Region segmentation according to motion-based criteria

Javier Santillana Rivero, Patrick Bouthemy

UNITÉ DE RECHERCHE
INRIA-RENNES

Institut National
de Recherche
en Informatique
et en Automatique

Domaine de Voluceau
Rocquencourt
B.P. 105
78153 Le Chesnay Cedex
France
Tél.:(1) 39 63 55 11

Rapports de Recherche

N° 629

REGION SEGMENTATION ACCORDING TO MOTION-BASED CRITERIA

Santillana RIVERO
Patrick BOUTHEMY

Février 1987

REGION SEGMENTATION ACCORDING TO MOTION – BASED CRITERIA
SEGMENTATION EN REGIONS SELON DES CRITERES DE MOUVEMENT

Javier SANTILLANA RIVERO and BOUTHEMY Patrick

Publication Interne n°340 - Janvier 87-36 Pages

IRISA / INRIA
Campus de Beaulieu
35042 Rennes Cedex

Résumé

Ce rapport présente une nouvelle approche pour réaliser la segmentation en régions de mouvement apparent différent dans une séquence d'images. Une modélisation du mouvement est introduite. Le schéma défini ne s'appuie en fait que sur une observation partielle de mouvement. Le critère d'homogénéité se traduit par un test de rapport de vraisemblance et s'inscrit dans une structure de segmentation de type split – and – merge. La partition en régions homogènes au sens du mouvement est mise – en – oeuvre de façon hiérarchique . Plus précisément, un premier modèle de type mouvement constant est pris en compte et dans un deuxième temps, une étape complémentaire est appliquée en considérant un modèle de type linéaire. Des résultats obtenus sur des données synthétiques bruitées et des images réelles sont présentés.

Summary

This report presents a new approach to the motion – based segmentation problem. The designed formalism includes some motion model and relies on an explicit partial motion information through a stochastic approach. Indeed, it comes to the computation of some likelihood ratio test embedded in a split – and – merge procedure. Hence, regions are structured according to motion homogeneity criteria considered in a hierarchical way. We mean that segmentation starts with a motion model as simple as possible (i.e. constant apparent velocity), and after a complete iteration cycle, a more elaborated motion model can be taken into account (i.e. linear one). Results obtained with synthetic noisy data and real images are shown.

Table des Matières

Chapitre 1

INTRODUCTION: PROBLEM STATEMENT AND RELATED WORK

Optic flow determination is an important cue for mobile robot vision and others applications such as meteorology, image coding, biology,... [HUA-83,NAG-86]. In a framework for determining optic flow fields from image sequences, a key-problem to be solved can be stated as a spatio-temporal segmentation problem. Obtaining such a segmentation may come to determine moving contours, [BOU-86], [BUX-84], or/and to differentiate between moving and fixed regions and between moving regions with different motion. This work deals with image segmentation into regions according to motion-based criteria.

Such a process can be an attractive pre-requirement to obtain good and consistent optic flow fields. It enables to delimit image areas where some smoothing procedure can be effectively performed to estimate such fields, without worrying about potential discontinuities. A spatio-temporal segmentation process can also be perceived as a tool which can be directly applied to problems related to 3D-dynamic scene analysis (such as moving object tracking, obstacle avoidance...) or to TV-broadcasting domain (such as HDTV transmission).

Several papers address this problem from different points of view. JAIN ET AL. [JAI-79] used spatio-temporal criteria and a heuristic to segment images into moving and fixed areas. Their method is based on change detection by considering temporal differences in image sequences. The object masks (i.e, moving regions) are built up over time by tracking intensity changes and using some heuristic.

A different way to segment images using spatio-temporal information was developed by HSU ET AL. [HSU-84]. They have constructed a change detector based on a likelihood test where the spatial gray value variations were modeled using constant, linear or quadratic approximation.

For instance, for the constant approximation case, the likelihood ratio test is given by the following hypotheses :

- H_0: g_1 and g_2 come from the same Gaussian distribution $N(\mu_0,\sigma^2)$
- H_1: g_1 and g_2 come from different Gaussian distributions $N(\mu_1,\sigma^2)$ and $N(\mu_2,\sigma^2)$

where g_1 and g_2 are the observed values from the two test areas in images I_t and I_{t+1} respectively.

Two remarks can be made. First, the observed values are composed by two sets of spatial

informations which are somehow merely compared ; second, only change detection is taken into account without any higer – level interpretation (data structuration, motion discrimination).

Another way to cope with spatio – temporal segmentation is to assume that optic flow fields have already been determined. Then, for instance, segmentation can be performed applying a Laplacian operator over the optic flow components v_x and v_y [THO – 85]. An other approach has proposed a least square identification of a motion model, by the use of a pyramid linking algorithm, combined with a split – and – merge procedure [HAR – 85]. However, the problem is still present. Indeed, how to obtain consistent optic flow fields without previous segmentation? In this report, we present a new approach to the motion – based segmentation problem. The criterion is based on an explicit **partial** motion information. It is embedded in a split – and – merge framework and expressed by a likelihood ratio test. Then regions are structured according to motion homogeneity criteria. An interesting point of this approach is that segmentation can be executed in a hierarchical way. We mean that segmentation starts with a motion model as simple as possible (i.e. constant velocity) and after a complete iteration cycle, a more elaborated motion model can be considered.

The remainder of this report is organized as follows. In chapter 2, our likelihood approach including a partial – information – based criterion will be described for the constant velocity model. Chapter 3, will show how the split – and – merge algorithm has been implemented and used in our application. Various results with synthetic and real data obtained with the constant motion model are shown in chapter 4. In chapter 5, the motion model will be completed considering a first order velocity gradient term and results will be shown. Finally, chapter 6, contains conclusion and future developments will be outlined.

Chapitre 2

SPATIO – TEMPORAL SEGMENTATION CRITERION BASED ON A CONSTANT – MOTION MODEL

2.1 Motion model and observation function

We assign to a spatio – temporally consistent region the following motion model:

constant $\underline{T}$ + *noise*

Assuming this motion model, we will derive the spatio – temporal homogeneity criterion for a given region as follows. According to the well – known aperture problem, [SCH – 86], [HIL – 84], some local process only yields one component of the velocity vector $\underline{v} = (dx/dt, dy/dt)$ at one image point (x,y) : the component parallel to the image spatial gradient $\underline{\nabla f}(x,y)$, which can be written as $\underline{v}(x,y).\underline{\nabla f}(x,y)$. Let us consider the expression given by :

$$e(x,y) = [\underline{T} - \underline{v}(x,y)].\underline{\nabla} f(x,y) \qquad (1)$$

We assume that e is a random variable, whose distribution is Gaussian of mean 0 and variance σ^2. If $\underline{T} = (a,b)$, let us denote e as $e_{a,b}$. Let us precise that $\underline{\nabla f} = (\partial f/\partial x, \partial f/\partial y)$.

We can infer from the image flow constraint equation, derived by differential techniques and described for instance in [SCH – 86], $\underline{v}.\underline{\nabla f} = -\partial f/\partial t$ (2), the following form for $e_{a,b}$, by substituting their components for vectors :

$$e_{a,b} = a.\frac{\partial f}{\partial x} + b.\frac{\partial f}{\partial y} + \frac{\partial f}{\partial t} \qquad (3)$$

Two points can be discussed concerning this basic equation (2). The relation is known to behave poorly in the neighborhood of occlusions, [ADI – 85], [MIT – 84], due to continuity assumptions related to differentiation which do not hold for such areas. This would certainly lead spatio – temporal homogeneity criterion to diverge. This is not particularly unsuitable as these areas are precisely discontinuities we want to detect.

On the other hand, at points interior to untextured regions, relation (2) and then expression (3) provide weak information. Hence the spatio – temporal homogeneity criterion will tend to zero. This

situation is still suitable because an uniform spatial region is also a spatio – temporal homogeneous region apart from singular cases.

2.2 Likelihood ratio test

The spatio – temporal segmentation criterion is achieved by the use of a log – likelihood ratio test. Let us consider two successive images from a sequence and a given area A in the reference image. Two hypotheses can be acting concerning function e within area A :

- Hypothesis H_0 : for each point within A, $e := e_{a0,b0}$
- Hypothesis H_1 : area A is divided into two sub – areas A_1 and A_2 such that :

  for each point in A_1; $e := e_{a1,b1}$

  for each point in A_2; $e := e_{a2,b2}$

  with $(a_1,b_1) \neq (a_2,b_2)$.

The problem now is how to select one hypothesis versus the other one. With each hypothesis H_0, H_1 is associated a likelihood function L_0, L_1. We infer the likelihood function from the expression of the joint probability density function of the *e's* within the considered area. We assume that the *e's* are independent. Recall that they come from some Gaussian distribution. Let us denote $\Theta = (a_0,b_0,a_1,b_1,a_2,b_2)$. After some mathematical developments reported in Appendix A, the log – ratio ξ of L_1 and L_0 is given by :

$$\xi(\Theta) = \frac{1}{2\sigma^2}\left[\sum_{(x,y)\in A} e_{a0,b0}(x,y)^2 - \sum_{(x,y)\in A_1} e_{a1,b1}(x,y)^2 - \sum_{(x,y)\in A_2} e_{a2,b2}(x,y)^2\right] \quad (4)$$

The test which will enable to select hypothesis H_0 versus hypothesis H_1 is expressed by :

$$\max_{a_1b_1a_2b_2} \min_{a_0b_0} \xi(\Theta) \leqq \lambda \quad (5)$$

Clearly, this means that if $\xi \leqq \lambda$ for proper optimal values of parameters $a_i, b_i, i = 0,1,2$, area A is declared to be spatio – temporally homogeneous. Optimal values $\hat{\Theta}$ are obtained by solving the following linear system for $i = 0,1,2$:

$$\begin{cases} \dfrac{\partial \xi}{\partial a_i} = 0 \\ \dfrac{\partial \xi}{\partial b_i} = 0 \end{cases} \quad (6)$$

Closed forms for the analytical solutions of (6) are given in Appendix A.

Chapitre 3

SPLIT – AND – MERGE ALGORITHM

The original split – and – merge algorithm such as it was developed by HOROWITZ and PAVLIDIS [HOR – 76] cannot be directly applied to our problem because of the nature of the homogeneity test we have defined.

3.1 Splitting stage

Our split procedure processes as follows.

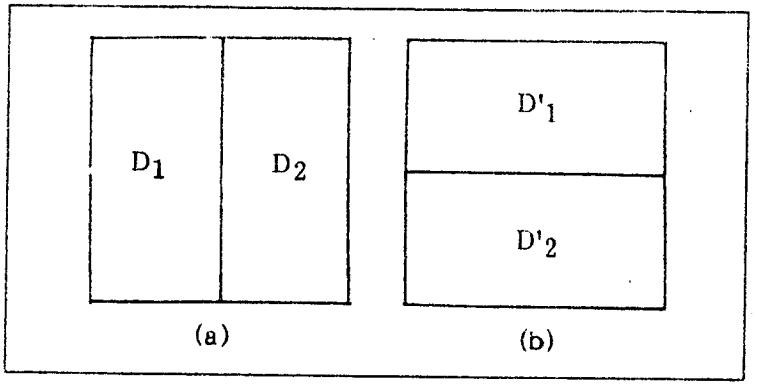


Figure 1. Considered complementary configurations of block B_i, (a) vertical one, (b) horizontal one

Let $I(x,y)$ be the picture domain of N by N elements and let us consider its initial partition into blocks B_i of size S_i (n by n elements). According to criterion (5), a block is then subdivided into sub – blocks of size s_i if and only if,

$$\xi(\hat{\Theta}) > \lambda \qquad (7)$$

where ξ is the likelihood ratio defined in (4) and λ is a predetermined threshold. Let us precise how ξ is computed. In order to test for alternative hypotheses given a block B_i, two complementary configurations are defined as shown in Fig.1. The following hypotheses are considered for each configuration as explained in Chapter 2, identifying A to B_i, A_1 to D_1 or D'_1, A_2 to D_2 or D'_2

The additional size constraint was included to avoid merging two regions of too different sizes at initial iterations. This means that η_k is initially set to $2s_i$ (twice the size of the smallest regions) and doubled at the end of each merge iteration.

When two regions R_i and R_j are merged together, the adjacency list is updated and all lists and arrays are updated as follows:

$$S_f^i = S_f^i + S_f^j$$
$$S_f^j = 0$$
$$Region_j = Region_i$$
$$\sum_{R_i} G_x G_y = \sum_{R_i} G_x G_y + \sum_{R_j} Gx G_y$$
$$\sum_{R_i} G_x G_t = \sum_{R_i} G_x G_t + \sum_{R_j} G_x G_t$$
$$\sum_{R_i} G_y G_t = \sum_{R_i} G_y G_t + \sum_{R_j} G_y G_t$$
$$\sum_{R_i} G_x^2 = \sum_{R_i} G_x^2 + \sum_{R_j} G_x^2$$
$$\sum_{R_i} G_y^2 = \sum_{R_i} G_y^2 + \sum_{R_j} G_y^2$$

Otherwise the adjacency between R_i and R_j is inhibited until the end of the merge procedure.

The merge procedure stops when η_k is greater than the image size. Next chapter shows results obtained using this approach.

- H_0: the same constant motion model M_0 is acting within B_i
- H_1: D_1 (resp.D'_1) and D_2 (resp.D'_2) correspond to two different constant motion models M_1 and M_2.

Let us denote L_0 (resp.L'_0), L_1 (resp.L'_1), the likelihood functions corresponding to hypotheses H_0, H_1.

Two log-ratios are evaluated within block B_i, using proper optimal values $\hat{a}_i, \hat{b}_i, i = 0,1,2$; these ratios are given by

$$\xi_V = \ln \frac{L_1}{L_0}; \qquad \xi_H = \ln \frac{L'_1}{L'_0}$$

And ξ is then given by

$$\xi = \max(\xi_V, \xi_H)$$

If a block is declared to be non uniform ($\xi > \lambda$), it is subdivided into sub-blocks b_i of size s_i. The same test is performed again. However, instead of splitting again these sub-blocks b_i, when the condition given by (7) is true, they are only declared as non uniform blocks. This size limitation was set, because the likelihood ratio test is not significant when the regions are too small. Therefore, the split stage consists of only two steps.

As a result of this split procedure, the image is initially separated into locally uniform and non-uniform blocks.

3.2 Merging stage

Now, the second stage is the merge procedure. First, a region adjacency list is created as in [HOR-74]. A region array and a final size array S_f are initialized. For each region R_i, eight informations $(X, Y, S_f, \Sigma G_x G_y, \Sigma G_x G_f, \Sigma G_y G_f, \Sigma G_x^2, \Sigma G_y^2)$ have been stored, where Σ means sum over points p of region R_i, $X = \{x_p\}$ and $Y = \{y_p\}$.

The same kind of approach is considered for the merge procedure. At the kth merge iteration, two adjacent regions R_i and R_j, identified respectively to A_1 and A_2 and their eventual union to A, can be merged together according to criterion (5), if and only if

$$i) \; \xi(\hat{\Theta}) \leq \lambda \qquad (8i)$$

$$ii) \; S_f^i + S_f^j \leq \eta_k \qquad (8ii)$$

Chapitre 4
SEGMENTATION RESULTS

4.1 Data description

In order to evaluate the performance of the designed segmentation algorithm, three image sequences were used. The first sequence, shown in Fig. 2, which is the simplest one, consists of two computer–generated images. The luminance function I of the disk, of radius r and centered on (h,k), is defined by

$$I(x,y) = L - \frac{L-b}{r^2}\left[(x-h)^2 + (y-k)^2\right] + offset$$

where L is the max gray level within the disk and b the background gray level, to which some uniform noise has been added.

It moves parallel to the image plane at constant velocity, $\underline{v} = (\Delta x, \Delta y) = (5,5)$, over the noisy fixed uniform background.

To simulate more complex motion situations, the second sequence of Fig.3 represents the same kind of disk that, in this example, moves parallel to the axis of view (hence, a dilatation is observed). Moreover it partially occludes an other disk that is shifted by a translation $\underline{v} = (3,3)$ parallel to the image plane.

The third sequence, presented in Fig.4, includes two images of a printer acquired by a CCD camera. Only the printer has been moved from one image to the next, camera and background remain fixed.

4.2 Computation details

4.2.1 Image gradient computation

The spatial derivatives of the image intensity function f, $G_x = \partial f/\partial x$ and $G_y = \partial f/\partial y$, are estimated by mask convolution as follows :

$$G_x = \frac{1}{3} f * M_x \qquad \text{and} \qquad G_y = \frac{1}{3} f * M_y \qquad \text{with}$$

$$M_x := \begin{bmatrix} 1 & 0 & -1 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{bmatrix}; \qquad M_y := \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix}$$

and, the temporal derivative $G_t = \partial f / \partial t$ is computed as a simple temporal difference,

$$G_t(i,j,t) = f(i,j,t) - f(i,j,t+1)$$

Then the observed processus e corresponding to the constant motion model must be rewritten as follows,

$$e_{a,b} = aG_x + bG_y + \nu G_t$$

with $\nu = 2$. ν is used to keep the same scale factor for spatial and temporal gradients.

4.2.2 Threshold determination

The homogeneity test was developed assuming that the variance of the random variable e is always the same over considered areas A, A_1 and A_2. It is an approximation to simplify the computation load. Therefore, since the variance is not available as a possible indication to "modulate" the threshold level, we have looked for an other parameter.

It seems reasonable that, when two areas of sizes S_i and S_j respectively are considered in the context of the homogeneity criterion described in section 2, the threshold value would be proportional to the size of the smallest region. Then, we have introduced a weighting factor κ proportional to the size of the smallest region. The likelihood ratio test is then expressed, in the case of hypothesis H_0 acting, as $\xi \leqq \kappa \lambda$.

4.3 Results

In Fig.5, the likelihood ratio function ξ is drawn in the case of the one-disk image data. Two examples of a complete iteration cycle of the merging stage is presented in Fig.6a-b.
Figure 7 shows the initial and final segmentation for each test sequence. Size of initial blocks B_i is $S_i = 16x16$, and for sub-blocks b_i, $s_i = 4x4$. The elementary 4x4 non uniform blocks are "painted" in black. They have not been taken into account during the merge procedure.

Chapitre 5

SPATIO – TEMPORAL SEGMENTATION CRITERION BASED ON A LINEAR – MOTION MODEL

5.1 Motion model

For regions which have not been declared as fixed regions, i.e. whose representative parameter values $(\hat{a}_0,\hat{b}_0) \neq (0,0)$, a second merge stage is attempted. It is based on a more complex motion model, indeed a linear – motion model; that – is – to – say of the kind
"constant velocity gradient (Dc,Dl) + noise".

Let (x_g,y_g) be the centroid of a given region R for which $(\hat{a}_0,\hat{b}_0)\neq(0,0)$. For each point (x,y) of R, let us write the first order expansion of the velocity vector $\underline{v}(x,y)$ referred to (x_g,y_g):

$$\underline{v}(x,y) = \underline{v}(x_g,y_g) + (x-x_g)\partial\underline{v}/\partial x + (y-y_g)\partial\underline{v}/\partial y + o^2$$

Let us introduce $\underline{Dc}=(\alpha,\beta)$ the model velocity gradient with respect to x and $\underline{Dl}=(\gamma,\delta)$ w.r.t. y. Let us denote $\underline{v}_g=(a_g,b_g)=\underline{v}(x_g,y_g)$. This leads to the new following expression for the variable e primary defined in (1):

$$e(x,y) = [\underline{v}_g + (x-x_g)\underline{Dc} + (y-y_g)\underline{Dl} - \underline{v}(x,y)] \cdot \underline{\nabla f}(x,y) \qquad (9)$$

where only first order terms are considered. Considering again the relation $\underline{v}.\underline{\nabla f} = -\ \partial f/\partial t$, e is then given by :

$$e = a_g.\partial f/\partial x + b_g.\partial f/\partial y + \partial f/\partial t + \alpha(x-x_g)\partial f/\partial x + \beta(x-x_g)\partial f/\partial y + \gamma(y-y_g)\partial f/\partial x + \delta(y-y_g)\partial f/\partial y \qquad (10)$$

Therefore e depends on six parameters $\Lambda=(a_g,b_g,\alpha,\beta,\delta,\gamma)$ with respect to which the new criterion will be to optimized. The same mathematical development is applied as previously, which leads to a similar likelihood ratio test. Details concerning this likelihood ratio ζ and the computation of corresponding optimal parameter values $\hat{\Lambda}$ are described in Appendix B. Of course, the computational load is more important considering this motion model than the previous one. However, it is minimized in the sense that this stage takes place after the constant motion model stage; hence, only few and appropriate regions are considered.

Let us point out that this linear model is rather a comprehensive one. For instance, it includes the

In sequence 2 (Fig. 3), the object moving perpendicularly to the image plane has been segmented into five regions, Fig.7e. The reason of this over-segmentation is that the constant motion model cannot fit this kind of motion, yet the way it has occurred is coherent. Somehow, parameters $\hat{a}_0$ and $\hat{b}_0$ represent the mean motion vectors for those sub-regions.
Notice that for the example of sequence 3, (Fig.7f), the whole fixed background has been well merged into one region, despite different spatial regions can be observed.

Improved results are obtained when the non uniform blocks are taken into account by the merge procedure, as shown in Fig. 8. The performance improvement is certainly due to the criterion sensibility to noise when the blocks are too small (i.e. 4x4). So, several blocks that are classified as non uniform blocks by the splitting procedure can be recovered by an other region at the merge step.
Table 1 gives, for some significant regions of Fig.8e, the parameter values $\hat{a}_0$ and $\hat{b}_0$ estimated when the given region has been built up. It can be noted that these parameters are quite consistent with the true motion.
Other experiments have been performed once data have been pre-processed using some Gaussian filter. However no significant differences have been observed.

Further comments can be made concerning those elementary non-uniform blocks. They correspond in particular to areas which are appearing or disappearing, from the first considered image to the next one, in the neighbourhood of occlusion contours. We have let presented final partitions just as they are on purpose, i.e. some labeled significant spatio-temporal regions and a set of resulting unstructured elementary non-uniform blocks. Indeed, the further process which could be undertaken depends on the seeked aim. For example, those blocks could be aggregated apart from other regions, if some particular information is searched concerning appearing and disappearing areas in the image sequence. On the contrary, they could be merged with proper adjacent significant regions, according to some spatial criterion upon intensity function like min-max, variance, .. as exposed for instance in [GAG-86].

We have mentioned that an over-segmentation has happened in the case of the disk which undergoes some dilatation, (see Fig.7e-8e). In order to cope with such varying apparent motion, a more complete motion model is then introduced; a linear motion model, as explained in the next section.

Chapitre 6
CONCLUSION

We have described a region segmentation method according to motion – based criteria. It relies on a partial motion information observation and follows a stochastic approach. Spatio – temporal homogeneity decision comes to computing a proper likelihood ratio test based on some motion model. Two such models are considered in a hierarchical manner within a split – and – merge procedure; first a constant motion model, then a linear one. Let us outline that other subsequent complex models could also be handled through the same formalism, e.g. quadratic or 3D models.

Future developments include the design of a complete cooperative framework for spatio – temporal segmentation of successive images integrating both the contour – based approach described in [BOU – 86],and this region – based approach. Moreover, such a scheme will be applied to navigation problems in the context of robot vision.

case of the combination of a translation, a rotation in the image plane and an uniform dilatation considered in [HAR 85], which corresponds to $\alpha = \delta$ and $\beta = -\gamma$.

5.2 Results

Results concerning the "two-disks" sequence of Fig.3 are shown in Fig.9. The second merge stage based on the linear motion model started from the partition achieved with the constant motion model presented in Fig.8e. The five significant subregions previously obtained within the disk, which undergoes some dilatation (see regions 2 to 6 in Fig.8e), well become one unique homogeneous moving region (region 1 in Fig.9). As the same threshold ($\lambda = 6$) is used, the merge actions are actually due to the supplementary information yielded by the linear motion model. On the other hand, no false merge action occurs.

The parameters $\hat{\Lambda}_0$ are given in Table 2 for the region 1 of Fig.9. Notice that they are consistent

[SCH – 86] B.G. SCHUNCK, *The image flow constraint equation*, Computer Vision, Graphics and Image Processing, vol. CVGIP – 35, 1986, pp.20 – 46.

REFERENCES

[ADI – 85] G. ADIV, *Determining three – dimensional motion and structure from optical flow generated by several moving objects*, IEEE Trans. on Pattern Analysis and Machine Intelligence, vol.PAMI – 7, no4, July 1985, pp.384 – 401.

[BOU – 86] P. BOUTHEMY, *Determining displacement fields along contours from image sequences*, Proc. Conf. Vision Interface '86, Vancouver, May 1986, pp. 350 – 355.

[BUX – 84] B.F. BUXTON AND H. BUXTON, *Computation of optic flow from the motion of edge features in image sequences*, Image and Vision Computing, vol.2, no2, May 1984, pp.59 – 75.

[GAG – 86] A. GAGALOWICZ AND O. MONGA, *A new approach for image segmentation* Proc. 8th Int. Conf. on Pattern Recognition, Paris, Oct. 1986, pp.265 – 267

[HAR – 85] R. HARTLEY, *Segmentation of optical flow fields by pyramid linking*, Pattern Recognition Letters, vol.3, July 1985, pp.253 – 262.

[HIL – 84] E.C. HILDRETH, *Computations underlying the measurement of visual motion*, Artificial Intelligence, vol.23, 1984, pp.309 – 354.

[HOR – 74] J.L. HOROWITZ AND T. PAVLIDIS, *Picture segmentation by a direct split – and – merge procedure*, Proc. 2nd Int. Conf. on Pattern Recognition, 1974, pp. 424 – 433.

[HSU – 84] 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.

[HUA – 83] T.S. HUANG (ED.), *Image Sequence Processing and Dynamic Scene Analysis*, NATO – ASI Series F2, Springer – Verlag, 1983.

[JAI – 79] R. JAIN, W.N. MARTIN AND J.K. AGGARWAL, *Segmentation through the detection of changes due to motion*, Computer Graphics and Image Processing, vol. CGIP – 11, 1979, pp.13 – 34.

[MIT – 84] A. MITICHE, *Computation of optical flow and rigid motion*, Proc. 2nd Workshop on Computer Vision: Representation and Control, Annapolis, 1984, pp.63 – 71.

[NAG – 86] H.H. NAGEL, *Image sequences – Ten (octal) years – From phenomenology towards a theoretical foundation*, Proc. 8th Int. Conf. on Pattern Recognition, Paris, Oct. 1986, pp.1174 – 1185.

[THO – 85] W.B. THOMPSON, K.M. MUTCH AND V.A. BERZINS, *Dynamic occlusion analysis in optical flow fields*, IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. PAMI – 7, no4, July 1985, pp.374 – 383.

APPENDIX A: Mathematical developments for the constant motion model

Let us consider the case of the constant motion model, and more precisely the function e defined in (3) as

$$e_{a,b} = a.G_x + b.G_y + G_t$$

Likelihood functions L_0 and L_1 associated with hypotheses H_0 and H_1 described in the main text, are given by :

$$L_0 = (2\pi\sigma^2)^{\frac{-(n1+n2)}{2}} \prod_{(x,y)\in A_1 \cup A_2} \exp\left\{-\frac{1}{2\sigma^2} e_{a0,b0}(x,y)^2\right\}$$

$$L_1 = (2\pi\sigma^2)^{\frac{-(n1+n2)}{2}} \prod_{(x,y)\in A_1} \exp\left\{-\frac{1}{2\sigma^2} e_{a1,b1}(x,y)^2\right\} \prod_{(x,y)\in A_2} \exp\left\{-\frac{1}{2\sigma^2} e_{a2,b2}(x,y)^2\right\}$$

n_1, (resp. n_2), is the number of points of A_1, (resp. A_2). The likelihood ratio test is given by L_1/L_0. Indeed, we take the logarithm of this ratio. Substituting for L_1 and L_0 from their above-mentioned expression, we get :

$$\xi(\Theta) = \frac{1}{2\sigma^2}\left[\sum_{(x,y)\in A} e_{a0,b0}(x,y)^2 - \sum_{(x,y)\in A_1} e_{a1,b1}(x,y)^2 - \sum_{(x,y)\in A_2} e_{a2,b2}(x,y)^2\right] \quad (A1)$$

with $A = A_1 \cup A_2$, and $\Theta = (a_0, b_0, a_1, b_1, a_2, b_2)$.

To obtain the optimal parameters, $\hat{a}_i, \hat{b}_i; i = 0,1,2$, that maximize the likelihood functions, we set the derivatives of ξ with respect to these parameters to zero :

$$\frac{\partial \xi}{\partial a_i} = 2\sum_{A_i}(a_i G_x^2 + b_i G_x G_y + G_x G_t) = 0$$

$$\frac{\partial \xi}{\partial b_i} = 2\sum_{A_i}(a_i G_x G_y + b_i G_y^2 + G_y G_t) = 0$$

Solving for the parameters gives

$$\hat{a}_i = \frac{\Sigma G_x G_y \ \Sigma G_y G_t - \Sigma G_y^2 \ \Sigma G_x G_t}{\Sigma G_x^2 \ \Sigma G_y^2 - \left[\Sigma G_y G_x\right]^2}$$

$$\hat{b}_i = \frac{\Sigma G_x G_y \ \Sigma G_x G_t - \Sigma G_x^2 \ \Sigma G_y G_t}{\Sigma G_x^2 \ \Sigma G_y^2 - \left[\Sigma G_y G_x\right]^2}$$

where Σ is a simplified notation for $\Sigma_{(x,y)\in A_i}, i = 0,1,2$, ($A_0$ denotes A by definition).

This implies that, in order to calculate $\hat{a}_i, \hat{b}_i$, needed quantities are :

$$\Sigma G_y G_x,\ \Sigma G_y G_t,\ \Sigma G_x G_t,\ \Sigma G_x^2 \text{ and } \Sigma G_y^2$$

These quantities are easily updated by simple addition after each merging operation.

APPENDIX B: Mathematical developments for the linear motion model

Let us consider the case of the more elaborated motion model, i.e. linear motion model, defined in section 4, and the likelihood ratio test $\zeta(a_g, b_g, \alpha, \beta, \gamma, \delta)$ associated with it. It is straightforward to derive ratio ζ from expressions (A1) and (10).

To obtain the optimal parameters $\hat{a}_{gi}, \hat{b}_{gi}, \hat{\alpha}_i, \hat{\beta}_i, \hat{\gamma}_i, \hat{\delta}_i, (i = 0,1,2)$ that maximize the likelihood functions, we set the derivatives of ζ with respect to these parameters to zero :

$$\frac{\partial \zeta}{\partial a_{gi}} = 0; \qquad \frac{\partial \zeta}{\partial \beta_i} = 0;$$
$$\frac{\partial \zeta}{\partial b_{gi}} = 0; \qquad \frac{\partial \zeta}{\partial \gamma_i} = 0;$$
$$\frac{\partial \zeta}{\partial \alpha_i} = 0; \qquad \frac{\partial \zeta}{\partial \delta_i} = 0.$$

Then, the optimal parameters are given by the solution of the resulting set of linear equations :

$$\sum_{(x,y)\in A_i} \Omega\Omega^T \Lambda_i^T = - \sum_{(x,y)\in A_i} G_t\, \Omega$$

where Ω and Λ are given by

$$\Omega^T = \left[G_x, G_y, (x - x_g) G_x, (y - y_g) G_x, (x - x_g) G_y, (y - y_g) G_y\right]$$

and

$$\Lambda = \left[a_g, b_g, \alpha, \beta, \gamma, \delta\right].$$

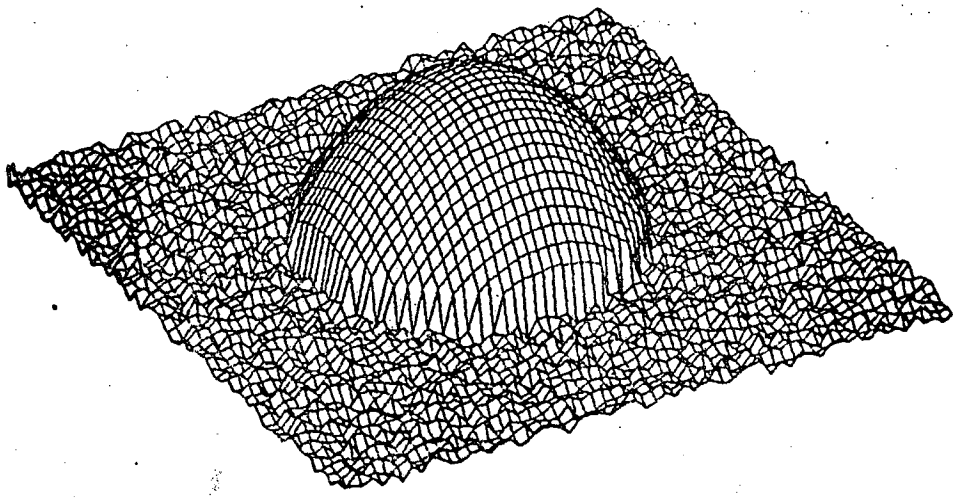


Figure 2 : First image of the sequence including only one disk (3D-graphic representation of the intensity function)

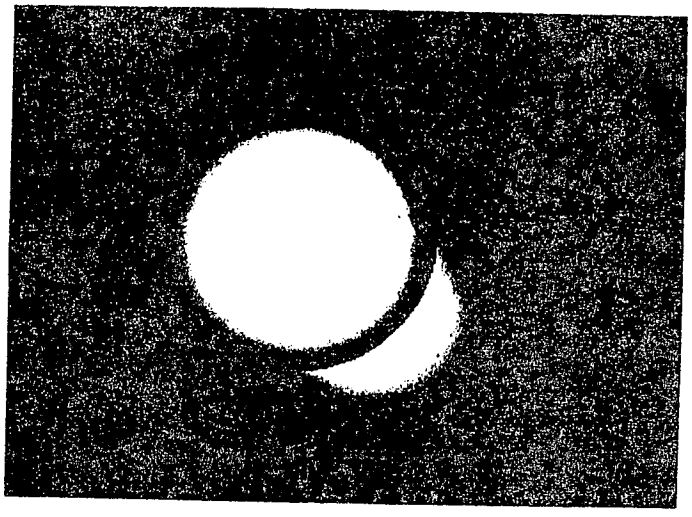
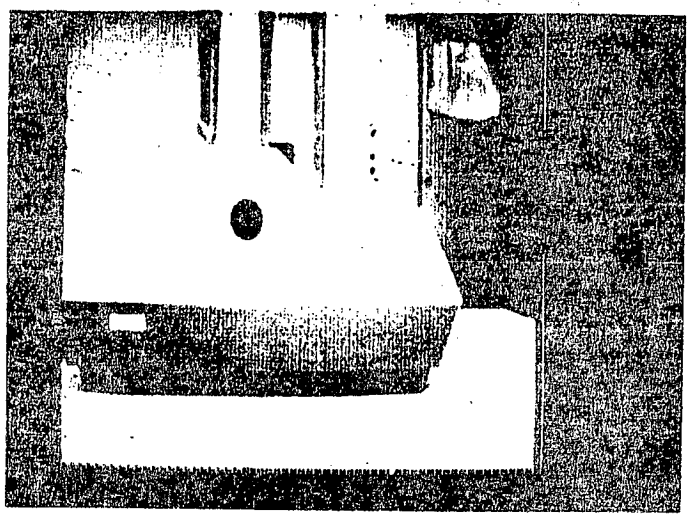
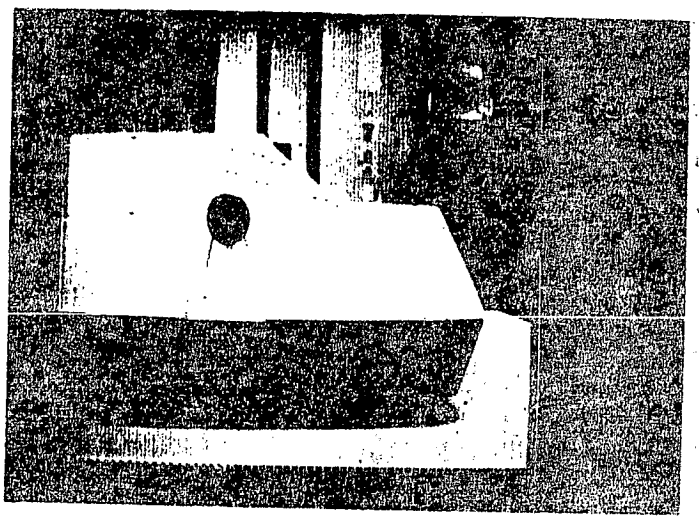


Figure 3 : First image of the sequence of the "two disks"



(a)



(b)

Figure 4 : "Printer" sequence
(a) first image, (size 256x256)
(b) second image

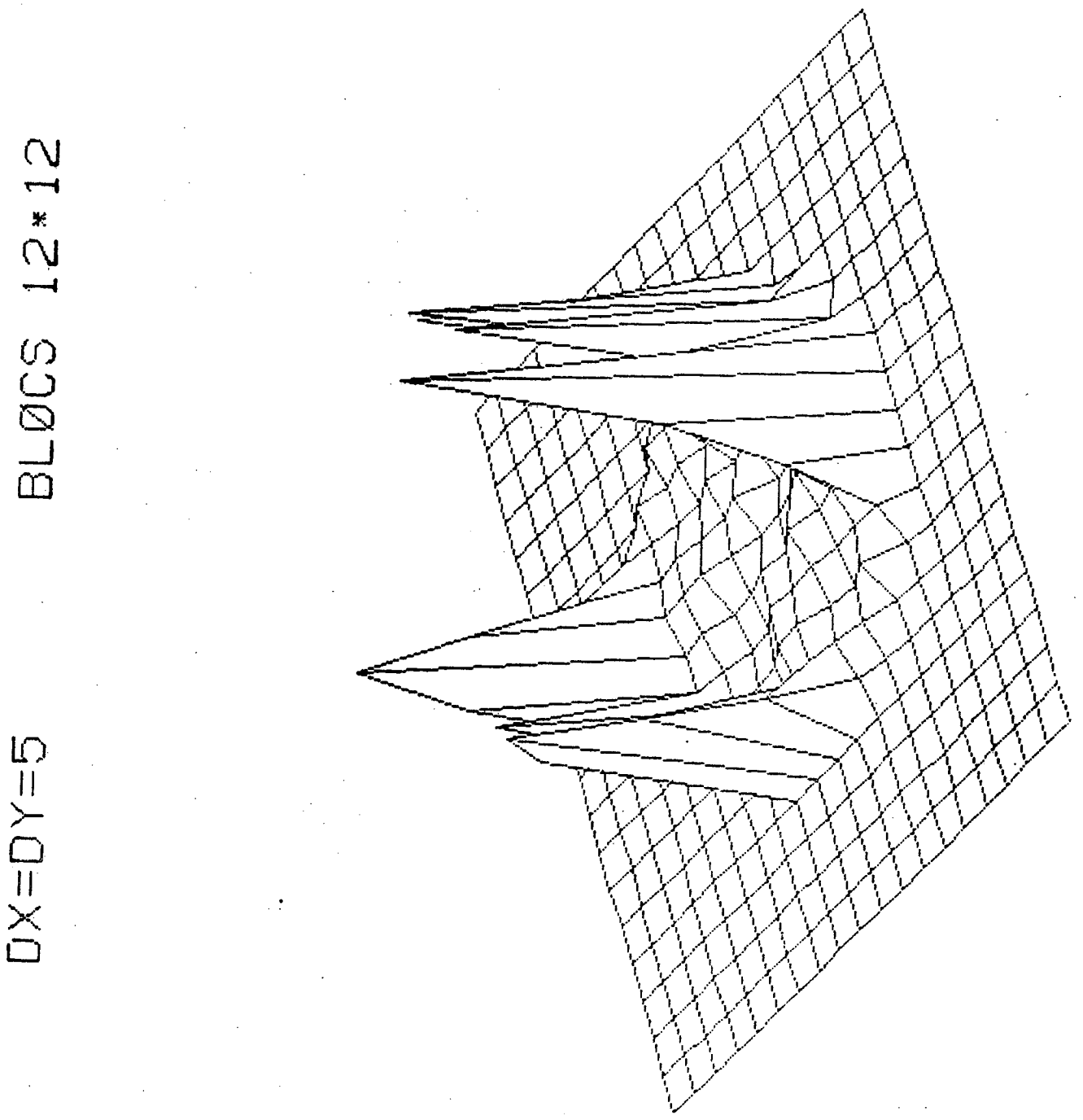


Figure 5 : Tracé du rapport de vraisemblance ξ pour un exemple similaire **à celui de la figure 2** (fond non bruité).

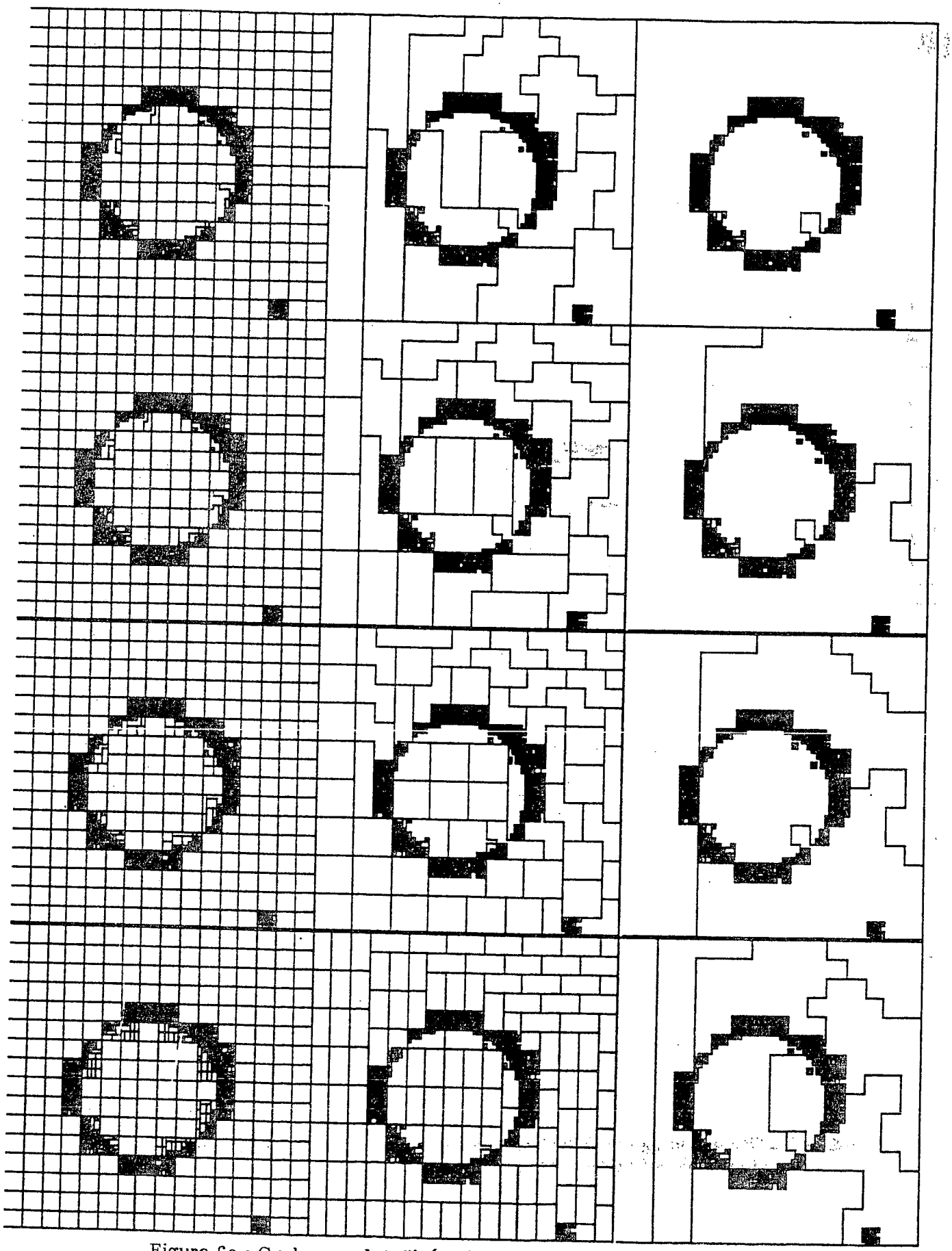


Figure 6a : Cycle complet d'itérations de la procédure "merge" correspondant au cas de la figure 2 (blocs 4x4 non-uniformes non pris en compte)

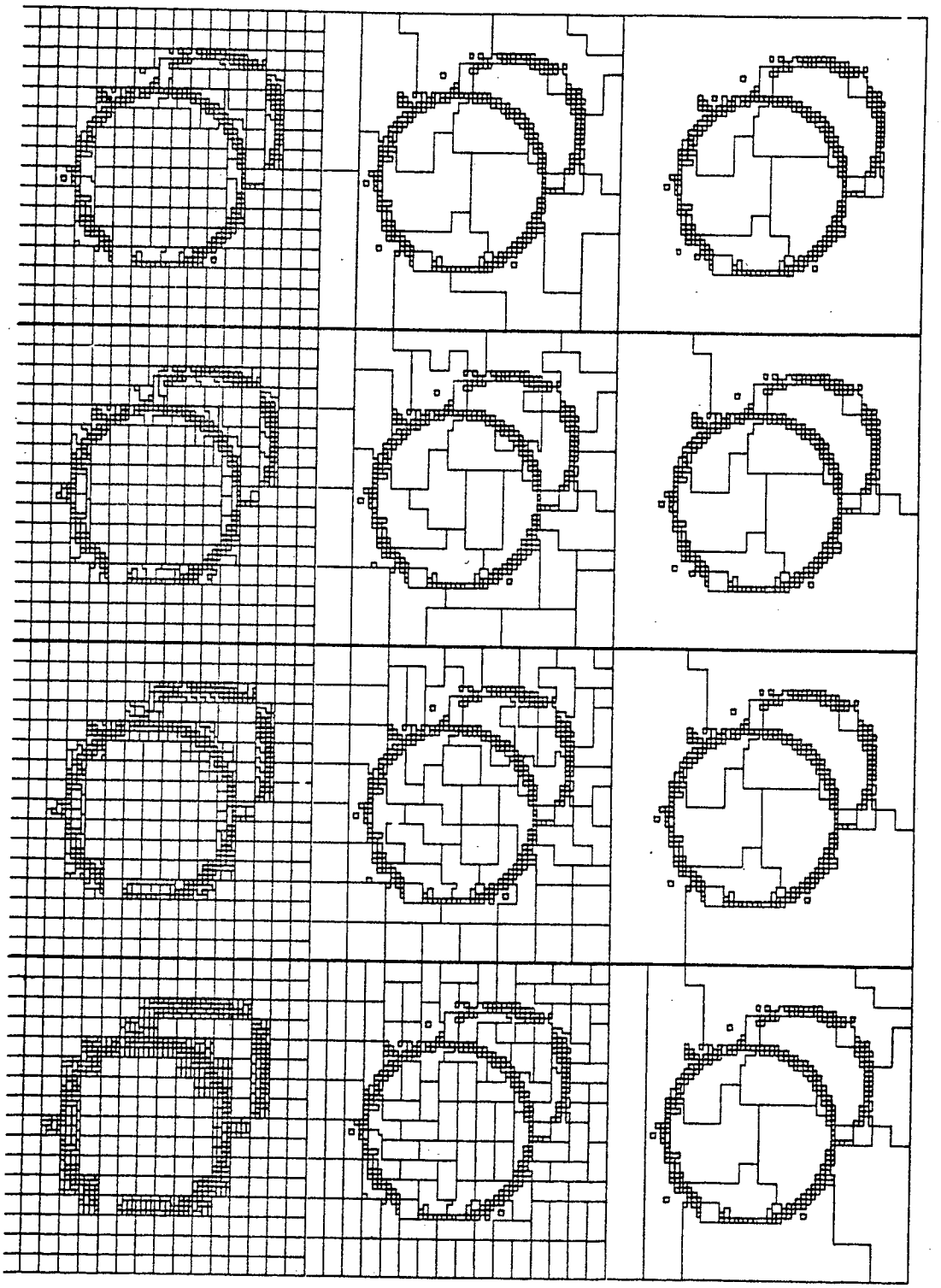
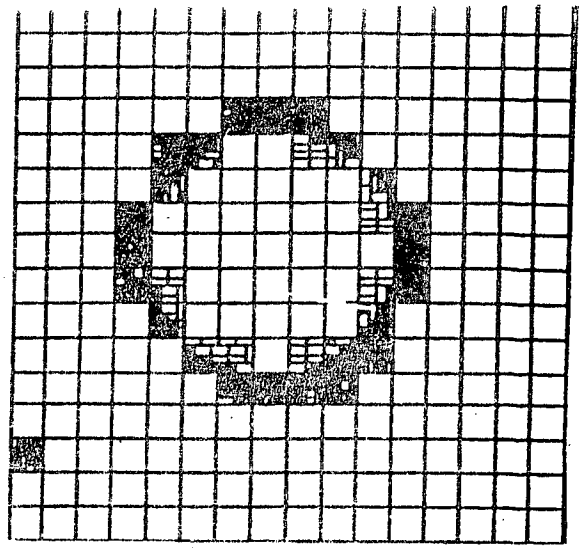
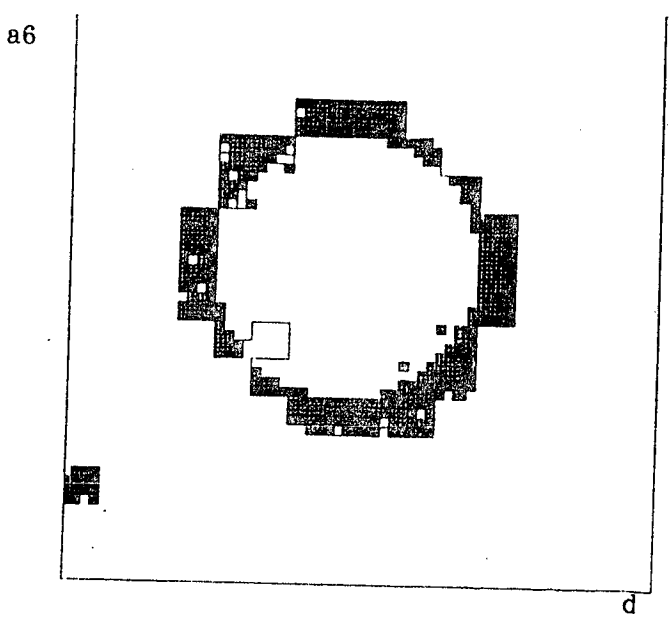


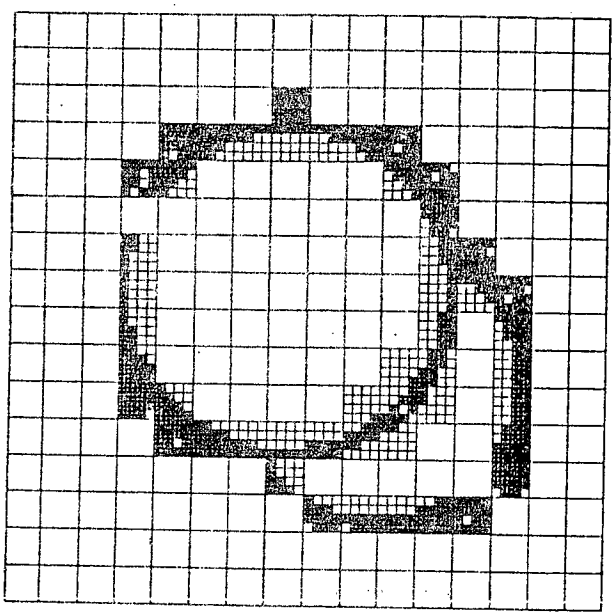
Figure 6b : Idem pour le cas de la figure 3 (blocs 4x4 non uniformes pris en compte)



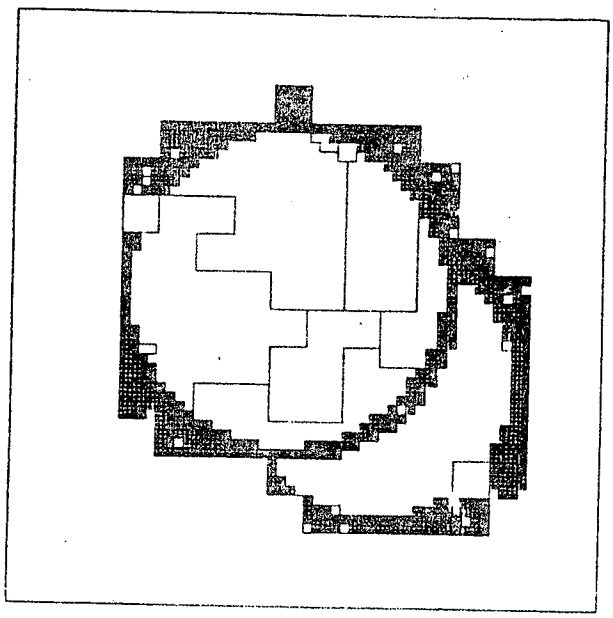
a



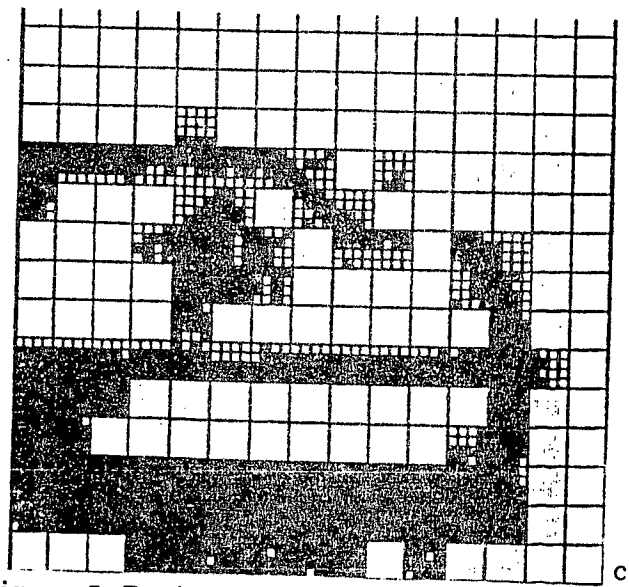
d



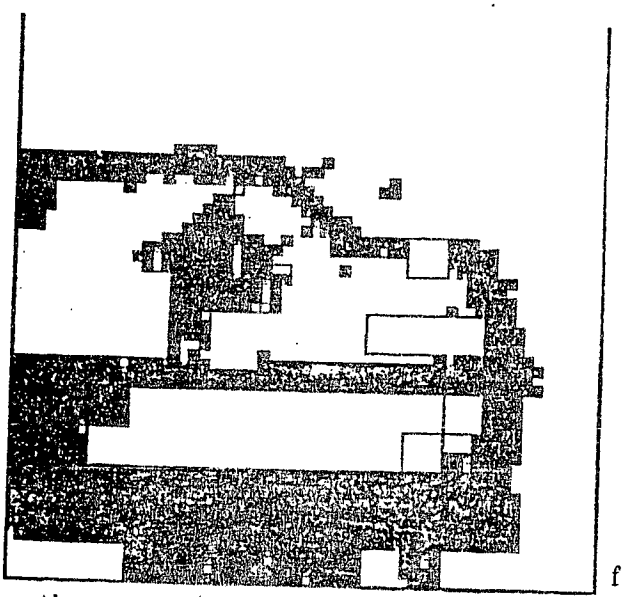
b



e

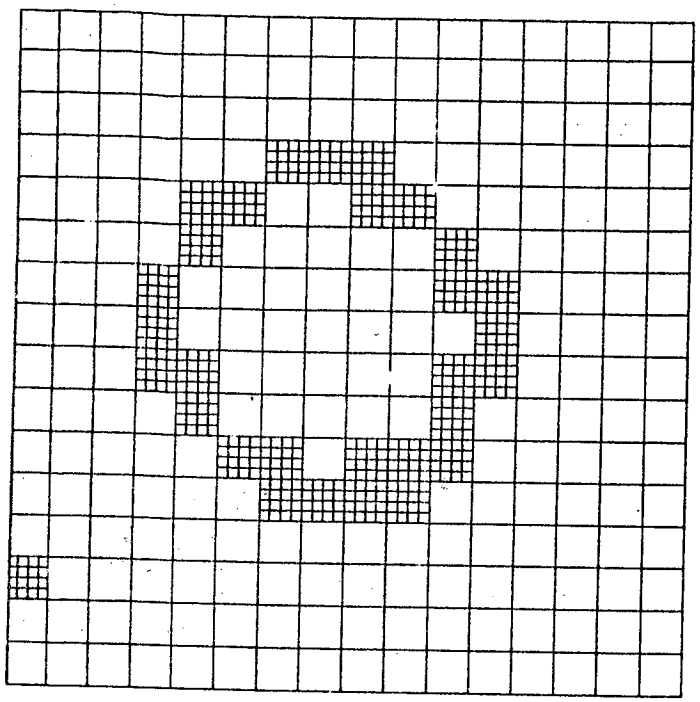


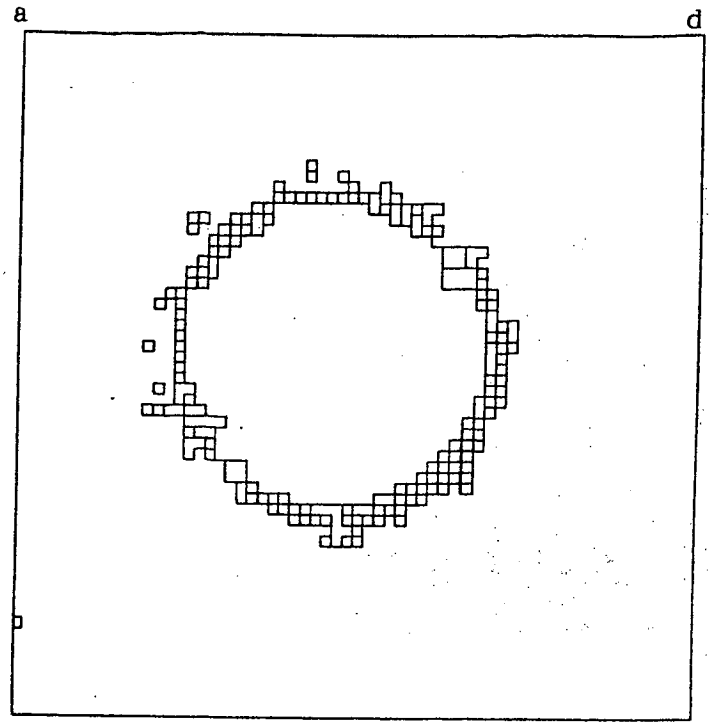
c

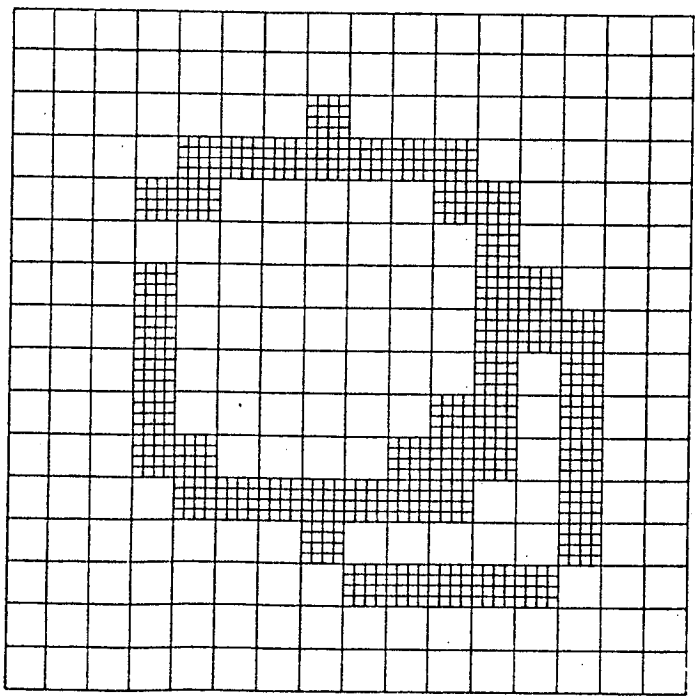


f

Figure 7: Region segmentation based on the constant motion model (black non uniform 4x4 blocks are not taken into account)
a,b,c : initial segmentation after the split stage with threshold λ = 6,6, 12 respectively
d,e,f : final partition after the merge stage using the same threshold λ







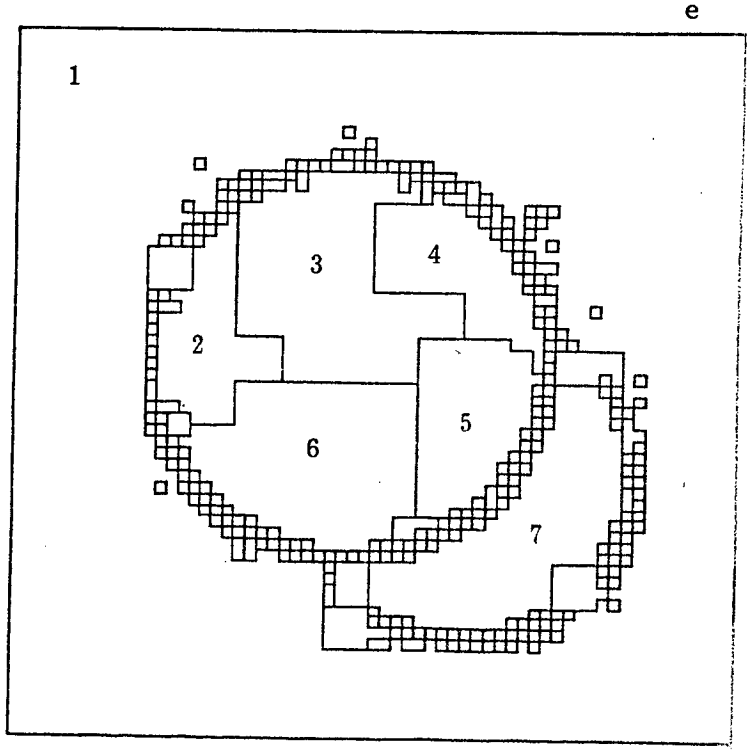


Figure 8 : Region segmentation with the constant motion model (elementary 4x4 non-uniform blocks are included in the merge procedure)
a,b,c : initial segmentation (same thresholds as in figure 7) after the split stage
d,e,f : final segmentation after the merge stage

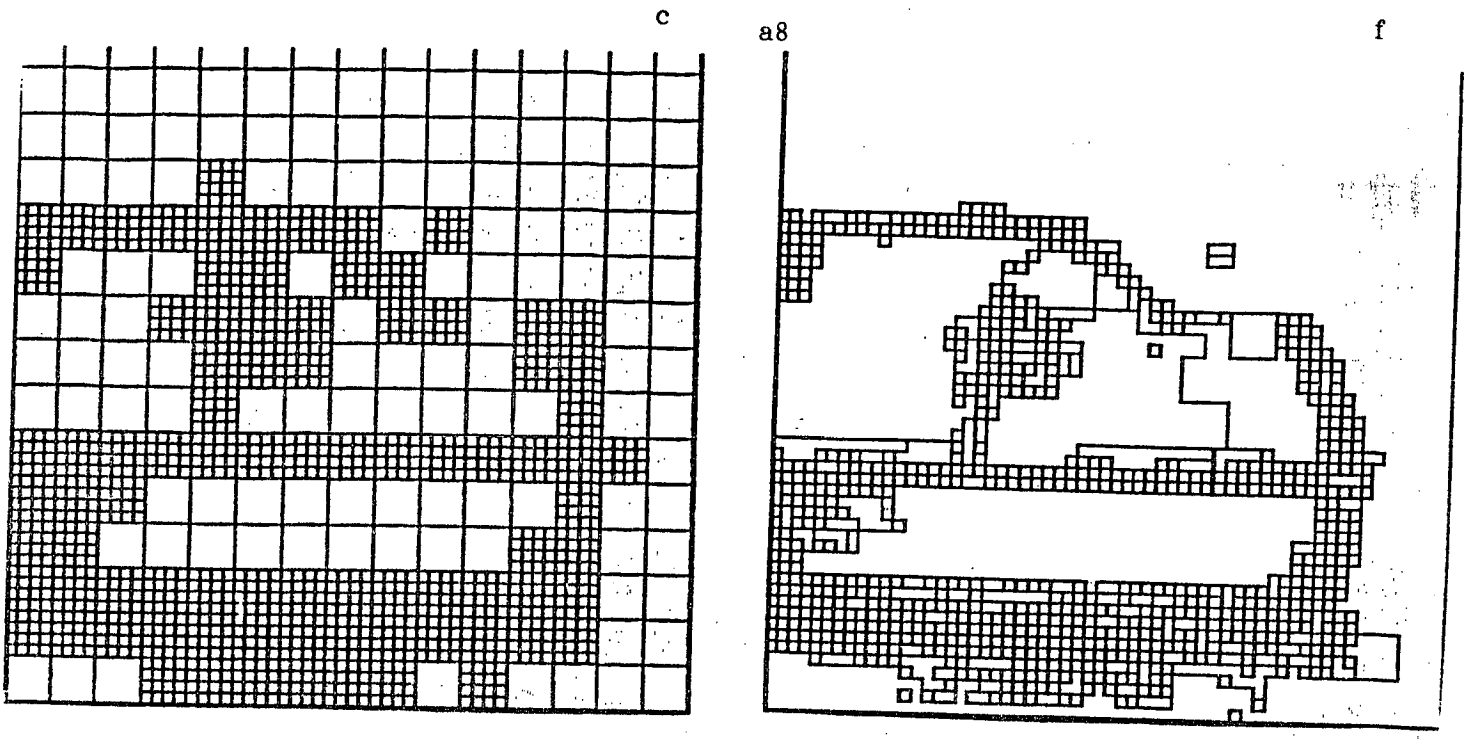


Figure 8 : (continued)

REGION	a_0	b_0
1	0.0272	-0.0188
2	-0.1421	-2.2561
3	-2.2814	-0.7074
4	2.092	-1.215
5	1.936	1.280
6	-0.8120	2.190
7	2.792	2.7921

Table 1 : Representative parameters $\hat{a}_0$, $\hat{b}_0$ corresponding to constant motion model for significant regions of figure 8e.

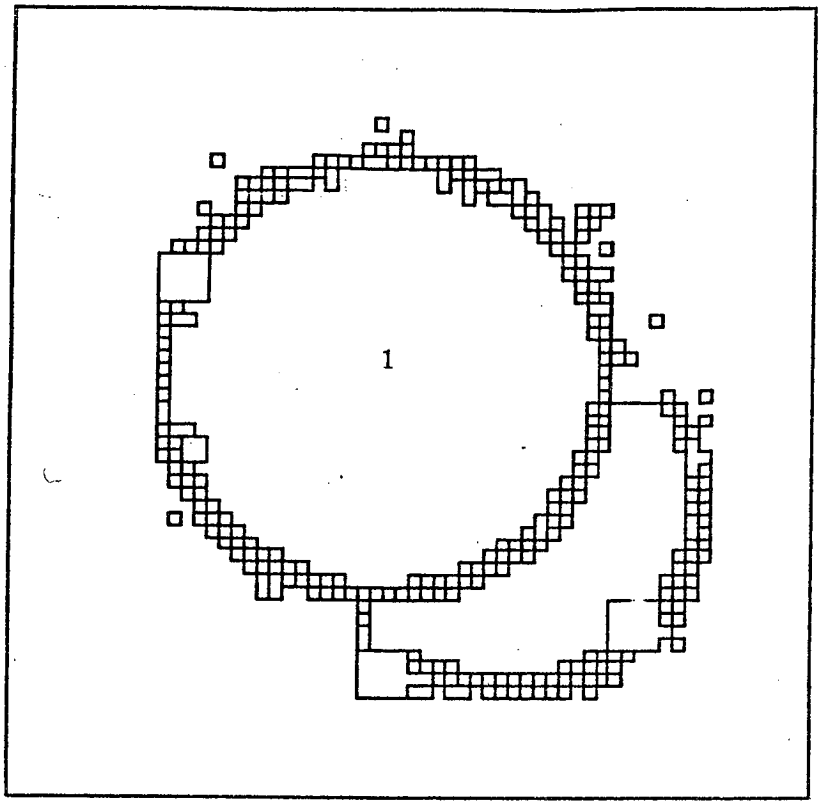


Figure 9 : Final segmentation after the second merge stage, based on the linear motion model, and starting from partition of figure 8 e ($\lambda = 6$)

REGION	a_g	b_g	α	β	γ	δ
1	0.0823	0.0426	-0.04	-0.0002	-0.0001	-0.040

Table 2 : Optimal parameters corresponding to the linear motion model for the region 1 of figure 9.

Imprimé en France
par
l'Institut National de Recherche en Informatique et en Automatique