

## DETERMINATION OF APPARENT MOBILE AREAS IN AN IMAGE SEQUENCE FOR UNDERWATER ROBOT NAVIGATION

Patrick BOUTHEMY and Patrick LALANDE

IRISA / INRIA-Rennes

Campus de Beaulieu, 35042 Rennes Cedex, France

### Abstract

This paper deals with the determination of moving areas in an image sequence from the variations in time of the intensity distribution. Beyond this problem we are concerned with an application whose aim is the estimation of the trajectory of an underwater vehicle by means of a video sensor on the robot pointing to the sea bottom. The trajectory will be piecewisely reconstructed by tracking a series of appropriate elements, i.e. apparent moving areas, throughout the image sequence. Contextual information is explicitly introduced in the moving-area detection-and-recovery process, which is stated as a statistical labeling one. The label field is modeled as a Markovian field using Gibbs distributions defined on a spatio-temporal neighborhood system. A solution to this labeling problem is formulated according to the maximum a posteriori (MAP) criterion. We have adopted a deterministic iterative algorithm to optimize the criterion at hand. Experiments with an underwater video image sequence are presented.

### INTRODUCTION

This study addresses a spatio-temporal analysis problem which is relevant to a project whose aim is to realize the trajectory estimation of an underwater vehicle equipped with a video sensor, [1]. This camera points to the bottom and it is maintained vertically. The use of complementary data supplying height and course information allows to restrict the problem to the measurement of 2D displacements in the image plane. The trajectory will be piecewisely reconstructed by tracking a series of appropriate elements, i.e. apparent moving areas, throughout the image sequence.

The situation involved by this application does not correspond to the usual one which associates motion detection with change detection, that is static camera and moving objects. Normally the case of a moving camera implies a more complex algorithmic solution to achieve motion-based segmentation and moving object detection as shown in [2]. Nevertheless in the case at hand the apparent motion of image points due to the camera displacement is far from being perceptible for all image points, because of the intrinsic properties of underwater images. Signal-to-noise ratio is low and image content is poorly structured and often uniform. Thus it is necessary to select the most reliable areas in the image, which are called 'tracers'. These areas must be spatially circumscribed in the image, and they must show a perceptible apparent motion which is entirely due to the underwater vehicle displacement. The vehicle trajectory will be piecewisely reconstructed by tracking successive tracers throughout the image sequence (long-term process).

As far as we are concerned, we only deal with the short-term process, that is determining adequate tracers at a given time and estimating their velocity field. Although the camera moves, we take advantage of the previous remarks to set up the tracer determination problem as if the matter would be to determine moving areas over a stationary background. (Among detected moving areas, it may happen that some correspond to elements moving by themselves, e.g. fish;

they can be subsequently eliminated). Besides, the appearance of underwater images incites to resort to a statistical framework. Therefore the method developed in this paper finds a valorizing application in this context. The information of the instantaneous motion of each tracer is represented by the velocity field along its boundary. It is obtained by first estimating at every point the velocity vector component perpendicular to the border using and adapting the model-based method described in [2], and then by reconstructing the complete velocity field according to the stochastic gradient algorithm also presented in [2].

This paper focuses on the determination of moving areas in an image sequence through the study of variations in time of the intensity distribution. Detection of changes in time of the gray value distribution in an image sequence corresponds to one of the basic problems to be addressed in dynamic scene analysis, [7]. Change detection can be realized by considering intensity differences in time, [10,11], or by introducing more elaborate statistical tests, such as maximum likelihood tests, [4,6,7]. Relevant interpretation of these changes remains in the general case an open question. Nevertheless for a large class of applications they can be correlated to motion information owing to some assumptions about the image formation process, i.e. mainly, constant illumination. As a matter of fact, the need is not limited to change detection but encompasses the recovery of the masks of moving areas in the image. Usually, these two processing stages are performed one after the other by means of separate techniques. Change images may include false detection because of noise, but above all it comprises three kinds of regions among change regions, first, due to objects covering up background, second, due to background uncovered by objects and third, due to each object moving over itself. Hence a complementary process is required to select areas corresponding at a given time  $t$  to real moving objects. These areas will be concisely called *moving object masks* or *moving areas*. To this end solutions have been proposed which essentially rely on heuristics, [7,10,11].

Our approach is in particular distinguished by treating conjointly detection of temporal changes, and reconstruction of mobile object masks. Therefore we will mention it as the *moving-area detection-and-recovery* (MADR) process. To this end, spatial and temporal contextual information is explicitly introduced in the MADR process according to a probabilistic formulation. Besides no assumptions are made on the respective intensity levels of objects and background. More formally, this problem is stated as a statistical labeling one. To decide whether a point belongs or not to a moving area is equivalent to assign to it a given label. The label field is modeled as a Markovian field using Gibbs distributions. A solution to this labeling problem is formulated according to the maximum a posteriori (MAP) criterion. We have adopted a deterministic iterative algorithm to optimize the criterion at hand, based on an efficient manner of selecting sites to be visited, which drastically limit the computational load.

## LABELING AND CRITERION

We propose a probabilistic formulation of the *MADR* process, which provides both a convenient framework for modeling prior expectations concerning significant spatio-temporal properties of a moving area against stationary background, and with an efficient criterion for specifying the most likely interpretation when dealing with noise-corrupted observations. Interpretation must be understood here as the recognition of moving areas in the image. As already outlined, interpretation is stated as a statistical labeling problem.

Let  $E$  be the label field and  $e$  a possible realization of this random field. The label set is given by  $\Omega = \{-1, 0, +1\}$ ; 0 corresponds to stationary background, -1 and +1 to change areas. Moving areas will be defined as regions of connected (+1)-labeled points, (resp. (-1)-labeled), validated at the end of the optimization process supplying the most likely interpretation,  $\hat{e}$ . Let  $I_t$  be the intensity image array at time  $t$ , and  $p = (x, y)$  be a sample or image point. The intensity value at point  $p$  in  $I_t$  is denoted by  $f(p, t)$  or  $f(x, y, t)$ . Besides  $e_t(p)$  will designate the label of point  $p$  in image  $I_t$ , and  $e_t$  is the label array  $\{e_t(p)\}$ . The next step is to define how the observation must be chosen.

As we assume that motion appearance is closely related to change in time of the intensity function, we consider as observation the temporal derivative of the intensity function, denoted  $f_t$ . As we deal with digitized images, we use an approximation of  $f_t$ , that is the finite difference  $\tilde{f}_t(x, y) = f(x, y, t + dt) - f(x, y, t)$ , where  $dt$  is the time interval between successive images. In fact a filtered version of  $\tilde{f}_t$ ,  $\tilde{g}_t$ , will be introduced. Let  $O$  denote the observation field, which is supposed to be a random field, and let  $o$  be a possible realization of it. We wish to find the most likely interpretation,  $\hat{e}_t$ , in terms of moving areas, of the changes in time of the intensity. To this end, we resort to the maximum a posteriori (MAP) criterion, which leads to maximize the probability of the random label field given the observation data,  $P(E = e/O = o)$ . Using Bayes's rule and neglecting normalizing factor which is constant as far as the maximization is concerned, we get:

$$\max_e P(O = o/E = e)P(E = e) \quad (1).$$

## MODELING WITH GIBBS DISTRIBUTIONS

The point is now how to formalize the prior probability distribution of the interpretation  $P(E)$ . We model the label field as a Markovian random field. Markovian modeling effectively offers the appropriate mathematical concept to express local interactions in the labeling process. Besides we resort to Gibbs distribution, which allows to manipulate an explicit analytical expression of this prior probability. The use of Gibbs distribution in the context of image analysis and its interest were originally strengthened in [5]. Recent attempts to apply this approach to motion problems are reported in [9], concerning scene segmentation from visual motion.

The distribution of the label field is given by the following expression:

$$P(E = e) = \frac{1}{Z} \exp\left(-\frac{U(e)}{T}\right) \quad (2);$$

where  $Z$  is a normalizing constant called the partition function, and  $T$ , often referred to as temperature, may act as a global control parameter during the optimization process.  $U(e)$  is called the energy function. The main efficiency of this modeling approach lies in the fact that  $U$  is given by the sum of local potentials defined on so-called cliques  $c$ ,  $c \in C$ , where  $C$  is the set of cliques for a given neighborhood system.  $U$  is of the form:  $U(e) = \sum_{c \in C} V_c(e)$ . Indeed this modeling approach offers easy and mathematically well-mastered means of adapting the algorithmic solution to different situations while keeping its framework.

We have defined potentials on a set of cliques derived from a spatio-temporal neighborhood system shown in Fig.1. A clique is a subset of sites which are mutual neighbors. In the case at hand interpretation sites are pixels  $p$  and two types of cliques are possible, either a clique consists of one single pixel, or it comprises two sites

$(p_1, p_2)$ . There are four such spatial cliques, horizontal one (\*\*), vertical one ( $\begin{smallmatrix} * \\ * \end{smallmatrix}$ ), diagonal ones (\*\*), (\*\*\*) and one temporal one. There is no theoretical means to choose the best family of potentials associated with possible configurations of labels for each clique. This choice must be guided by the knowledge of the problem at hand. For the while we suppose the label field to be piecewise constant and we take a set of constant-level potentials. For the spatial cliques, they are defined as follows:

$$V_c = \beta_s \cdot \text{if } e(p_1) \neq e(p_2); \quad (3i)$$

$$V_c = -\beta_s \cdot \text{if } e(p_1) = e(p_2). \quad (3ii)$$

This kind of potentials favours spatial continuity of the label field; the optimization process will be inclined to eliminate very small regions, which can be assumed to correspond to false detection due to noise, and to form compact regions (stationarity principle). In consequence a complementary item must be introduced to cope with discontinuities. Discontinuity is materialized by points belonging to borders of two adjacent regions. Therefore edge location takes place at pixels of the interpretation site network. Edge state can be considered as an additional quality for an interpretation site. It is a deterministic information supplied by an edge detection operator applied in each image  $I_t$ . Accordingly potentials will be slightly modified to take into account this additional information, as explained in [8]. They are chosen in such a way to encourage to label the point of interest with the same label as points in its neighborhood which are not border points and which are situated on the same side of the contour. Three levels are defined  $\{-\beta_s, 0, \beta_s\}$ . A positive potential discourages the corresponding label configuration, a negative one encourages it and a potential equal to zero expresses a neutral opinion.

Before describing potential associated with the temporal cliques, let us point out a few aspects concerning change areas and moving areas, which will make more explicit the choice of these potentials. For this purpose, let us introduce the notation  $F_t$  for the image of intensity changes between image  $I_{t+dt}$  and  $I_t$ , validated for instance by thresholding the image difference or by using one of the likelihood tests proposed in [4,6]. One solution to be sure to keep only points corresponding to the position of the moving object in image  $I_t$  is basically to perform a logical-AND between  $F_{t-dt}$  and  $F_t$ , as described for instance in [10,11]. In the case of objects supposed to be projected as uniform intensity regions, this logical operation typically acknowledges the following succession of events at a given location  $p$  of the image: *background* (at  $t-dt$ ), *object* (at  $t$ ), *background* (at  $t+dt$ ). (If significant overlapping between successive projections of the object occurs, a complementary stage based on spatial properties of the intensity function is usually required in this case.) The above-mentioned remarks can give an intuitive insight into the definition of potentials associated with the temporal clique, knowing that we deal with a ternary labeling process. If site  $p_1$  is supposed to correspond to time  $t$  and site  $p_2$  to time  $t-dt$ , they are given by:

$$V_c = \beta_\tau \cdot \text{if } e(p_1) = e(p_2) \cdot \text{OR} \cdot e(p_2) = 0; \quad (4i)$$

$$V_c = -\beta_\tau \cdot \text{if } e(p_1) = -e(p_2) \cdot \text{OR} \cdot e(p_1) = 0. \quad (4ii)$$

The last case corresponding to  $e(p_1) = 0$  expresses that it is preferred to miss a detection if doubtful rather than to induce a false alarm. We do not take into account the unit clique in the optimization process. This means that all labels are assumed to be of equal probability.

The last step is to define the conditional probability of the observation. The observation is given by the filtered temporal derivative of the intensity function,  $\tilde{g}_t$ . We model this observation as a piecewise constant function corrupted by an additive zero-mean Gaussian noise. This model is a tractable one, even if sometimes unrealistic. Nevertheless it is suited enough for the application we deal with. Results reported in this article indicate that the use of such a model is acceptable. The observation at each point  $p$  follows a Gaussian law of mean  $\mu$  and variance  $\sigma^2$ . The mean depends on the label according to a simple law, [8]. Variance  $\sigma^2$ , which is assumed to be independent of the label, and only due to noise effect, is a constant with respect to the optimization process. It is supposed to be space invariant. If we assume that observation variables are independant

from one pixel to the other, the conditional probability of the observation  $O$  is of the form:

$$P(O) = \frac{1}{(2\pi\sigma^2)^N} \exp\left\{-\sum_{\mathbf{p} \in I_t} \frac{(g_t(\mathbf{p}) - \mu_{e_t(\mathbf{p})})^2}{2\sigma^2}\right\} \quad (5)$$

where  $N$  is the number of pixels in  $I_t$ .

The resulting a posteriori probability distribution of the label field given by the product of expressions (2) and (5), is again a Gibbs one. If we take the opposite of the logarithm of it and neglect constant terms, we come to:

$$W(e) = \frac{1}{2\sigma^2} \sum_{\mathbf{p} \in I_t} (g_t(\mathbf{p}) - \mu_{e_t(\mathbf{p})})^2 + \frac{1}{T} \sum_{c \in C} V_c(e) \quad (6)$$

$W$  can be called the total energy function. The most likely interpretation will then result from the minimization of this function  $W$  with respect to  $e$ . The next section will present the procedure designed for this purpose.

### OPTIMIZATION PROCEDURE

The formulation of the labeling problem through the MAP criterion gives rise to several intrinsic difficulties, in particular the number of variables involved in the minimization of (6), equal to the size of the image. A stochastic optimization procedure called simulated annealing ensures that the global minimum can be reached asymptotically speaking, regardless of the initial state, [5]. In practical situations, this algorithm presents the major drawback of requiring a tremendous amount of iterations. This explains several efforts for designing other procedures indeed sub-optimal but computationally more efficient, as surveyed in [3]. We have adopted a deterministic iterative algorithm to optimize the criterion at hand, which could be defined as a deterministic version of the Metropolis sampling algorithm. Temperature parameter  $T$  is set to a constant value. We start from a given initial configuration of the label field,  $e_t^?$ . By iteratively minimizing relation (6), we are supposed to reach the most likely interpretation  $\hat{e}_t$ ; it is delivered by the final configuration denoted  $e_t^f$ . This minimization procedure works as follows. After selecting a site  $\mathbf{p}$ , we compute the energy change induced by modifying the current label of site  $\mathbf{p}$ ,  $\omega_c$ , for all  $\omega \in \Omega \setminus \{\omega_c\}$ , that is:

$$w(\mathbf{p}, \omega) = W(e_t^?) - W(e_t) \quad (7)$$

where the field  $e_t^?$  is defined as equal to  $e_t$  except for site  $\mathbf{q} = \mathbf{p}$  where  $e_t^?(q) = \omega$ . Of course the evaluation of  $w$  involves only local computations. If all variations  $w(\mathbf{p}, \omega)$  are positive, label  $\omega_c$  is kept at site  $\mathbf{p}$ , otherwise  $e_t(\mathbf{p})$  is set to  $\omega_m$  that minimizes  $w(\mathbf{p}, \omega)$ . This systematic minimum search is reasonable, since we deal with only three labels and we use constant-level potentials. Computational cost is rather low.

In fact sites are ordered according to a stability criterion as introduced in [3]. Let us denote  $\psi$  this stability function.  $\psi$  is merely defined as follows:

$$\psi(\mathbf{p}) = w(\mathbf{p}, \omega_m) \text{ if } w(\mathbf{p}, \omega_m) < 0; \psi(\mathbf{p}) = 0, \text{ otherwise.} \quad (8)$$

This function  $\psi$  measures how significantly a more stable labeling, in the sense of a lower energy state, can occur from an alternative decision. At the beginning of the minimization process, given an initial configuration  $e_t^?$ , a stack is constructed which contains sites  $\mathbf{p}$  such that  $\psi(\mathbf{p}) \neq 0$ . These sites are ranged according to the value of  $\psi(\mathbf{p})$ ; the site at the top,  $\mathbf{p}^*$ , corresponding to the lowest value of  $\psi$ , will be considered by the minimization process. Once the label of  $\mathbf{p}^*$  has been updated, function  $\psi$  is computed again for neighbors of  $\mathbf{p}^*$ . The stack is reorganized accordingly and new sites may be added to it. This operation is iterated until the stack is empty leading to the final configuration  $e_t^f$ . This way of selecting sites, by constantly focusing on sites which are the most likely to be mislabeled, enables to drastically limit the total number of iterations. In return the construction and the maintenance of the stack must be taken into account. The smaller the initial size of the stack the more efficient this technique. This militates for taking care of the step of initialization, that is determining  $e_t^?$ . Moreover the optimization process must start not too far from the global minimum in order to prevent it from being

captured by irrelevant local minima, since a deterministic method is used.

### INITIALIZING STEP

The determination of  $e_t^?$  subdivides into two stages. First an augmented set of labels is considered,  $\bar{\Omega} = \{-1, 0, +1, NIL\}$ .  $NIL$  signifies that no labeling decision has been made for the while. The intermediate field label is denoted  $e_t^{init}$ . It is obtained through the use of a slightly-modified version of one of the change detectors based on a likelihood test, described in [6]. We model the intensity distribution as a piecewise constant function corrupted by an additive zero-mean Gaussian noise. We assume the noise variance to be constant over the two windows  $M_1$  and  $M_2$ , centered on  $\mathbf{p}$  respectively in  $I_t$  and  $I_{t+dt}$ , which are considered to validate or not a change in time at point  $\mathbf{p}$ . This assumption of constant variance leads to a simplified expression for the likelihood ratio. This ratio corresponds to change hypothesis versus no-change hypothesis. Once optimal estimators of the means of the different Gaussian laws involved are derived, we can substitute these estimators for means in the likelihood ratio. Then the optimized likelihood ratio can be rewritten in the following form, after some mathematical developments and after taking its square root, [8]:

$$|\xi(\mathbf{p})| = \frac{1}{2n\sigma} \left| \sum_{q \in M_p} \tilde{f}_t(q) \right| \quad (9)$$

where  $M_p$  designates the set of common spatial locations within  $M_1$  and  $M_2$ , and where  $n$  is the width of these windows if they are assumed to be square ones. The label field  $e_t^{init}$  is now defined as follows. A two-threshold strategy is used for the test. If the likelihood ratio is inferior to the lowest one or superior to the highest one, one of the labels  $0, -1, +1$ , is assigned to  $\mathbf{p}$  according to the case at hand, [8]. Otherwise no decision is taken, that corresponds to the  $NIL$ -label. The principle of this first stage is to proceed with decisions as less committing as possible while making use of the low computational load of this test to give pre-initial labels from the set  $\bar{\Omega}$  to the maximum of sites corresponding to unambiguous situations. Then, in the second stage, the available contextual information will be used to assign a label from the set  $\bar{\Omega}$  to sites still with the  $NIL$ -label. Starting from the pre-initial label field  $e_t^{init}$ , we apply the MAP criterion (6) to the subset  $\chi$  of points  $\mathbf{p}$  in  $I_t$  whose pre-initial label is  $e_t^{init}(\mathbf{p}) = NIL$ . (The stack-based version is not used in this case). Then we get the label field  $e_t^?$  with labels from  $\bar{\Omega}$ . As for the determination of the final configuration  $e_t^f$ , the minimization of expression (6) involves two complementary data arrays, that is the border map  $B_t$  and the change map  $F_{t-dt}$  between images  $I_t$  and  $I_{t-dt}$ . This last one is obtained again using the likelihood ratio given in (9), but without considering the  $NIL$  label. Therefore only one threshold is taken into account and three labels  $\{-1, 0, +1\}$ . The label field  $e_{t-dt}$  which is concerned with the determination of the temporal-clique potential, formula (4i-4ii), is precisely given by  $F_{t-dt}$ , then denoted by  $e_{t-dt}^{ef}$ . It is not the label field  $\hat{e}_{t-dt}$  which yields the moving areas referred to image  $I_{t-dt}$ .

### RESULTS

Experiments have been carried out with a real underwater video image sequence acquired in a dock at IFREMER-Brest (French National Research Center in Oceanography). Three images ( $I_{t-dt}, I_t, I_{t+dt}$ ) out of one sequence are shown in Fig.2. Oblong shapes correspond to objects sunk in the dock and lying on the bottom surface. The border image  $B_t$  is derived from the zero-crossings of the difference of two filtered versions of image  $I_t$ , obtained with two Gaussian filters of respective standard deviation,  $\sigma_1 = 4$  and  $\sigma_2 = 2$ . Potentials are given by  $\beta_s = 64$  and  $\beta_t = 500$ . The observation  $g_t$  is simply obtained by averaging within a  $5 \times 5$  spatial neighborhood the function  $\tilde{f}_t$ . Fig.3 shows the "pre-initial" label field  $e_t^{init}$  obtained with the modified likelihood test before any contextual information is considered. The final result delivered by the complete optimization process is presented in Fig.4. Moving areas resulting from  $e_t^f$  are delineated. Complementary results corresponding to the different

intermediate steps of the procedure can be found in [8]. By comparing Fig.3 and Fig.4 it can be pointed out that nearly all change regions due to uncovered background have been eliminated, and non-reliable detections have been swept out. Among validated moving areas the most significant ones coincide with the objects lying on the bottom, and almost all the others with distinguishable blobs on the bottom surface. A sequence of twenty images have been processed. It was found that the choice of potential values is not a critical matter, and for a given set of parameter values, rather stable and consistent results have been obtained.

**CONCLUSION**

The final goal of the whole project, which is conducted by the French National Research Centre in Oceanography (IFREMER), is to achieve a prototype, which would prove the feasibility of such an on-board equipment for underwater robot navigation. Hence a tradeoff had to be found between the algorithmic complexity of a solution able to cope with underwater video images and its implementation as efficient as possible. This leads us on one hand to achieve an elaborate statistical modeling of a spatio-temporal tracer using Gibbs distributions, and on the other hand to resort to a deterministic optimization procedure to solve the labeling problem designed. Current work deals with the processing of other sequences of underwater video images, which were acquired off-shore. We will also investigate a few extensions of this study, such as the consideration of other models for the observation, and of other families of potential functions.

**References**

1. J.M. AUGUSTIN, J.M. BOUCHER, P. BOUTHEMY, J.P. HUE AND G. LOAEC, "Trajectographie sous-marine et mosaïques d'images video", *Proc. Colloque Espace et Mer*, Marseille, Nov. 1987
2. P. BOUTHEMY, "Extracting dense motion information from an image sequence: optic flow estimation and related segmentation issues", *NATO Advanced Research Workshop: Kinematic and Dynamic Issues in Sensor-based Control*, II Ciocco, Italy, Oct. 1987
3. P.B. CHOU AND R. RAMAN, "On relaxation algorithms based on Markov random fields", TR 212, Comp. Sci. Dpt, Univ. of Rochester, July 1987
4. G.W. DONOHOE, D.R. HUSH, AND N. AHMED, "Change detection for target detection and classification in video sequences", *Proc. ICASSP-88*, New-York, pp.1084-1087
5. S. GEMAN AND D. GEMAN, "Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. PAMI-6, No 6, Nov. 1984, pp.721-741
6. Y.Z. HSU, H.-H. NAGEL AND G. REKERS, "New likelihood test methods for change detection in image sequences", *Computer Vision, Graphics and Image Processing*, Vol.CVGIP-26, 1984, pp.73-106
7. R. JAIN, "Dynamic scene analysis", in *Progress in Pattern Recognition 2*, L. Kanal and A. Rosenfeld (eds.), Machine Intelligence and Pattern Recognition Series, Vol. 1, North-Holland, 1985, pp.125-167
8. P. LALANDE AND P. BOUTHEMY, "Détermination spatio-temporelle de traceurs dans une séquence d'images sous-marines à des fins de trajectographie", Research Report PI-416, IRISA, June 1988, 40p.
9. D.W. MURRAY AND B.F. BUXTON, "Scene segmentation from visual motion using global optimization", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. PAMI-9, No 2, March 1987, pp.220-228
10. S.A. SUGIMOTO, H. MATSUKI, AND Y. ICHIOKA, "Implementation of tracking and extraction of moving objects in successive frames", *Applied Optics*, Vol. 25, No 6, March 1986, pp.990-996
11. J. WIKLUND AND G.H. GRANLUND, "Image sequence analysis for object tracking", *Proc. 5th Scandinavian Conf. on Image Analysis*, Stockholm, June 1987, pp.641-648

**Acknowledgements**

The whole project, called TOS "Trajectographie Optique Sous-marine", includes five labs, IFREMER, ENSTBr, IRISA, THOMSON-LER and THOMSON-SINTR. It is supported by the Brittany County Council and the French Ministry of Research and Technology, in the context of the "Collaboration Bretagne Image".

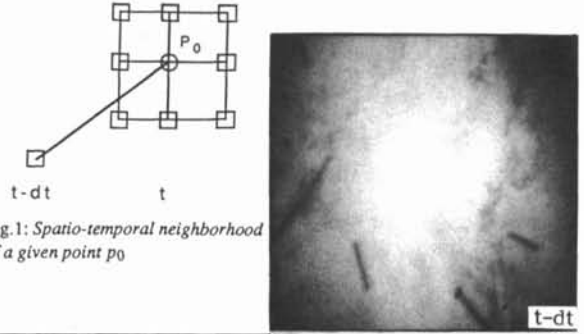


Fig.1: Spatio-temporal neighborhood of a given point  $p_0$

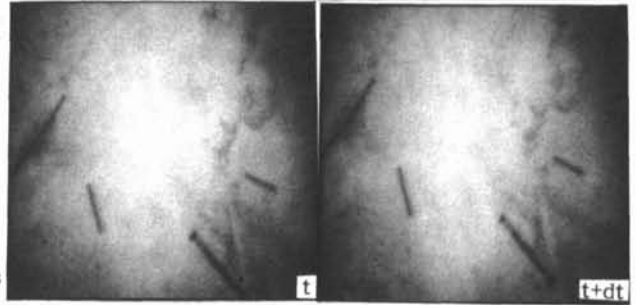


Fig.2: Three successive images from an underwater video image sequence acquired in a dock; size is 256x256 pixels

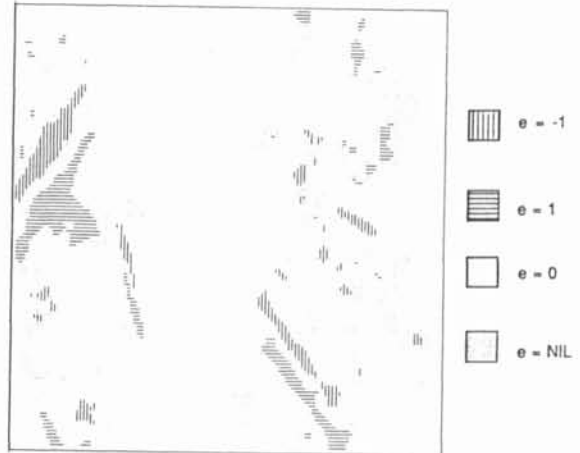


Fig.3: Pre-initial label field,  $e_i^{init}$ , considering four labels  $\Omega = \{-1, 0, +1, NIL\}$ , with  $\lambda_1 = 200$ ,  $\lambda_2 = 350$



Fig.4: Delineation of validated tracers, as subsets of connected points all labeled by -1, or by +1 in the final label field  $e_f$