MVA '92 IAPR Workshop on Machine Vision Applications Dec. 7-9,1992, Tokyo

Cloudy sky velocity map based on matched filter

COLLET C. *, QUINQUIS A.(1)*, BOUCHER J.M **

(*) Groupe de Traitement du Signal - Batiment des laboratoires
ECOLE NAVALE - 29240 BREST NAVAL - FRANCE -
telephone number : 98.23 40 44
telefax number : 98.27.57.04

(**)ENSTBr Département MSC
B.P 832 - 29283 BREST Cédex - FRANCE
telephone number : 98.00.11.11
telefax number : 98.45.51.33

In an infrared surveillance system in an aerospace environment (which is tasked with the detection of remote sources thus with a very low resolution), the cloudy sky velocity estimation should generate a lower false alarm rate in discriminating the motion between different moving shapes, by means of a background velocity map. The optical flow constraint equation, based on a Taylor expansion of the intensity function, is often used to estimate the motion for each pixel. However one of the main problems we faced with motion estimation is that we cannot have for one pixel a complete knowledge of the real velocity because of the aperture problem. Another kinematic estimation method is based on matched filter (Generalized Hough Transform (GHT)), it gives a global velocity estimation for a set of pixels. On the one hand we obtain a local velocity estimation for each pixel with little credibility, because of the sensitivity to noise of the optical flow; on the other hand, we obtain a global robust kinematic estimation identical for all pixels. This paper aims to adapt and improve the GHT in our typical application in which one has to discern the global movement of objects (clouds) whatever their form may be (clouds with hazy edges or distorted shapes or even clouds that have very little structure). We propose an improvement of the GHT algorithm through an image segmentation using polar constraints on spatial gradients : one pixel, at time t, will be matched with another one at time $t+\Delta T$, only if the direction and modulus of the gradient are similar. This technique, which is very efficient, sharpens the peak and improves the motion resolution. Each of these estimations is calculated within windows belonging to the image, these windows being selected by means of an entropy criterion. Using the property of the aperture problem, all the pixels selected for the GHT will have their own optical flow as a velocity component. This kinematic vector will be computed accurately by means of the optical flow constraint equation applied on the displaced window. We showed that, for small displacements, the optical flow constraint equation sharpens the results of the GHT. Thus a semi-dense velocity field is obtained for cloud edges. A velocity map computed on real sequences with these methods will be shown. In this way, a kinematic parameter will allow discrimination between a target and the cloudy background.

KEYWORDS : **Motion estimation, Pattern recognition, Generalized Hough Transform, Crosscorrelation, Optical flow.**

I) Introduction

This paper covers the detection and the estimation of cloudy kinematic fields. It will present the two main existing methods of measuring the motion and the main physical limitations to a correct measurement of the velocity. The first section propose a method based on an entropy criterion that detects and localizes the image areas (called windows) where there is motion. The second section will propose three methods belonging to the set of global methods : the Fourier analysis, the crosscorrelation method and the Generalized Hough Transform (GHT). This last one will be improved by using a polar constraint during the matching process. Two credibility vectors will be used in order to weigh the estimates. The segmentation within the windows between matching pixels and pixels that are eliminated, will be presented in this section. The third section deals with a differential method which gives further velocity information for each pel. This local method is not able to take into account large displacements but can give a velocity estimation in the gradient direction for each pel. We will show that this optical flow estimation can be computed after the GHT method. The combination of the two will sharpen the velocity component within a window. An example demonstrating the performance of this technique will be presented.

I.1) Motion Estimation Methods

A study of the different motion estimation methods in an image sequence quickly shows that these methods must be classified into two major types (1) : features based methods (overall view) and differential methods (local view).

These methods are called global methods when the velocity estimation is a displacement mean value for a set of pixels. Primarely the crosscorrelationmethod is used. Global methods can use workspace transformations where the problem is supposed to be simplier to resolve. These techniques take into account the Fourier and the Hough transform methods. The kinematic vector is global and similar for all the pixels in the window. The method of selection of the windows will be presented at the end of this section (cf. I.3). Local methods are interesting in the sense that they provide a dense kinematic field of estimates : a motion measurement for each pixel is theoretically possible. The techniques used usually refer to the optical flow constraint equation which links (through a first order Taylor expansion) the spatio-temporal derivatives of the irradiance function to the kinematic vector. Sensitive to noise because of their differential nature, these local methods have the great advantage of not imposing any constraint on the type of displacement to be quantified. Unfortunately, only small displacements can be correctly estimated.

Indeed, the object of all these methods is to define an entity, which is not in itself immedietly quantifiable, through different filters or mathematical tools.

I.2) Aperture problem

The aperture problem is the main limitation when displacement components are computed (2) : because of the aperture problem we can only locally estimate the velocity , projected on the gradient direction for an oriented shape, as shown on figure 1. The velocity components parallel to the edge direction remains invisible.

In our application we have no information about the depth of or distance between clouds. This means that some occlusion phenomenon may occur because of relative velocity due to perspective projection (figure 2). Moreover the velocity estimation is a function of the azimuth angle and the direction of the wind vector : we will not have to estimate an ego-motion over the whole image. For this reason we propose to place small windows in our image where there is motion. This should decrease the probability of having an occlusion on the set of pixels on which we are working. A second problem is that the clouds have hazy edges, distorted shapes and may have very little structure. Here we form the hypothesis that distorsion between two images, due to the temporal sampling rate, is unperceptible. This hypothesis will be later verified on real image sequence.

I.3) The choice of working windows solved by an entropy criterion

Fistly we first applied a 3x3 median filter to eliminate impulsionnal noise without blurring the cloud edges. We detect motion through a difference of irradiance in time, that is why we propose to compute

(1) ENSIETA - 29240 Brest Naval

the derivative in time as follows:

$$\text{diff}(x,y) = \text{abs}(\, f(x,y,t) - f(x,y,t+\Delta T)\,)$$

where f represents the image sequence and ΔT the sampling time.
In our application we want to lead our process into regions of greatest disturbance. Our hypothesis is based on the fact that motion induces more disturbance than noise does in the picture obtained using the above formula.
The main principle is to place the working windows in areas in which temporal evolution of the irradiance is most stochastic (least determined). The entropy factor, as defined below, is a measurement of the mean incertitude in a stochastic set:

$$H = -\sum_{i=-2^{m-1}}^{2^{m-1}-1} P_i.\log(P_i)$$

where Pi represents the probability of pixels (with irradiance i) appearing in the window, each pixel being coded in m bits. The picture is divided into four quadrants.

On each picture, we apply a recursive split algorithm that cuts every window down into four smaller ones. At each step we compute four new entropy criteria which are compared with the preceding entropy values (in the father window) and with those in the three brother windows. The largest entropy factor is always chosen. The algorithm stops when a window smaller than 30x20 pixels appears.
The maximum entropy is obtained with equiprobability in the histogram of diff(x,y). This sets the upper limit of the entropy criterion:

$$0 \le H \le \log(255).$$

This entropy technique gives good results for real sequences.
We observed that our windows were created in areas of greatest disturbance (figure 3).

II) Global Methods

This section shows three global methods.

II.1) The Fourier transform method

The Fourier transform method generates the motion information by computing the phase difference between two time samples (only in the case of translatory motion (9)). However, not only does this method entail a very long computing time when applied to a bidimensional signal, but it can only be used in the case of a translation on a single object. In addition, the moving source spectrum features must be known. In conclusion, there are too many constraints for use in our application.

II.2) The crosscorrelationmethod

The crosscorrelationmethod can be used to match the irradiance function f(x,y,t) in time : $f(x+\Delta_X, y+\Delta_y, t+\Delta_T) = f(x,y,t)$) where $(\Delta_X\Delta_y)$ represents the displacement and Δ_T is the temporal sampling rate. The crosscorrelationcan also be calculated for two previously binarized images (crosscorrelationon segmented images) to decrease processing time.
The normalized mean crosscorrelationexpression is :

$$C'_{ff}(\vec{\Delta}) = \sum_{m=-M}^{M}\sum_{n=-N}^{N}\left[\left(f(m-\Delta_x,n-\Delta_y,t-\Delta T) - \mu_f(t-\Delta T)\right).\left(f(m,n,t) - \mu_f(t)\right)\right]$$

$\vec{\Delta}\left[\Delta_x, \Delta_y, \Delta_t\right]$ is a spatiotemporal vector of the displacement and $\mu_f(t)$ represents the average intensity value of the image f at time t.
The correlation factor is also defined as:

$$C'_{ff}(\vec{\Delta}) = \sum_{m=-M}^{M}\sum_{n=-N}^{N}\left[\left(f(m-\Delta_x,n-\Delta_y,t-\Delta T) - \mu_f(t-\Delta T)\right).\left(f(m,n,t) - \mu_f(t)\right)\right]$$

$$C''_{ff}(\vec{\Delta}) = \frac{C'_{ff}(\vec{\Delta})}{\sigma_f(t).\sigma_f(t-\Delta T)} \quad \text{where} \quad -1 \le C''_{ff}(\vec{\Delta}) \le 1$$

and $$\sigma_f^2(t) = E\left\{\left(f(\vec{S},t) - \mu_f(t)\right)^2\right\}$$

This last formula gives good normalized results, but the required computation time is high because standard deviation has to be calculated repeatedly. Crosscorrelationon previously segmented and binarized images can be computed more quickly (boolean operations) which improves the quality of the correlation peak.

Only translation of a single rigid object on a stationnary background can be perfectly detected. It is therefore difficult to use this method on windows that are too large, because clouds can appear and move at different speeds and in different directions. However, this method seems to be applicable to isolated cirrostratus or stratocumulus clouds on blue sky background.

II.3) The generalized Hough transform method

The generalized Hough transform method must be applied to binarized pictures. It describes a mapping between points in feature spaces and points in parameter spaces. Sklansky (3) has shown that this transform method is equivalent to a matched filtering process between the selected pattern at time "t" and the image at the next time sample : "$t+\Delta_T$". There is also an equivalence between the Generalized Hough transform method and template matching. Of course, this method gives better results than the correlation technique because we more accurately define the pattern required in the new picture. We obtain a cluster array (called accumulator) where the peak position provides the estimation of the displacement vector. We have improved the performance of the Generalized Hough Transform method by using a polar constraint in the matching process which sharpens the peak emergence. The mathematical formulation of the GHT is as follows :

Let $\vec{\alpha}$ and $\vec{\beta}$ be two spatial vectors.

Let $b(\vec{\beta}) = b(x_\beta, y_\beta)$ the binarized window threshold by keeping N pels of the larger gradient modulus (cf. 2.4).

Let $h(\vec{\alpha}) = h(x_\alpha, y_\alpha)$ the N vectors belonging to the set of selected points :

for each $b(\vec{\alpha}) = 1$, let $h(x_0 - x_\alpha, y_0 - y_\beta) = 1$ with (x_0, y_0) representing the origin of the space.

Finally, let $A(\vec{\alpha}) = A(x_\alpha, y_\alpha)$ the accumulator matrix which has been initialized with zero values;

$$A(\vec{\alpha}) = \sum_{\vec{\beta}\,(x_\beta, y_\beta)} h\,(x_\alpha - x_\beta, y_\alpha - y_\beta)\,.\,b(x_\beta, y_\beta)$$

Each cell of the matrix A has a counter and for each pixel selected in the binarized window, N cells are reached by the process (figure 5). At the end, the position of the larger counter correponds to the pattern position in the next frame.
Here a question is raised : how are the matching point of interest defined ?
Clouds have hazy edges, distorted shapes and no internal structure, thus it is very difficult to know what pixel irradiance will be kept in the new picture. We resolve this problem by using Sobel's eight direction edge detector (4) and by only considering the highest gradient modules. These are the points of interest and are used for pattern matching. We also neglect the cloud description using string edges which are unrealizable when describing cloud features (5).

II.4) Thresholding by a correlation criterion

With the two preceeding methods there is a threshold choice to do (how many gradients with high modulus are to be kept ?). In our application, we want to extract the clouds from the background to obtain a binary image that can be used by the generalized Hough transform method. This threshold choice can be heuristic (invariable) or processed (matched in time). Some methods are based on the bimodal histogram search by means of iterative methods. They don't always provide satisfactory results and processing time depends on the distance between the bimodal histogram model and the real histogram. Here the process consists of a binarization that preserves the same mean value for two images, and that tries to have the correlation factor between them, near to 1 (6).

let $P(\alpha_s) = \sum_{i=0,\, x[i] > s}^{N} P(\,x[i]\,)$ and $P(\beta_s) = \sum_{i=0,\, x[i] \le s}^{N} P(\,x[i]\,)$

where P(x[i]) is the probability of having the intensity x[i] in the raw image x and s is the variable threshold.
The picture is binarized with two values : y1 (if x(i) > s) and y2 (if x(i) ≤ s).

$$y_1 = \frac{\alpha_s}{P(\alpha_s)} \qquad y_0 = \frac{\beta_s}{P(\beta_s)}$$

We obtain a crosscorrelationfunction which is a function of s :

$$\rho_{(x,x_s)} = \frac{\alpha_s^2/P(\alpha_s) + \beta_s^2/P(\beta_s) - E[x].(\alpha_s + \beta_s)}{\sigma_x.\sigma_{x_s}}$$

where E[x] represents the average of the raw picture; σ_X the standard deviation of the raw picture and σ_{XS} the standard deviation of the binarized picture (function of s).
The high value of ρ provides the threshold value S, which

maximizes the SNR between the raw picture and the thresholded image. This method is interesting because of its constant processing time : we always obtain a threshold adapted value. An example showing the performance of this technique is presented in figure 6.

II.5) Improvements and credibility vector

II.5.1) Crosscorrelation

A two component credibility vector is defined which describes the quality of the estimation (figure 7).

α describes the quality of the resolution and is normalized between 0 and 1. In the perfect theoretical case, we obtain a Dirac's pulse and $\alpha = 1$ (infinite resolution).

$$\alpha = \frac{C^{*}_{ff}(\Delta_x,\Delta_y)}{\sum_{i=\Delta_x-3}^{\Delta_x+3} \sum_{j=\Delta_y-3}^{\Delta_y+3} C^{*}_{ff}(i,j)}$$

where $\max(C^{*}_{ff}(i,j)) = C^{*}_{ff}(\Delta_x,\Delta_y)$

β describes the matching process quality and is normalized between 0 and 1. In case of perfect matching, we obtain $\beta = 1$ (perfect matching).

$$\beta = \frac{C^{*}_{ff}(\Delta_x,\Delta_y)}{N}$$

The results, obtained directly on a real image sequence, show a low resolution at the peak (α) of the crosscorrelation matrix for the motion estimation, but the likelihood measurement (β) between these pictures is excellent (α and β are often written as a percentage). This implies that there is little distortion between clouds from one sample to the next. We can improve the resolution with binarized images but the likelihood decreases.

II.5.2) Generalized Hough Transform

We have defined a credibility vector with two components which describes the quality of the cluster array estimation. As before, these criteria have been normalized so that comparisons may be made.

The definition of α is :

$$\alpha = \frac{A(\Delta_x,\Delta_y)}{\sum_{i=\Delta_x-3}^{\Delta_x+3} \sum_{j=\Delta_y-3}^{\Delta_y+3} A(i,j)}$$

where $\max(A(i,j)) = A(\Delta_x,\Delta_y)$.

The definition of β is :

$$\beta = A(\Delta x,\Delta y) / N$$

where N stands for number of points to match.

β describes the set of points which group together to create the high array cluster value, in a 7x7 pixel neighborhood around the peak.

α and β are complementary thus it is not necessary to obtain a very good resolution (where α tends toward 100%) with a low β value (β represents the number of points actually matching). Some results showing this are in figure 8. We propose a new method for improving the peak emergence using a polar constraint. The high gradients are classified into 6 groups of intensity and 6 groups of direction. Two points, each belonging to two successive images cannot be matched if they don't belong to the same intensity and direction groups. This, the cluster array becomes more robust as shown in figure 9. Different numbers of polar groups have been tested and only the ones which improves the resolution without markedly diminishing the β criterion were kept (In our test on real sequence, we have only considered results for which β remains higher than 20 %).

II.6) Conclusion

We proposed and tested a new approach to match feature points by means of a polar constraint on the generalized Hough transform method. We compared the results using the crosscorrelation method and proposed two normalized credibility factors. We show now that optical flow techniques are actually tested to define accurately and more locally the velocity components in order to obtain a dense kinematic field estimation. This module will still allow a fairly rough estimate of cloud motions. The translation aspect is now satisfactorily calculated and its other aspects (homothetic, rotationnal motions or others) will be studied more accurately by means of optical flow.

III°) Optical flow estimator

Based on a first Taylor expansion of the intensity function, the optical flow constraint equation links the kinematic parameter to the spatio-temporal gradient as follows (7)(8):

$$\dot{f}_x \dot{x} + \dot{f}_y \dot{y} + \dot{f}_t = 0$$

where $(\dot{f}_x, \dot{f}_y, \dot{f}_t)$ are the first derivatives of the the intensity function in space and time.

and

$$\dot{x} = \frac{\Delta_x}{\Delta T} \text{ and } \dot{y} = \frac{\Delta_y}{\Delta T}$$

ΔT is the temporal sampling rate.

Because of the aperture, the preceding equation has a unique solution vector that satisfy :

$$\dot{f}_x \dot{y} - \dot{f}_y \dot{x} = 0$$

Thus, the optical flow coefficients at pixel i are defined as :

$$\dot{x}_i = \frac{\dot{f}_{x_i} \dot{f}_{t_i}}{\sqrt{\dot{f}^2_{x_i} + \dot{f}^2_{y_i}}} \qquad \dot{y}_i = \frac{\dot{f}_{y_i} \dot{f}_{t_i}}{\sqrt{\dot{f}^2_{x_i} + \dot{f}^2_{y_i}}}$$

III.1) Validity of the constraint equation

The constraint equation assumes that a modelization of the irradiance function f(x,y,t) according to a first order Taylor's expansion is verified. In our study, f is a digital signal; i.e. a discrete and quantified signal. If the constraint equation is always valid when mathematical limits are computed we have to verified that these hypothesis are really repected. The easiest mean consists in chcking that the irradiance function slope does not exceed the pattern displacement :that is why we have low pass filtered the image before computing the optical flow coefficients.

III.2) Interest of computing optical flow

We proposed a method based on an optical flow estimation that will locally sharpen the global translatory motion estimation. This method, based on a differential analysis is obviously less robust than the matched filter and limited to low displacements : as the speed increases, the robustness decreases. But after the global displacement estimation, the estimation of the optical flow values for the pixels precedly selected for the GHT computation allow to sharpen the kinematic map. The optical flow estimation takes into account homothetic or rotational movement, discriminates the motion between different pattern within a window.

III.3) conclusion

This method that combines optical flow estimation after global displacement gives a semi-dense velocity map as shown in figure 10.

IV) Some results

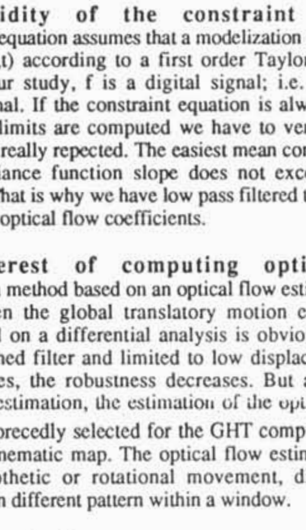
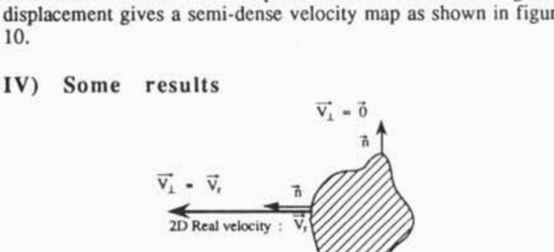


Fig 1 Aperture problem : main limitation to a complete knowledge of the real displacement



The velocity perceived is not the same for the two clouds

Figure 2 : perspective projection and occlusion phenomenon

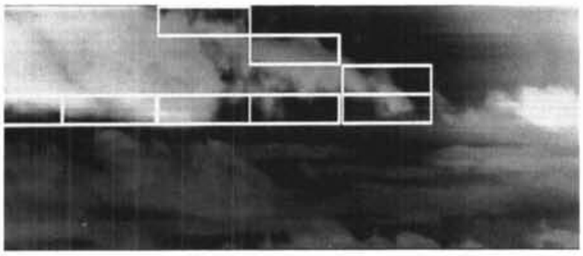
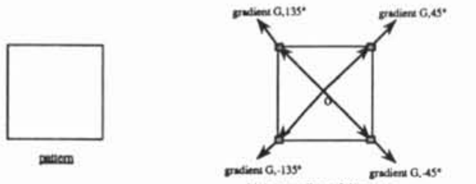


Fig 3 image with windows



pattern

highest gradient of the pattern

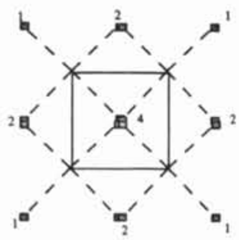


Fig 4 : GHT

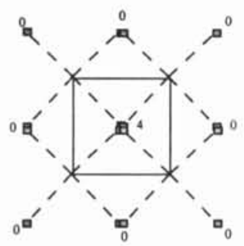


Fig 5 GHT with polar constraint

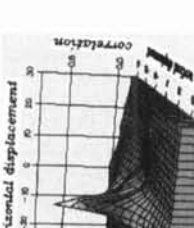


Fig 7 : Crosscorrelation :

$\dot{x} = 2$ pel.s^{-1} ; $\dot{y} = 0$ pel.s^{-1} ; $\alpha = 4{,}2\%$; $\beta = 95\%$

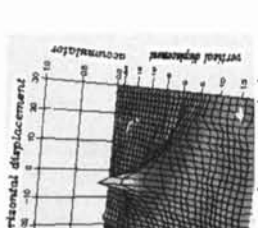


Fig 8 : GHT (without polar constraint) :

$\dot{x} = 1{,}98$ pel.s^{-1} ; $\dot{y} = 0{,}01$ pel.s^{-1} ; $\alpha = 4{,}3\%$; $\beta = 92\%$

Fig 9 : GHT (6 direction groups and 6 gradient groups) :

$\dot{x} = 2{,}01$ pel.s^{-1} ; $\dot{y} = 0{,}02$ pel.s^{-1} ; $\alpha = 12{,}7\%$; $\beta = 59\%$

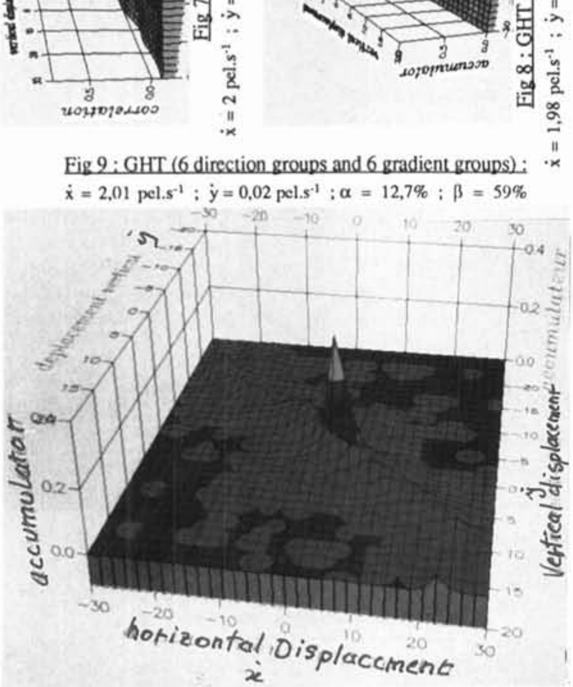
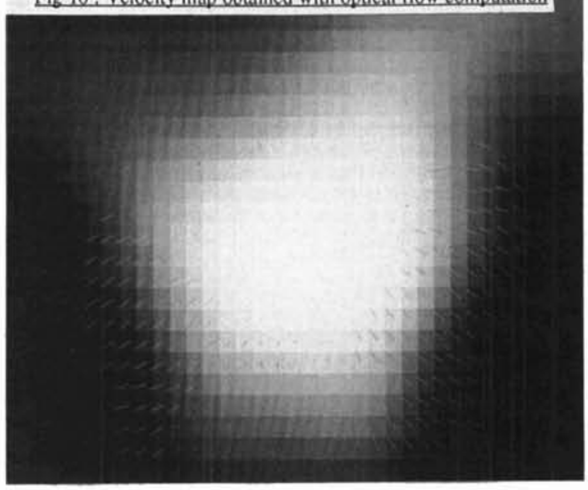


Fig 10 : Velocity map obtained with optical flow computation



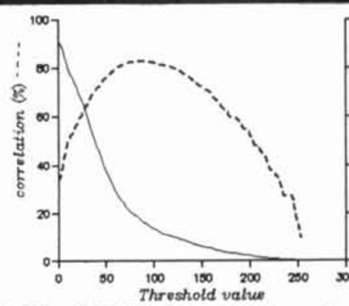


Fig 6 : Thresholding (16,9% of highest modulus gradient are selected)

V) Conclusion

As a conclusion, we proposed a new approach with motion extraction in the case of cloud velocity estimation. We first proposed and tested an entropy criterion which leads our process in windows where there is a lot of disturbance in time (i.e. edge of moving clouds). We tested a new approach to match feature points by means of a polar constraint on the generalized Hough transform method. We compared the results using the cross-correlation method and proposed two normalized credibility factors. Meanwhile we tested optical flow techniques to define accurately and more locally the velocity component to obtain a dense kinematic field estimation. The translation aspect is satisfactorily calculated by means of the GHT method and the other aspects (homothetic or rotationnal motions) will be take into account with optical flow estimation.

We tested these methods on actual pictures and showed the improvements brought by these techniques. Within the framework of our processing chain, optical flow techniques are implemented to accurately and more locally define the velocity components in order to obtain a dense kinematic field estimation, after a global recalage by means of a global matched filter. A new contribution toward understanding motion estimation has been presented. Since then, a continuous interaction between theorctical and experimental investigations has resulted in a gradual development of how local and global methods might be combined and how such a formulation could be evaluated in an autonomous processing chain.

Acknowledgments : The authors wish to express their deep gratitude to the Direction des Recherches, Etudes et Techniques (DRET) for partial support of this work.

Reference :

1 VEGA-RIVEROS J.F. - JABBOUR K. : Review of motion analysis techniques IEEE proceedings, Vol 136, Pt1, N°6 (1989)
2 PRAZDNY K. : On the information in optical flows Computer vision, graphics, and image processing 22, 239-259 (1983)
3 SKLANSKY J. : On the Hough technique for curve detection - IEEE transactions on computers Vol 27, N°10, pp 923 (1978)
4 GONZALEZ R.C. : Digital image processing - second edition, Addison Wesley (1987)
5 BOUTHEMY P. : A new scheme for motion computation along contours in image sequences Second image symposium-Nice(1986)
6 OTSU N. : A threshold selection method from gray-level histograms IEEE Transaction on systems,man and cybernetics - VOL SMC-9, n°1 (1979)
7 NAGEL H.H. : On the estimation of the optical flow Artificial Intelligence 33 pp 299-324 (1987)
8 SCHUNCK B.G. : The image flow constraint equation - Computer vision, graphics, and image processing 35, 20-46 (1986)
9 RAJALA S.A. - RIDDLE A.N. - SNYDER W.E.: Application of the one-dimensional Fourier transform for tracking moving objects in noisy environments Computer vision, graphics, and image processing 21, 280-293 (1983)