A Vision-Based Sensor of Position and Rate for Path Tracking of Autonomous Underwater Vehicles in Environments with Regular Patterns

Mario A. Jordán, Carlos Berger and Jorge L. Bustamante

Abstract— **In this paper a vision-based algorithm for online estimation of position control errors in a guidance system for path tracking is presented. The algorithm uses techniques of pattern recognition, however with a high degree of morphological simplifications. The algorithm detects the path line stretch which is displayed in the actual frame, and estimates the slope of the line and its midpoint. Moreover, it identifies a confidence zone in where the line stretch would be certainly. *Ad-hoc* experiments with a subaquatic vehicle in an test tank show the features of the algorithm proposed under strong conditions of light perturbation and cloudy water.**

Keywords— **Vision-based sensor - Path tracking - Autonomous underwater vehicles - Pattern recognition - Guidance systems.**

I. Introduction

The interest for subaquatic vehicles applications has been continuously increasing. Applications embrace not only the classical ones of the off-shore industry but also have begun to get widely into oceanographic and scientific missions [Narimani *et. al*, 2009].

One of the most relevant technical application concerns the path tracking of pipelines on the sea bottom with the goal of inspection [Inzartev, 2009, Wang *et. al*, 2007]. In the general case control actions are constructed by the navigation system from sonar signals and/or vision. Particularly, subaquatic vision-based systems are generally cheaper and provide superior rates in the signal processing than sonar systems.

With an onboard camera in the vehicle, the relative image motion can give useful information of tracking path errors in real time. However the visual conditions of the environments (often cloudy and muddy water) put restrictions to the quality of measures and the feasibility of gathering these [Sattar and G. Dudek, 2006]. Therefore, the design of vision-based guidance systems with blurred vision come into consideration in scenarios when the altitude to bottom is relatively small. This in turn restricts the applications to a particular class of vehicles that possesses the ability to react rapidly in order to bypass potential obstacles on the sea floor.

Corresponding Author: Mario A. Jordán: E-mail: mjordan@criba.edu.ar. Address: CCT-CONICET. Florida 8000, B8000FWB, Bahía Blanca, ARGENTINA

Mario Jordán and Jorge Bustamante are with the Argentine Institute of Oceanography (IADO-CONICET) and Dto. de Ingeniería Eléctrica y de Computadoras (DIEC-UNS), Bahía Blanca, Argentina. Carlos Berger is with the Argentine Institute of Oceanography (IADO-CONICET).

From a sensorial viewpoint, vision-based systems provide frames with features to be extracted at regular intervals of times. So their integration is natural in digital control systems. The sampling time determined by the image processing rate may negatively influence the stability of the control system in case this is not sufficiently small. So, for large cruise velocities the image processing has to be efficient and timesaving to accomplish stability and performance demands.

In any described scenario above, an intelligent vision sensor to abstract motion properties from moving images on the sea floor could be significant for autonomous navigation and in any case could be a complement of other navigation sensors [Sattar *et. al*, 2008, Huster *et. al*, 2002].

Image-based sensors have the ability to provided information of position and velocity by image processing with certain index of confidence. A combination of techniques like pattern recognition and optical flow is a common way to build up this kind of sensors [van der Zwaan and Santos-Victor, 2001, Caimi *et. al*, 2008].

One of the problems to design robust functions of visual sensors is commonly the growth of uncertainty in the estimations under extreme situations in subaquatic environments. The usually permanent changes of light intensity which make the information retrieval from image motion very difficult (for instance, caustic waves on the sea floor). Also blurred scenes despite adequate lighting are common in fluid medium so that the object recognition can not easily be sustained with continuity like in the aerial medium. Finally the sporadic lost of frames in the context of control must at least be detected in order to resort to predictions upon past states or, in the worst case, to break the vehicle motion [Garcia-Aracil *et. al*,2009].

In this paper, the design of a vision-based sensor for simultaneous spacial and kinematics measurements is focused. All the difficulties in the image processing emerging from a blurred subaquatic vision are taken into account for attaining a robust estimation. So we develop independently two algorithms for position and velocity and concatenate them interactively for navigation in the context of path tracking of lines. Experiments are also set up for this goal with the evaluation of results in the short-term perspective of including this sensor in autonomous navigation vehicles.

II. Vehicle Dynamics

The dynamics of the underwater vehicle is [Fossen, 1994]

$$M\dot{\mathbf{v}} = -C(\mathbf{v})\mathbf{v} - D(|\mathbf{v}|)\mathbf{v} + \mathbf{g}(\boldsymbol{\eta}) + \boldsymbol{\tau}_t \quad (1)$$

$$\dot{\boldsymbol{\eta}} = J(\boldsymbol{\eta})(\mathbf{v} + \mathbf{v}_c), \quad (2)$$

with $\boldsymbol{\eta}$ being defined as the generalized position in some earth-fixed frame, $\mathbf{v}$ the generalized velocity vector in a vehicle-fixed frame, $\mathbf{v}_c$ is the current flow rate in vector form. Also there are system matrices, namely: the inertia matrix M, the Coriolis matrix $C(\mathbf{v})$ and drag matrix $D(|\mathbf{v}|)$. Besides, $\mathbf{g}$ is the net buoyancy force and $\boldsymbol{\tau}_t$ the generalized force of the thrusters.

For autonomous vehicles like there are focusing here, $\boldsymbol{\eta}$ is composed by significant and less important modes. The first set contains the position variables x, y, z and the yaw angle ψ. The second set contains the roll θ and pitch ϕ angles. These former modes are inherently damped or automatically regulated independently of the control goal of path tracking. Also the altitude z is regulated by autopilot to a fixed value.

III. Path tracking with vision servoing

In the effort of tracking a line, actually only the relative position of the vehicle to the line is significant. This includes the location of the vehicle over the line and its tangential orientation to it. Moreover the quantification of the motion is necessary to push the vehicle at desired cruise velocities.

In our goal, the visual servoing is accomplished by an inboard camera that has to provide the relative position x, y, the course ψ, the cruise velocity v and the rotation rate $\dot{\psi}$. All these estimations must be provided by the image moving (the so-called egomotion).

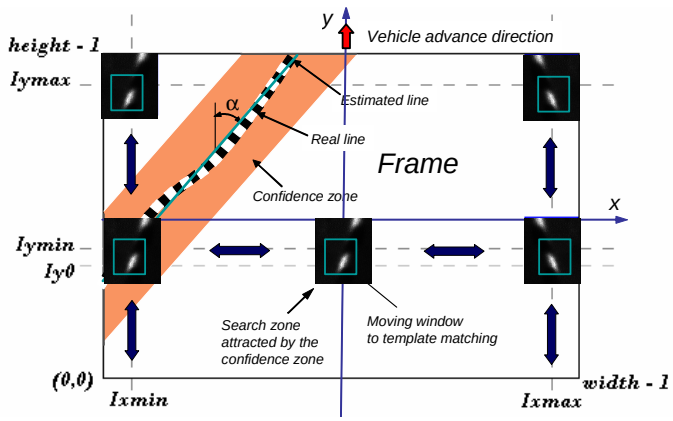


Fig. 1 - Pattern matching procedure when the estimated line moves crossing completely the frame or partially sidelong

IV. Spatial feature abstraction

Imagine a frame containing a line with regular patterns like in Fig. 1. The line is first recognized (i.e., detected) and their parameters: the slant α (estimation of the course ψ) and the coordinates of the line stretch which is embraced in the frame are on-line identified (see Fig. 1). The coordinates are calculated with respect to the frame center.

These estimations have to be provided to the control system as geometric path errors and the controller would attempt after that to guide the vehicle by placing the line vertically at the center of the image in the next steps of the guidance.

The features of the line are extracted by an algorithm that runs whenever a frame arrives the estimator.

To this end, we develop in this work a method that performs a number of systematic operations and transformations over the image. These are summarized in Fig. 2. The reader can find a mathematical description in [Jordán *et. al*, 2010a].

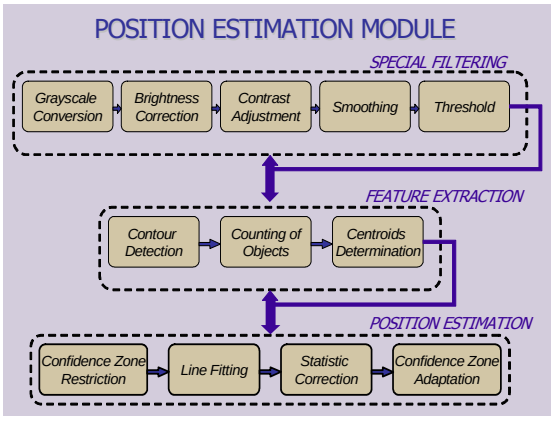


Fig. 2 - Algorithm to estimation of geometric path errors

The module contains tasks that are divided in three groups which are executed one stage at a time.

The first group attempts to simplify the frame morphologically in order to enhanced the contrast between patterns and environment.

Usually, an RGB-format image is received and transformed into a grey scale and then the brightness highlighted. Commonly pattern appear as white nonconnected sets (white speckles). Also the differentiation of exogenous speckles from the sought-after ones in a binary image must be done morphologically by filtering process. After two successive erosions, a dilation operation is carried out in order to recuperate the size of the original speckles that are not erased. Small speckles are generally produced by caustic waves on the floor in shallow water.

The second group is advocated to image feature extraction.

The resulting binary image is further simplified disallowing unimportant information. Here the contours are next identified. The best method to this end is based on the morphology of the speckles. At this stage small contours (little regions that pass the previous erosions) are eliminated. A further reduction of information is attained by replacing the contours by their centroids.

So we arrive to the third group that supplies the position parameters of the line, also the geometric path errors in position and alignment.

From all calculated centroids there are only those taken into account that are in the so-called confidence zone. This zone is a band delimited by two parallel lines. The outtake of centroids is performed by an ad-hoc flag. So the slope of the estimated line (referred to as α) and the midpoint of the visible segment of the line determine the parameters of the estimated line. Additionally, with the sake of reducing the number of wrong estimations and giving the algorithm certain continuity in the identification of the line, some statistical modifications are introduced. The modification takes into account past results from previous cycles employing standard methods for instance, of the affine data averaging or of the forgetting factor.

It is important to adapt permanently the confidence zone, because the conditions of the image are commonly changeable. One possibility is to adjust the width in the proportion of the speckle areas, or alternatively, the simple count of the pixels conforming any contour gives idea of the order of magnitude of the change of the width. Another possibility could be to let the border lines of the confidence zone be not parallel.

When the group tasks are fully executed, the algorithm wait for the next frame.

V. Kinematics feature abstraction

Now we present our approach to estimate kinematics properties from egomotion. Generally spoken, the previous algorithm to estimate relative position of the line is taken as basis to work up to a new combined procedure.

Once again, we will emphasize the requirements of real-time calculations so as to estimate measures on line for a guidance control system. This goal is hardly achieved when parameters have to be extracted from image processing on-line. In this sense, we will take full advantage of the previous algorithm. The readers are referred to [Jordán *et. al*, 2010a, Jordán *et. al*, 2010b] for mathematical details of two different algorithms, respectively.

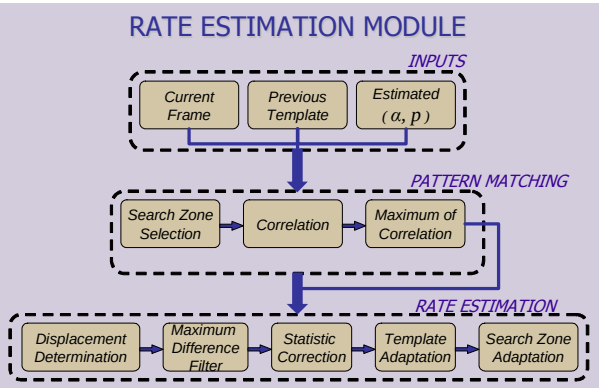


Fig. 3 - Algorithm to estimation of translation velocity and sway rate

Clearly, the recognition of motion properties is related to techniques of optical flow. Since in general common optical flow techniques are markedly time-consuming and considered to be not fit for control systems with rapid response, we will develop instead a simple but robust correlation-based method to this end.

The key idea is to define a search zone over the patterned line and overlapping the confidence zone as close as possible (see Fig. 1). Inside it, there is a small moving window in where the pattern matching of a template will take place. This window slides slightly everywhere in all directions to pattern matching, but always enclosed in the search zone.

The module for kinematics estimation is illustrated in Fig. 3. Similarly it contains three groups of tasks.

In the first group the template is located at the confidence zone according to possible displacement regions. The search zone is defined for the next frame around template location. Here the estimated position parameters of the line are required.

The second group performs basically a match based in correlation of the previous template with contents in the moving window inside the search zone. This group is the core of the module.

Let us suppose the search region stays overlapping the confidence zone and a template with a pattern was selected in the moving window a sampling time before. This template has specific coordinates (x_{t-1}, y_{t-1}). After Δt seconds a new frame enter the algorithm to be considered at the present sampling time. So, the search zone slides eventually a bit so as to overlap the new position of the confidence zone. Then the moving window inside the search zone begins the matching process by correlation between the previous template and the window contents. Perturbing the template coordinates for the moving window in all directions, a maximum of the correlation is searched for. This maximum occurs by successful matching, say at coordinates (x_t, y_t). From this value on, an actual value of the velocity is estimated as the incremental quotient $\hat{v}_t = \frac{(x_t, y_t) - (x_{t-1}, y_{t-1})}{\Delta t}$.

The third group determines the displacement of the template as the slid point that produces the maximum correlation value.

It is worth noticing that the estimation so far suffers from quantification errors and false perspective appreciation which are proper from 2D image approaches. In fact, effects of radial and tangential distortions are not taken into account. So the evolution of the estimates may be irregular and a filtering is needed. This is accomplished in two ways. On one side, a maximum-difference filter is used to reduce high fluctuations that appears in some cases concerning images with elevated noise level (change of contrast, caustic waves, etc.). On the other side, a strong perturbation may be induced when the estimated position of the line results incorrect. In such a cases, the calculated value $\hat{v}_t$ is averaged with the previous value $\hat{v}_{t-1}$.

Moreover, this filtering is not sufficient enough to smooth the time evolution of $\hat{v}_t$ according to the expected quality for control purposes. Therefore a second filter with a forgetting factor allows a much more soft time evolution.

It is to remark that the estimations are displacements expressed in pixels, so the velocity is measured in units of pixels/frame. Knowing the frames per second of the camera

, one can obtain velocity in units of pixels/second.

Finally, a commissioning face will allow us to determine parameters to tune the algorithm in order to give the vehicle velocity in the physical units meters per second. Among these parameters are the altitude, the tilt angle of the camera and proper parameters of the image processing. Moreover, it is supposed that an autonomous vehicle possesses an autopilot to regulate altitude as well as to select a proper camera tilt angle according to a convenient shortsighted or large-sight vision of the bottom scene. Due to space limitations in the paper we will not address the calibration procedures here.

VI. Supervision algorithm

It is quite important for successful vision-based control applications that the sensor can give the controller a certain confidence about the quality of measures.

One of the problems to be tackled in our goal is to share the search zone in the confidence zone when, for instance, the line is moving fast in the vision frame. This can occur, above all, when perturbations affect the control system and control actions do not avoid that the line slides rudely from the frame center. Obviously, this problem is quite alien to the vision-based sensor. However it is aimed by design to confer the search procedure such robust properties that makes the sensor reliable. To this goal the search zone is continuously being attracted by the confidence zone. In that way, if the line is visible in the frame through the confidence zone, a pattern matching is possible.

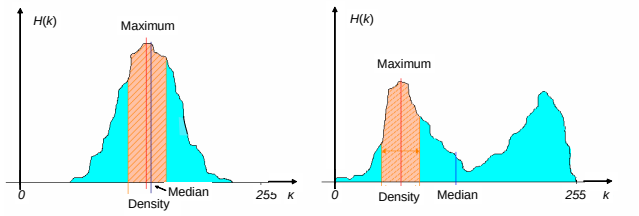


Fig. 4 - Left: histogram of sea floor. Right: histogram of pattern

In customary operations, pattern estimations may fail: *a*) due to sporadic bad image quality or *b*) simply when patterns actually ran over the frame. In order to identify such conditions, supervision is needed. This should provide a flag to stop position and rate estimations in the two abnormal conditions, and also to continue estimating when such conditions have disappeared (case *a*) or have been remedied by control (case *b*).

To this end, a new module that performs histogram-based operations is included in the whole algorithm. It works upon the observed fact that both images of patterns and of the sea floor alone, have typically distinct statistic properties when contrasted. For instance, as seen in Fig. 4, the histogram of a pattern is typical bimodal (two local maximums) and well extended over the range [0, 255] due the high contrast between black and white zones. On the other hand, the environment on the bottom has a narrow range of intensities instead (typical one maximum) with poor contrast. This last scenario can also appear by blurred image even when the pattern is in focal plane. Accordingly these differences will be exploited here for detecting pattern presence/absence or bad/good measure quality.

The algorithm of supervision is described in Fig. 5, which is self explained together with Fig. 4.

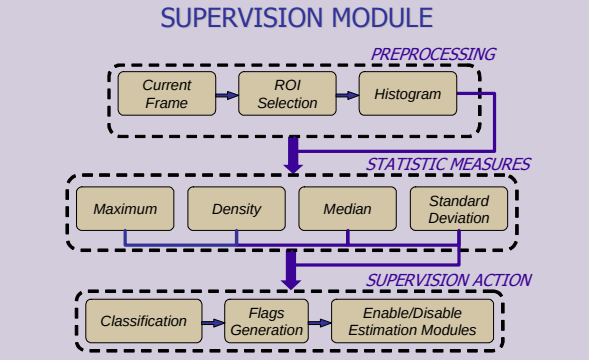


Fig. 5 - Supervision module for the sensor

Finally, the full algorithm to vision-based sensor is depicted in Fig. 5. For vision servoing, the estimations of position and rate are feedback to the controller. Here ξ means the width of the confidence zone.

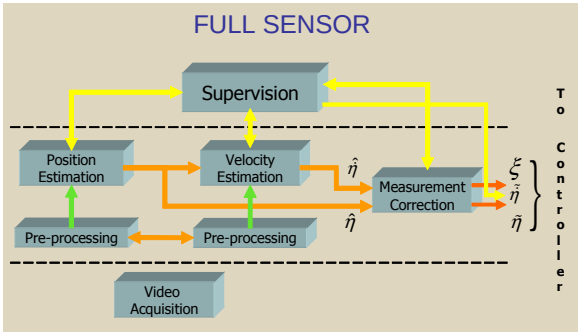


Fig. 6 - Full algorithm to vision-based estimation of position and rate

VII. Experiments

A. Setups

To test the implementation of the vision-based approach a series of experiments in a test tank were set up. These consisted in the employment of a subaquatic vehicle (AUV prototype, see Fig. 7 and 9) that can navigate by telecontrol, following a visible path on the tank bottom. The vehicle possesses a wireless camera onboard that transmit the images during the egomotion (see Fig. 7). For the path tracking problem a patterned line (see Fig. 8) with a metallic wire inside to bent it plastically according to any circuital form (see Fig. 9). In many performed experiments there were perturbations of wind producing currents from some slanted direction, or of caustic waves on the floor produced by the sun rays at midday when trespassing an undulating free surface.

One fact to be highlighted in the work is the altitude of the vehicle with respect to the bottom which influences

directly the focal distance. Motived by the different states of water transparency, we have mounted the vehicle motion at two different constant altitudes. These in turn have forced us to set the experiments up in a shortsighted and longsighted vision-based navigation.

B. Control goal

The path tracking was performed according to the navigation along a circuital path. An ideal control would have to maintain the line totally vertical crossing the center of the frame at all times. A good practical control however attempts to achieve this goal despite perturbations revealing always path following errors. As we are not evaluating the control but the performance of a sensor instead, both large or small path errors are considered favorable in the evaluation. Indeed the line stretch sometimes went sidelong through the lateral borders or inclusive had temporally disappeared.

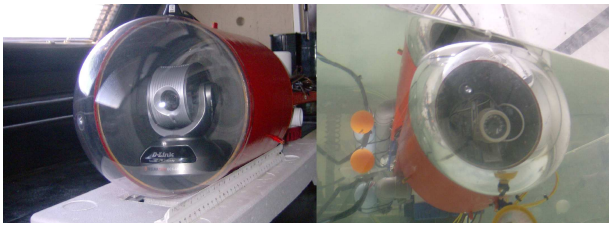
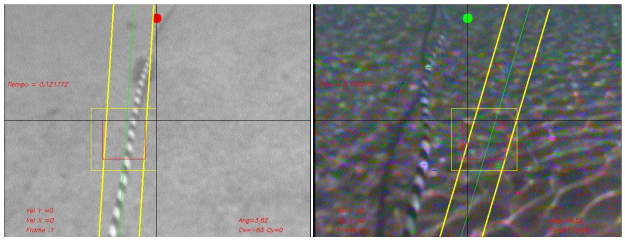


Fig. 7 - Wireless cameras in UV. Link: pan-tilt-zoom regulable camera. Right: Analog camera with digitalizer



Weather: cloudy and windless Weather: sunny and windy

Fig. 8 - Reference line in water

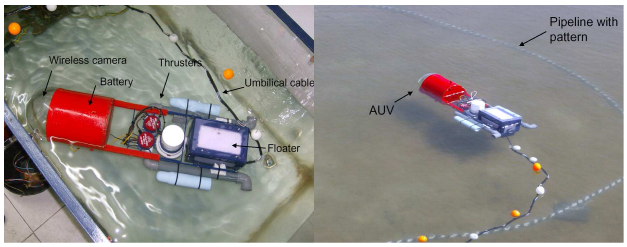


Fig. 9 - AUV navigation in tank with reference line

C. Position and rate identification

Several experiments were led with the described setups. We will illustrate one of them in Fig. 10. Here some cycles are selected to show the results of the modules in normal and critical situations of both blurred frames and coarse line displacements due to control. To the left and to the right of the figure, we can see photograms of the line in a shortsighted and a longsighted orientation of the camera, respectively.

The frames are illustrated with augmented reality during the screening of the results in real-time. Here, the crude image is superimposed with the estimated line, the confidence zone and the search zone for velocity estimation. The rate for frame processing was about 10 (fps) which gives small sampling times for control purposes at relatively large cruise velocities in subaquatic vehicles. However, when combining the position and the kinematic jointly, the measurement rate is reduced in half the time which is also good for the guidance.

We now describe in more details the Fig. 10. The frames to the left correspond to the navigation with shortsighted horizon and to the right with longsighted horizon.

For the shortsighted horizon, in a), the position estimation gave a false line position, and consequently the velocity value was incorrect at this frame. In b), the rapid movement of vehicle produces a lost of the line in the respective frame, thus also here nor a rate estimation neither the line location were possible. Once the line is well detected, the velocity estimation can be realized error free. This occurred in c) where a successful estimation resulted. In d), the estimation was also correct despite the curvature of the line. A vertical displacement restriction had occurred in e) when noted the line moved sidelong to the right border of the frame. This executed the pattern-matching process on the borders. In f)-g) two scenarios are shown related to when the templates turned over to the center of the image and the search zone had begun to slide horizontally for pattern matching.

For the longsighted horizon in Fig. 10, right, different results were obtained. In a) the scenario when the position estimation did not give the right line location is illustrated, but the velocity estimation could be realized anyway. In b), the search zone was attracted by the confidence zone according to rate module, which crossed vertically the frame causing the moving window to search for pattern over the horizontal middle line. In c), a limit case is shown when the confidence zone was moving to the right till to touch the lateral border. Up to here the line began to cut the lateral right border and from this time on the pattern matching was consequently be accomplished sidelong. This allowed continuity in the rate estimation. This last process ended when the control forced the confidence zone to enter again in the frame interior, see d)-e). Here the search zone went down and slid over the middle horizontal again. In f) the pattern went out of the image because of the poor quality of the frame, so the estimated values became spurious until the pattern could be detected in g). The navigation with camera in the longsighted orientation seems to provide a wide vision field and so the line may be much better detected than in the case of shortsighted horizon. Also the control could allow itself to compensate large path errors due to strong perturbations. However this scenario can change drastically by blurred waters in where the altitude

has to be diminished and the camera horizon has to be tuned to a rather shortsighted vision as occurred in our first experimental description.

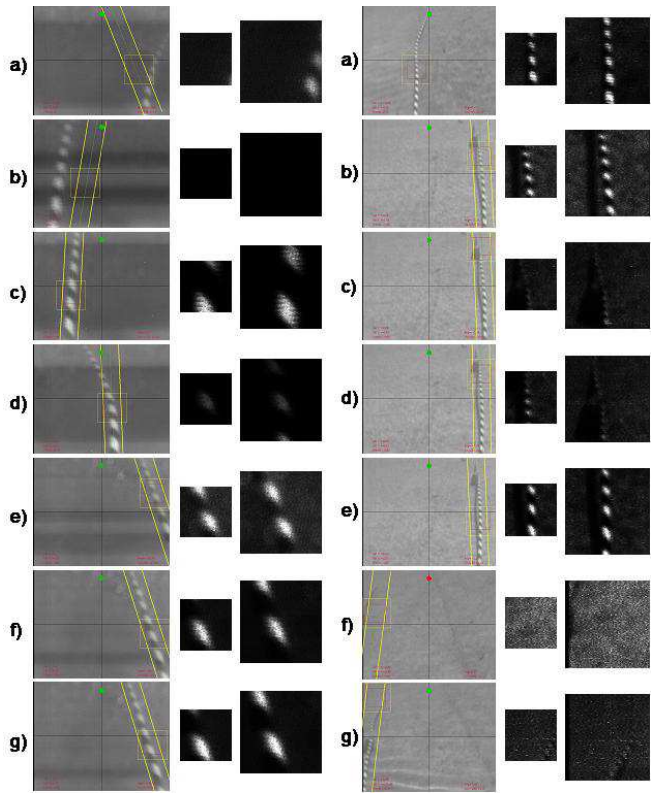


Fig 10 - Selected frames for navigation with shortsighted and longsighted horizons

The result of the supervision process is indicated in the frame as a dot in the upper border. The color green indicates the measures were possible, and the color red indicates that the sensor can not provide any reliable measure.

In Figs. 11 and 12 we will illustrate the evolution of the final conditioned estimated rate vector in their components for the shortsighted horizon (Fig. 11) and for the longsighted horizon (Fig. 12). Moreover, for the sake of clarity, there were described with labels for the cases a) up to g) according to Fig. 10. Also there are windows (in orange color) showing the periods when the pattern matching was accomplished on the lateral borders. Similar indication in windows (in blue color) was carried out for the cases when the line got lost or could not being identified from reasons of poor quality of the image. Clearly, these windows could be located manually after seeing the processing, and they helped us to analyze the flag evolution given by the supervision algorithm.

It can be seen clearly that the continuity of the rate estimation is maintained in the transitions when the search zone moves horizontally and then slides vertically sidelong and vice versa. However, when the line could not be located, the estimation showed signs of instability, becoming oscillatory and irregular. This had occurred only in the longsighted orientation of the camera. In this case, the flag indicating estimation quality performs adequately almost all the time. Wrong matching cases from the main algorithm can not be discriminate with the proposed supervision.

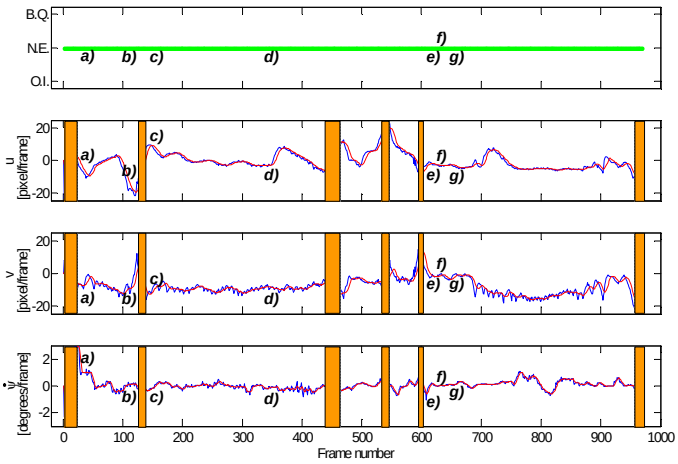


Fig. 11 - Estimation of the advance velocity from the vision-based sensor in a shortsighted-horizon setup

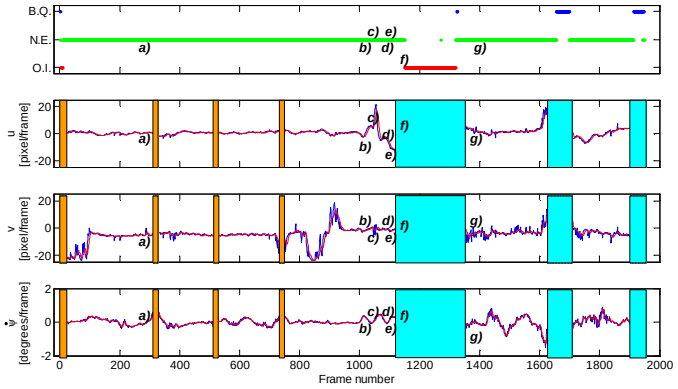


Fig. 12 - Estimation of the advance velocity from the vision-based sensor in a longsighted horizon setup

VIII. Conclusions

This paper was concerned about the design of a vision-based algorithm for on-line estimation of position and rate. The main application is to use this sensor to visual servoing in the guidance of autonomous underwater vehicles for path tracking of underwater lines.

A first identification algorithm employes techniques of pattern recognition with different degrees of morphological operations on the image. The result is the estimation of the line coordinates, its slant and a confidence zone about the line. Additionally, a second algorithm operates on the confidence zone and search for patterns in successive frames to estimate the rate.

A supervision algorithm finally attempts to ensure the continuity of the estimations and detect the lost of the line on the frame. *Ad-hoc* experiments with a subaquatic vehicle and a pattern-shaped line in a test tank had shown the feasibility of our approach in real applications under strong conditions of light perturbations and cloudy water. Here

vision with shortsighted and longsighted horizons were differenced.

REFERENCES

Caimi, F. M., Kocak, D.M., Dalgleish, F., Watson, J. (2008). Underwater imaging and optics: Recent advances. *In OCEANS'08*, Quebec, 1-9.

Fossen, T.I. (1994). *Guidance and control of ocean vehicles*, John Wiley&Sons, New York.

Garcia-Aracil, N.M., Azorin Poveda, J.M., Sabater Navarro, J.M., Perez Vidal, C. and Saltaren Pazmino, R. (2006). Visual Control of robots with changes of visibility in image features. *IEEE Latin America Transactions*, 4 (1), 27-33.

Huster, A., Frew, E.W., Rock, S.M. (2002). Relative position estimation for AUVs by fusing bearing and inertial rate sensor measurements. *In OCEANS'02,* Mississipi*, USA,* vol. 3, 1863 - 1870.

Inzartev A.V.(Ed.) (2009). *Underwater vehicles*, In-Tech, Vienna, Austria.

Jordán, M.A., Berger, C.E., Bustamante, J.L. and Hansen, S. (2010a). Two Vision-based Algorithms for Image-Properties Extraction in the Path Tracking of *Underwa*ter-Vehicle Navigation. *In XXII Congreso Argentino de Control Automático*, Buenos Aires, Argentina.

Jordán, M.A., Berger, C.E., Bustamante, J.L. and Hansen, S. (2010b). Path Tracking in Underwater Vehicle Navigation - On-Line Evaluation of Guidance Path Errors via Vision Sensor. Accepted to the *49th IEEE Conf. on Decision and Control*. Atlanta, USA.

Narimani, M., Nazem, S. and Loueipour, M. (2009). Robotics vision-based system for an underwater pipeline and cable tracker. *In OCEANS'09*, Bremen, Germany, 1-6.

Sattar, J. and Dudek, G. (2006). On the performance of color tracking algorithms for underwater robots under varying lighting and visibility. *In IEEE Int. Conf. on Robotics and Automation (ICRA 2006)*, Orlando, USA, 3550-3555.

Sattar, J, Dudek, G., Chiu, O., Rekleitis, I., Giguere, P., Mills, A., Plamondon, N., Prahacs, C., Girdhar, Y., Nahon, M. and Lobos, J.-P. (2008). Enabling autonomous capabilities in underwater robotics. *In IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS 2008)*, Nice, France, 3628-3634.

van der Zwaan, S and Santos-Victor, J. (2001). Real-time vision-based station keeping for underwater robots. *In OCEANS'01*, Honolulu, USA, vol.2, 1058-1065.

Wang, Y. and Hussein, I.I. (2007). Cooperative Vision-Based Multi-Vehicle Dynamic Coverage Control for Underwater Applications. *In IEEE Int. Conf. on Control Applications (CCA 2007)*, Singapore, 82-87.