair of lines indicated by the system is true, the following
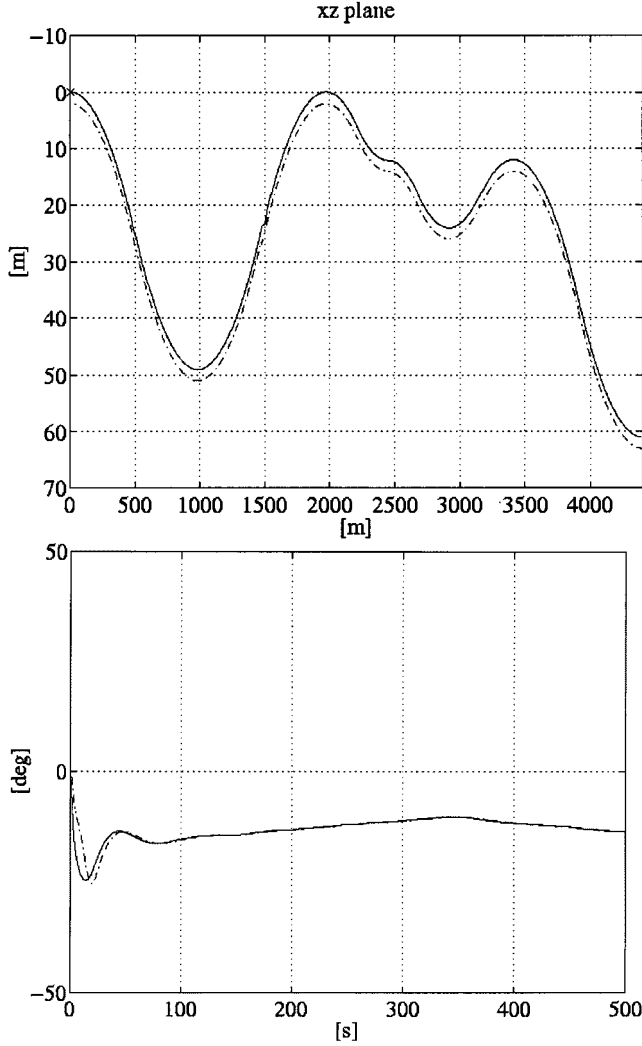
Fig. 16. Real (measured) and estimated trajectory of the AUV during a mission of 4500 meters. The dashed dot line represents the real behavior, the continue line the estimated one.

rule has been applied: a pair of lines is considered correctly detected if, for each line, the Euclidean distance between its midpoint and that of the related line on the ground truth image is lower than 5 pixels and the difference between the orientation angles is lower than 3°.

The values in Table V indicate that the methods proposed by Kasprzak *et al.* [25] and by Tang and Kasturi [26] show low detection accuracy. The first seem to be highly dependent on the vanishing point detection, while the second one seem to be very dependent on the quality of the image, e.g., it enjoys certain success only on images characterized by pipelines with well-defined borders. Both methods work well when the AUV is moving over the pipeline, while they show several limits when the pipeline is partially occluded by sand or other objects. The method proposed by Lai and Yung [27] reaches the 90% of the correct detection, the 4% of false detection, and the 6% of missed detection. It uses the camera parameters (they must be fixed or known for each acquired frame) to estimate into a 3-D reference system the compatibility of the detected pairs of lines, so it works well only for applications where the camera

TABLE V
PERCENTAGE OF CORRECT, FALSE, AND MISSED LINE DETECTION OBTAINED BY APPLYING THE METHODS PROPOSED BY KASPRZAK *et al.* [25], TANG AND KASTURI [26], LAI AND YUNG [27], AND OUR METHOD ON A SET OF ABOUT 2×10^3 REAL IMAGES

	Correct Detection	False Detection	Missed Detection
Kasprzak et al. [25]	77	13	10
Tang and Kasturi [26]	72	16	12
Lai and Yung [27]	90	4	6
Proposed method	92	3	5

is fixed (e.g., visual surveillance of fixed environments). It is worth noting that the use of the compatibility analysis reduces significantly the percentage of wrong line detection. Finally, our method reaches performances similar to these obtained by the method proposed by Lai and Yung [27], but thanks to the EKF model it is able to estimate in real-time the camera parameters (i.e., the AUV pose), so it can be adopted for applications that require the motion of the camera.

VII. CONCLUSION

A robust vision-based system able to drive an AUV during inspections of underwater pipelines has been described. The system has been implemented in C++ language and runs under the Windows NT operating system. Currently, for 256×256 resolution images, it runs at 4 frames/s on a PC with two PENTIUM II Processors at 400 MHz, thus allowing the AUV to move at an average speed of about 1.5 m/s. The proposed system uses a 3-D geometric model of the environment where the AUV operate and an EKF to estimate the AUV pose and orientation in real-time by observing the projection on the image plane of the pipeline edges. Many experiments done on long sequences of real underwater images and comparisons with other methods used for lane mark detection for autonomous land vehicles have confirmed the validity of the proposed system and its efficiency also in case of critical and complex situations, e.g., pipeline partially or totally covered by sand.

APPENDIX

The function g_1 is achieved by solving (2b) for the P_{z1} value

$$P_{z1} = \frac{\frac{d}{2} T_4 - T_7}{T_6} \tag{18}$$

and by substituting the obtained result in (2a)

$$i_1 = h_1(C_x, C_y, \varphi, \theta, \xi)$$

$$= \frac{-\frac{d}{2} T_0 + \left(\dfrac{\frac{d}{2} T_4 - T_7}{T_6}\right) T_2 + T_3}{-\frac{d}{2} T_8 + \left(\dfrac{\frac{d}{2} T_4 - T_7}{T_6}\right) T_{10} + T_{11}}. \tag{19}$$

Analogously, we obtain for the other functions g_2, g_3, and g_4 as shown in (20)–(22) at the top of the next page.

$$i_2 = g_2(C_x, C_y, \varphi, \theta, \xi) = \frac{\frac{d}{2}T_0 + \left(-\frac{\frac{d}{2}T_4 + T_7}{T_6}\right)T_2 + T_3}{\frac{d}{2}T_8 + \left(-\frac{\frac{d}{2}T_4 + T_7}{T_6}\right)T_{10} + T_{11}} \tag{20}$$

$$i_3 = g_3(C_x, C_y, \varphi, \theta, \xi) = \frac{-\frac{d}{2}T_0 + \left(-\frac{\frac{d}{2}T_4 + T_7 - (h-1)\left(-\frac{d}{2}T_8 + T_{11}\right)}{(h-1)T_{10} - T_6}\right)T_2 + T_3}{-\frac{d}{2}T_8 + \left(-\frac{\frac{d}{2}T_4 + T_7 - (h-1)\left(-\frac{d}{2}T_8 + T_{11}\right)}{(h-1)T_{10} - T_6}\right)T_{10} + T_{11}} \tag{21}$$

$$i_4 = g_4(C_x, C_y, \varphi, \theta, \xi) = \frac{\frac{d}{2}T_0 + \left(\frac{\frac{d}{2}T_4 + T_7 - (h-1)\left(\frac{d}{2}T_8 + T_{11}\right)}{(h-1)T_{10} - T_6}\right)T_2 + T_3}{\frac{d}{2}T_8 + \left(\frac{\frac{d}{2}T_4 + T_7 - (h-1)\left(\frac{d}{2}T_8 + T_{11}\right)}{(h-1)T_{10} - T_6}\right)T_{10} + T_{11}}. \tag{22}$$

REFERENCES

[1] F. Aguirre, J. M. Boucher, and J. J. Jacq, "Underwater navigation by video sequence analysis," in *Proc. 10th Int. Conf. Pattern Recognition*, Atlantic City, NJ, June 16–21, 1990, pp. 537–539.

[2] J. O. Hallset, "A vision system for an autonomous underwater vehicle," in *Proc. 11th Int. Conf. Pattern Recognition*, The Hague, The Netherlands, Sept. 1–3, 1992, pp. 320–323.

[3] A. Branca, E. Stella, and A. Distante, "Autonomous navigation of underwater vehicles," in *Proc. Int. Conf. OCEANS*, Nice, France, Sept. 29–Oct. 1, 1998.

[4] J. Kojima *et al.*, "Development of autonomous underwater vehicle AQUA EXPLORER 2 for inspection of underwater cables," in *Proc. Int. Conf. OCEANS*, Halifax, NS, Canada, 1997.

[5] B. Balasuriya *et al.*, "Vision based autonomous underwater vehicle navigation: Underwater cable tracking," in *Proc. Int. Conf. OCEANS*, Halifax, Nova Scotia, 1997, pp. 1418–1424.

[6] B. Balasuriya and T. Ura, "Vison based object following for underwater vehicles," *J. Soc. Naval Arch.*, vol. 180, pp. 663–668, Dec. 1996.

[7] A. Grau, J. Climent, and J. Aranda, "Real time cable tracking using texture descriptors," in *Proc. Int. Conf. OCEANS*, Nice, France, Sept. 29–Oct. 1, 1998.

[8] S. Negahdaripour, X. Xu, and L. Jin, "Direct estimation of motion from sea floor images for automatic station-keeping of submersible," *IEEE J. Oceanic Eng.*, vol. 24, pp. 370–382, July 1999.

[9] G. L. Foresti *et al.*, "A voting-based approach for fast object recognition in underwater acoustic images," *IEEE J. Oceanic Eng.*, vol. 22, pp. 57–65, Jan. 1997.

[10] S. Negahdaripour, C. H. Yu, and A. Shokrollahi, "Recovering shape and motion from undersea images," *IEEE J. Oceanic Eng.*, vol. 15, pp. 189–198, July 1990.

[11] R. L. Marks, S. M. Rock, and M. J. Lee, "Real-time video mosaicking of the ocean floor," *IEEE J. Oceanic Eng.*, vol. 20, pp. 229–241, July 1995.

[12] S. McMillan, D. E. Orin, and R. B. McGhee, "Efficient dynamic simulation of an underwater vehicle with a robotic manipulator," *IEEE Trans. Syst., Man, Cybern.*, vol. 25, pp. 1194–1206, Aug. 1995.

[13] N. H. Yang and C. Ye, "An intelligent mobile vehicle navigator based on fuzzy logic and reinforcement learning," *IEEE Trans. Syst., Man, Cybern. B*, vol. 29, pp. 314–321, Apr. 1999.

[14] E. D. Dickmanns, B. Mysliwetz, and T. Christians, "An integrated spatial–temporal approach to automatic visual guidance of autonomous vehicles," *IEEE Trans. Syst., Man, Cybern.*, vol. 20, pp. 1273–1284, June 1990.

[15] G. L. Foresti *et al.*, "A distributed approach to 3-D road scene recognition," *IEEE Trans. Vehi. Technol.*, vol. 43, pp. 389–406, Feb. 1994.

[16] G. L. Foresti, S. Gentili, and M. Zampato, "A vision-based system for autonomous underwater vehicle navigation," in *Proc. Int. Conf. OCEANS*, Nice, France, Sept. 29–Oct. 1, 1998, pp. 195–199.

[17] P. Zingaretti and S. M. Zanoli, "Robust real-time detection of an underwater pipeline," *J. Eng. Applicat. Artif. Intell.*, vol. 11, pp. 257–268, Apr. 1998.

[18] D. Koller, K. Daniilidis, and H. Nagel, "Model-based object tracking in monocular image sequences of road traffic scenes," *Int. J. Comput. Vision*, vol. 10, pp. 257–281, 1993.

[19] G. L. Foresti, "A real-time Hough-based method for segment detection in complex multisensor images," *J. Real Time Imag.*, vol. 6, pp. 93–111, 2000.

[20] O. Faugeras, *3-D Computer Vision: A Geometric Viewpoint*. Cambridge, MA: MIT Press, 1993.

[21] Z. Zhang and O. D. Faugeras, "Three-dimensional motion computation and object segmentation in a long sequence of stereo frames," *Int. J. Comput. Vision*, vol. 7, pp. 211–241, 1992.

[22] R. Y. Tsai, "Techniques for calibration of the scale factor and image center for high accuracy 3-D machine vision metrology," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 10, pp. 713–720, Oct. 1988.

[23] J. Canny, "A computational approach to edge detection," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 8, pp. 679–698, June 1986.

[24] G. Chen, "A simple treatment for suboptimal Kalman filtering in case of data missing," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 26, pp. 413–415, Feb. 1990.

[25] W. Kasprzak, H. Niemman, and D. Wetzel, "Adaptive road parameter estimation in monocular image sequences," in *Proc. Fifth British Machine Vision Conf.* York, U.K., Sept. 13–16, 1994, pp. 691–700.

[26] Y. L. Tang and R. Kasturi, "Runway detection in an image sequence," *Proc. SPIE*, vol. 2421, pp. 181–190, 1995.

[27] A. H. S. Lai and N. H. C. Yung, "Lane detection by orientation and length discrimination," *IEEE Trans. Syst., Man, Cybern. B*, vol. 30, pp. 539–548, Aug. 2000.

Gian Luca Foresti (S'93–SM'01) was born in Savona, Italy, in 1965. He received the Laurea degree (cum laude) in electronic engineering and the Ph.D. degree in computer vision and signal processing from the University of Genoa, Italy, in 1990 and 1994, respectively. His Ph.D. thesis dealt with distributed systems for analysis and interpretation of real image sequences.

In 1994, he was Visiting Professor at the University of Trento, Italy. Currently, he is an Assistant Professor in the Department of Computer Science (DIMI), University of Udine, Italy. Immediately after receiving the Laurea degree, he worked with the Department of Biophysical and Electronic Engineering (DIBE), University of Genoa, in the area of computer vision, image processing, and image understanding. His main interests involve artificial neural networks, multisensor data fusion, computer vision and image processing, and multimedia databases. He has proposed techniques that have found applications in the following fields: automatic video-based systems for surveillance and monitoring of outdoor environments (e.g., underground stations, railway lines, motorways, etc.), vision systems for autonomous vehicle driving and/or road traffic control, and 3-D scene interpretation and reconstruction. He is author or co-author of more than 100 papers published in international journals and conferences. He has co-organized several special sessions on video-based surveillance systems at international conferences. He has been Guest Editor of two special issues of international journals on the same topic (*Real Time Imaging Journal* and the PROCEEDINGS OF THE IEEE), and he has been co-editor of two books on multimedia video-based surveillance systems, *Multimedia Video-Based Surveillance Systems: From User Requirements to Research Solutions* (Kluwer: Boston, MA, 2000) and *Vision and Communications in the Third Generation of Advanced Video-Based Surveillance Systems: New Trends and Experiments* (Kluwer, Boston, MA, 2001).

Dr. Foresti is a member of IAPR. He has also been involved as an evaluator of project proposals in some CEC programs [MAST III (1995), Long Term Research (1995–1998), and BRITE-EURAM-CRAFT (1996)].