



NICE du 20 au 24 MAI 1985

ESTIMATION RECURSIVE DE LA MOYENNE DE SIGNAUX BIDIMENSIONNELS :  
VERS UNE APPROCHE PARALLELE DE LA SEGMENTATION D'IMAGES

Jean-Pierre GAMBOTTO

ETCA - 94114 ARCUEIL CEDEX

**RESUME**

Dans cet article, nous présentons un algorithme de regroupement hiérarchique qui traite toutes les régions d'une image en parallèle. Cet algorithme itératif permet de segmenter l'image en utilisant à la fois l'information statistique sur les régions homogènes et l'information du gradient sur le contour des régions détectées.

A l'initialisation, l'algorithme analyse directement l'image originale en considérant que chaque pixel est une région ; une itération consiste à regrouper en parallèle des régions adjacentes ayant des moyennes suffisamment proches. Ces regroupements sont effectués sous contrainte, et à une itération donnée un nombre limité de régions sont regroupées. La contrainte est relâchée de façon continue aux itérations suivantes afin d'autoriser d'autres regroupements. Un critère permettant de comparer les gradients moyens sur 2 régions adjacentes et sur leur contour commun constitue un test d'arrêt au processus de regroupement.

Après avoir analysé certaines propriétés importantes de cet algorithme, nous présentons des résultats expérimentaux avec des images synthétiques, ainsi que la segmentation d'une image naturelle.

**SUMMARY**

This paper describes an algorithm for region growing which process all the regions in the image in a parallel and recursive fashion. This segmentation algorithm makes use of the statistical properties of homogeneous regions and of the gradient information over the region boundaries. It starts processing the raw image by assuming that each pixel is a region. All pairs of adjacent regions whose mean grey levels are close enough are merged. However, merging is performed using a constraint and only a limited number of regions are merged in a given iteration. This constraint is iteratively modified in order to allow other regions to merge in the subsequent iterations. A criterion which is based on the analysis of the mean gradients over two adjacent regions and over their common boundary is used to terminate the growth.

The algorithm is described and some important properties are discussed. Next we present the result of applying this algorithm to synthetic and natural images.



## I - INTRODUCTION

Les méthodes classiques de modélisation monodimensionnelle ou vectorielle supposent que le signal est quasi stationnaire, et consistent en général à analyser les propriétés spectrales sur des zones juxtaposées du signal. Récemment, de nouvelles méthodes ont été proposées pour modéliser les non-stationnarités ; elles détectent les changements brusques [1], ou bien modélisent les évolutions lentes [2] des propriétés spectrales du signal.

Dans le cas des images naturelles, une partie importante de l'information est véhiculée en premier lieu par les non-stationnarités de moyenne du signal bidimensionnel ; aussi les techniques basées sur la recherche de transitions entre 2 niveaux moyens ont été très étudiées dans le domaine du traitement d'images, et sont principalement utilisées pour détecter le contour des objets. Mais la plupart de ces techniques (p.ex. méthode du gradient, ou méthode utilisant les passages par zéro de la dérivée seconde), présentent l'inconvénient de n'utiliser que l'information locale du signal bidimensionnel, leur application directe est souvent médiocre (présence d'éléments de contour parasites ou non détectés...) et des algorithmes complexes sont nécessaires pour arriver à une segmentation correcte de l'image.

Les techniques de modélisation du signal bidimensionnel sont assez bien connues (estimation linéaire, modèles de textures...), et permettent de représenter et même de synthétiser de façon assez satisfaisante des zones d'images ayant des propriétés statistiques [3-4] ou structurelles [5] bien définies. Le problème de la segmentation automatique d'une image en zones homogènes est plus complexe ; en particulier il semble que le choix de la fenêtre d'analyse soit aussi important que celui de la méthode de modélisation. Dans le cas monodimensionnel, la segmentation automatique du signal est souvent basée sur une comparaison entre 2 modèles : un modèle défini sur une fenêtre de référence (passé) et un modèle défini sur une fenêtre d'analyse (futur). Cependant, le choix de paramètres tels que la taille, la position relative, ou l'évolution des 2 fenêtres au cours du temps est souvent empirique et varie suivant les auteurs [6]. Une transposition simple au cas bidimensionnel consiste par exemple à utiliser des fenêtres rectangulaires de dimensions prédéfinies [3] ; mais il est évident que ce choix n'est pas optimal, et le problème se complique si l'on considère, que dans une image, une fenêtre peut être une région connexe de forme quelconque.

En pratique, la fenêtre doit être déterminée de façon à ce que l'échantillon du signal soit stationnaire et de dimension assez importante afin de pouvoir estimer un modèle statistique. Ces contraintes sont souvent contradictoires, et ont conduit certains auteurs [6-7] à utiliser une fenêtre de référence de taille croissante au cours du temps et une fenêtre d'analyse glissante ; le modèle de référence se construit alors itérativement et utilise les données de façon optimale.

Les méthodes de regroupement de régions ("région growing") permettent justement de segmenter une image en analysant des régions de formes et de tailles variées [8]. Dans l'étude de Muerle et Allen [9] la segmentation de l'image est obtenue par accroissement de régions élémentaires de façon séquentielle ; la méthode proposée par Brice et Fennema [10] est plus élaborée et consiste à regrouper des régions adjacentes par application de 2 heuristiques utilisant des informations sur les frontières communes et les tailles des régions. Récemment certains auteurs ont proposé des méthodes plus complexes ; citons: la méthode du "split

and merge" de Pavlidis et Horowitz [11] qui utilise à la fois un processus de regroupement et un processus de division ce qui permet d'éviter un certain nombre d'erreurs de segmentation ; et l'approche développée par Pong [12] qui utilise plusieurs algorithmes : l'images est tout d'abord approximée par un modèle à facettes [13], et la segmentation finale est ensuite obtenue par regroupement des facettes.

Nous proposons dans cet article un algorithme de regroupement qui analyse les régions de l'image en parallèle et de façon itérative. A l'initialisation cet algorithme traite directement l'image originale en considérant que chaque pixel est une région. Une itération consiste alors à regrouper en parallèle des régions adjacentes ayant des moyennes suffisamment proches ; ces regroupements sont effectués sous contrainte et à chaque itération un nombre limité de régions sont regroupées ; cette contrainte est relâchée de façon progressive aux itérations suivantes afin que d'autres regroupements aient lieu.

## II - ESTIMATIONS DE LA MOYENNE PAR REGROUPEMENT HIERARCHIQUES

Considérons un algorithme simple qui permet de partitionner une image en zones homogènes, et s'inspire des techniques classiques de regroupements hiérarchiques en reconnaissance des formes [15]. L'algorithme est itératif, et à chaque itération, les deux régions adjacentes les plus semblables sont recherchées dans l'image  $N \times N$ , et ensuite regroupées. A l'initialisation, les régions considérées sont les  $N^2$  pixels de l'image originale ; le processus de regroupement peut donc être au plus itéré ( $N^2 - 1$ ) fois, et dans ce cas extrême on obtient une seule région comprenant tous les pixels de l'image. Une propriété intéressante de cet algorithme provient du fait qu'une itération réalise une transformation minimale de l'image. Un tel processus risque évidemment d'être coûteux en temps calcul, même pour le traitement d'une image de taille moyenne. Nous proposons donc une stratégie qui est sous-optimale mais présente certaines propriétés intéressantes et en particulier permet l'utilisation d'une structure parallèle.

Remarquons tout d'abord que dans une image numérique codée sur 8 bits, la probabilité pour que 2 points voisins aient le même niveau est égale à  $1/2^8$  ; en utilisant la 8-connexité on aura en moyenne (cas d'un bruit blanc) environ  $4N^2/2^8$  paires de points voisins de même niveau dans l'image (soit 256 paires si  $N$  est égal à 128). Dans les images réelles, les pixels voisins étant fortement corrélés on obtient des valeurs nettement plus élevées. Plus généralement, le nombre de paires de petites régions adjacentes qui partagent la même propriété est souvent assez grand. Il semble donc intéressant de traiter en parallèle les paires des régions qui ont la même propriété.

### Le critère de regroupement

Le critère utilisé à chaque itération permet le regroupement d'un nombre limité de régions adjacentes.

Définissons une suite ordonnée de seuils  $S_0, S_1, \dots, S_k, \dots$  par :

$$S_0 = 0 ; S_k = S_{k-1} + \Delta \quad \text{si } k > 0 \quad (1)$$

où  $\Delta$  est une constante positive ; deux régions adjacentes  $R(i)$  et  $R(j)$  sont regroupées si leurs moyennes  $L(i)$  et  $L(j)$  vérifient le critère  $C(k)$  :

$$|L(i) - L(j)| \leq S_k \quad (2)$$



ESTIMATION RECURSIVE DE LA MOYENNE DE SIGNAUX BIDIMENSIONNELS :  
VERS UNE APPROCHE PARALLELE DE LA SEGMENTATION D'IMAGES

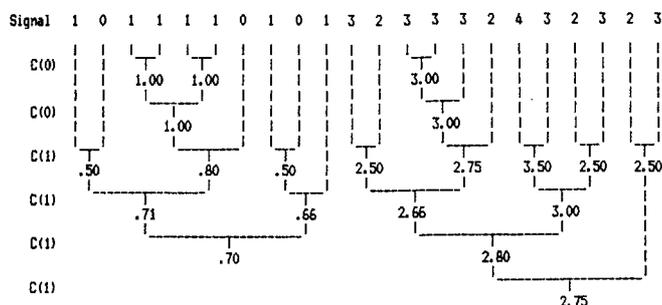


Figure 1 : Application de l'algorithme de regroupement à un signal monodimensionnel ; le paramètre  $C(k)$  est indiqué pour chaque itération

Les critères  $C(k)$  sont appliqués de façon séquentielle au cours de l'algorithme en commençant par  $C(0)$  à la première itération. A l'itération  $n$  toutes les paires de régions adjacentes sont testées en utilisant  $C(k)$  : si aucune paire ne vérifie le critère, à l'itération suivante le critère utilisé sera  $C(k+1)$  ; par contre si au moins une paire est regroupée, le critère  $C(k)$  sera à nouveau utilisé à l'itération  $n+1$ . Donc un même critère est en général appliqué pendant plusieurs itérations, tant que des regroupements inter-régions se produisent. Remarquons que, lors des premières itérations, seules les régions adjacentes ayant le même niveau sont regroupées et donc l'image n'est pas transformée ; mais au fur et à mesure que le paramètre  $k$  croît, des régions ayant des moyennes le plus en plus éloignées sont regroupées.

Ce processus de regroupement peu bien sûr aussi être utilisé pour la segmentation d'un signal monodimensionnel ; dans ce cas chaque élément a deux voisins et l'algorithme est très simple. La Figure 1 présente un exemple où la valeur du paramètre  $\Delta$  est égale à 1 ; lors de chaque regroupement la valeur moyenne sur la nouvelle fenêtre est calculée de façon itérative à partir des valeurs sur les fenêtres précédentes ; le critère  $C(0)$  est appliqué 2 fois, le critère  $C(1)$  4 fois, et après 6 itérations le signal est segmenté en 2 zones homogènes.

L'algorithme réalise donc une estimation de la moyenne en utilisant des fenêtres d'analyse de tailles croissantes, et chaque itération consiste à comparer l'information sur des fenêtres adjacentes. L'approche hiérarchique et parallèle permet de passer de l'information locale en tout point du signal à l'information globale de manière continue.

### III - UN ALGORITHME DE REGROUPEMENT DE REGIONS

#### Représentation des régions

Supposons qu'à l'itération  $n$  l'image est décomposée en  $M$  régions  $\{R(1), \dots, R(M)\}$  ; chaque région  $R(i)$  est décrite par son contour qui est une courbe fermée, sa surface, et sa moyenne  $L(i)$ . L'analyse des propriétés entre régions adjacentes conduit à s'intéresser à leur frontière commune : à chaque paire de régions adjacentes on associe l'ensemble  $C(i,j)$  qui contient les paires de pixels voisins  $((m,n), (h,p))$  tels que  $(m,n) \in R(i)$  et  $(h,p) \in R(j)$ .

#### Algorithme

L'algorithme de regroupement utilisé pour la segmentation des images est décrit en détail dans la référence [14] ; il est similaire à celui présenté dans le cas monodimensionnel avec toutefois un certain nombre de différences dues au caractère bidimensionnel des données. Chaque itération se décompose en 3 étapes :

- a) création de nouvelles régions par analyse du voisinage de chaque pixel non encore traité,
- b) test des pixels adjacents à la région  $R(i)$  puis regroupement,
- c) recherche et test de la meilleure région  $R_j^*$  adjacente à la région  $R(i)$ , puis regroupement.

Cette décomposition permet de réduire l'espace mémoire nécessaire ainsi que le temps de calculs. La recherche des régions et pixels adjacents à une région donnée  $R(i)$  est réalisée par suivi de contour uniquement aux étapes b) et c), et les regroupements pixel-pixel sont effectués séparément à l'étape a).

Contrairement au cas monodimensionnel, chaque région dans l'image peut avoir un grand nombre de voisins. Aussi les regroupements multiples (un pixel ou une région avec tous les pixels voisins) sont autorisés aux étapes a) et b) ; par contre, une région ne peut être regroupée qu'une seule fois avec une région voisine dans l'itération.

Cet algorithme a un certain nombre de propriétés importantes. Remarquons tout d'abord que pour une valeur de  $k$  fixée, l'étape a) ne doit être exécutée qu'une seule fois, alors que les étapes b) et c) sont en général exécutées plusieurs fois. L'étude de la structure de l'arbre (cf. Figure 1) est aussi intéressante, elle montre que les regroupements sont guidés par l'évolution des contraintes (2) au cours du temps ; une propriété importante de cette structure est maintenant démontrée.

**Propriété :** Soit  $R_g$  l'ensemble des régions regroupées à l'itération  $n$ , et  $V_g$  l'ensemble formé par toutes les régions adjacentes aux régions appartenant à  $R_g$  ; supposons ces deux ensembles non vides, alors à l'itération  $n+1$ , seules les régions appartenant à l'ensemble  $M_g = R_g \cup V_g$  peuvent être modifiées.

La démonstration est évidente ; en effet, considérons une région  $R \in M_g$ . La relation d'adjacence étant symétrique, il est clair que aucune région adjacente à la région  $R$  n'a été modifiée dans l'itération  $n$ . L'ensemble  $R_g$  étant non vide, le même critère  $C(k)$  s'applique aux itérations  $n$  et  $n+1$  ; on en déduit que la région  $R$  ne sera pas modifiée dans l'itération  $n+1$ .

Une conséquence de cette propriété est que les seules régions qui doivent être analysées à l'itération  $n+1$  sont celles qui ont été modifiées à l'itération  $n$ . Ainsi la première fois où le critère  $C(k)$  est utilisé, toutes les régions de l'image sont traitées, puis le nombre de régions à analyser diminue à chaque itération.

#### Arrêt automatique du processus de regroupement

Le processus de regroupement permet de construire un ensemble de solutions optimales en l'absence d'informations a-priori sur la loi statistique du signal, mais ne donne pas une segmentation finale. Des critères d'arrêt très simple peuvent être utilisés : par



exemple le processus de regroupement peut être arrêté lorsqu'un nombre donné de régions est obtenu, ou bien lorsque le paramètre  $k$  devient égal à une valeur fixée a priori  $k_{max}$ . Mais ces 2 heuristiques sont difficiles à utiliser pour la segmentation automatique d'une image naturelle, car le nombre de régions homogènes dans une image est très difficile à estimer, et un seuil approprié sur le critère  $C(k)$  varie souvent de façon importante d'un endroit à un autre de l'image.

Le système visuel détecte les bords des régions aux endroits où le gradient prend une valeur élevée relativement aux valeurs voisines. Cependant, cette propriété ne signifie pas toujours l'existence d'une transition brutale entre deux régions adjacentes. Le test d'arrêt proposé utilise donc à la fois l'information locale sur les contours et l'information globale des régions.

L'estimation du gradient sur le contour commun aux 2 régions  $R(i)$  et  $R(j)$  est donnée par :

$$GC(i,j) = \sum_{(p,q), (r,s) \in C(i,j)} |Y_{pq} - Y_{rs}| / N_{ij}$$

où  $Y_{p,q}$  est le niveau de gris et  $N_{ij}$  est le nombre de pixels de  $C(i,j)$ . Soit  $N_i$  le nombre de paires de points sur la région  $R(i)$ , l'estimateur du gradient moyen sur une région :

$$GR(i) = \sum_{(p,q), (r,s) \in R(i)} |Y_{pq} - Y_{rs}| / N_i$$

est calculé de façon itérative.

Donc la décision de regroupement comprend, en plus du test (2), une comparaison entre le gradient moyen  $GC(i,j)$  sur le contour commun et le gradient moyen sur les deux régions adjacentes  $R(i)$  et  $R(j)$  ; ce second critère est défini par :

$$GC(i,j) < F \{GR(i), GR(j)\} \quad (3)$$

où  $F$  est une fonction positive qui dépend des valeurs des gradients moyens sur les 2 régions. Le seuil de détection d'un contour varie donc en fonction du niveau de bruit moyen ou de la texture dans les régions voisines ; et la fonction  $F$  doit bien sûr augmenter lorsque le gradient augmente sur les 2 régions. On en déduit que  $F$  doit être une fonction croissante des variables  $GR(i)$  et  $GR(j)$ . Soient  $A = \max(GR(i), GR(j))$  et  $B = \min(GR(i), GR(j))$  la fonction utilisée est :

$$F(A,B) = \alpha \frac{A + B + \beta}{A - B + \gamma}$$

où  $\alpha, \beta, \gamma$  sont 3 constantes positives. Les constantes  $\beta$  et  $\gamma$  permettent simplement d'éviter une indétermination ou une valeur nulle à la fonction  $F$  lorsque les 2 régions ont un niveau de gris constant. L'équation  $S(k) = \alpha\beta/\gamma$  fixe le seuil au-dessus duquel 2 régions de niveaux constants ne sont jamais regroupées.

Supposons que le gradient moyen soit le même sur les 2 régions, alors  $A = B$ , et si on néglige le facteur  $\beta$ , la fonction  $F$  est tout simplement un multiple de la moyenne des gradients. D'autre part, en supposant  $(A+B)$  constant, on voit que la fonction  $F$  diminue lorsque  $(A-B)$  augmente ; cette fonction devrait donc permettre de séparer des zones peu bruitées de zones fortement bruitées.

Dans la pratique et afin que les statistiques soient significatives, le critère (3) n'est pas appliqué lorsque les régions adjacentes ont des surfaces trop petites. Le même, il est en général inutile de calculer ce critère lorsque le paramètre  $k$  a une valeur trop faible.

#### IV - RESULTATS

L'algorithme a tout d'abord été appliqué à plusieurs images synthétiques comprenant chacune 3 régions : un fond, un rectangle, et un triangle recouvrant partiellement le rectangle. Dans la première image (Figure 2a) le fond a un niveau constant égal à 128, le rectangle et le triangle ont un niveau qui varie avec une pente constante suivant l'axe vertical (de 130 à 94 avec une pente égale à  $-0.5$  pour le rectangle et de 155 à 183 avec une pente égale à  $0.5$  pour le triangle) ; un résultat intermédiaire est donné sur la Figure 2b et la Figure 2c montre que la segmentation finale est parfaite. Dans une seconde image (Figure 3a) chacune des régions est obtenue par addition d'un bruit blanc gaussien centré de variance égale à 7. à un niveau constant. Dans ce cas l'image a été lissée avec un filtre  $3 \times 3$  avant d'utiliser l'algorithme de regroupement. Les 3 régions sont correctement détectées ; on note cependant, la présence de quelques fausses alarmes qui pourrait être facilement éliminées dans une étape ultérieure du traitement. Ces premiers résultats montrent que l'algorithme ne voit pas les variations continues ou locales du signal mais par contre détecte avec grande précision les transitions brutales et étendues.

L'algorithme a aussi été appliqué à l'image naturelle de la Figure 4a ; les Figures 4b et 4c représentent des résultats intermédiaires après 64 et 95 itérations. La valeur du paramètre  $\Delta$  est ici égale à  $0.3$ , et à la première itération 5367 régions sont créées. La segmentation finale est présentée sur la Figure 4d : il reste 218 régions. La plupart des transitions importantes dans l'image ont été détectées ; remarquons que les ombres et les réflexions, bien que n'étant pas toujours délimitées par des contours très nets, sont aussi détectées.

Remarquons que le paramètre  $\Delta$  contrôle le nombre de régions regroupées à chaque itération. Lorsque  $\Delta$  diminue le nombre de régions regroupées diminue aussi et un plus grand nombre d'itérations sont nécessaires pour obtenir une segmentation finale. Inversement lorsque  $\Delta$  augmente le nombre de régions regroupées augmente de la même façon. D'autre part, on montre facilement que diviser le paramètre  $\Delta$  par 2 est équivalent à multiplier la dynamique de l'image originale par 2. ; ce résultat fournit une méthode pour trouver une valeur optimale au paramètre  $\Delta$ .

#### V - CONCLUSION

Un algorithme de segmentation qui analyse les différentes parties de l'image en parallèle a été développé. La méthode de regroupement sous contrainte, avec relâchement progressif de la contrainte au cours du temps, conduit à une estimation très précise de la moyenne. A chaque itération un certain nombre de régions rivalisent avec leurs voisines, et le parallélisme évite, dans une certaine mesure, qu'une région donnée ne soit privilégiée au dépend des autres.

Les résultats obtenus montrent l'importance d'une estimation correcte de la moyenne dans les images. D'autres informations statistiques telles que la variance, la densité de probabilité, ou bien des paramètres de textures peuvent être estimées en utilisant la même structure d'algorithme.



ESTIMATION RECURSIVE DE LA MOYENNE DE SIGNAUX BIDIMENSIONNELS :  
VERS UNE APPROCHE PARALLELE DE LA SEGMENTATION D'IMAGES

REFERENCES

- [1] Séminaire CNRS "Détection de ruptures dans les modèles dynamiques de signaux et systèmes", Paris, 21-22 Mars 1984
- [2] Séminaire CNRS "Modélisation de signaux non stationnaires", Nice, 18 Juin 1984
- [3] J.P. GAMBOTTO : "Time series for textures", in Digital Image Processing, J.C. SIMON et al. Eds) D. Reidel Pub. Company, Dordrecht, Holland, 1981
- [4] N. AHUJA, B.J. SCHACHTER : Patterns Models, Wiley, 1982
- [5] K.S. FU : "Syntactic pattern recognition and application" Prentice hall, pp 445-480, 1982
- [6] M. BASSEVILLE : "Quelques algorithmes de détection de changements de caractéristiques spectrales utilisés en traitement du signal", réf. [1] pp. 6.1 - 6.35
- [7] U. APPEL, A.V. BRANDT : "Performance comparison of two segmentation algorithms using growing référence windows", 6ème Conférence Internationale sur l'analyse et l'Optimisation des Systèmes, Nice, 19-22 Juin 1984
- [8] S.W. ZUCKER : "Survey. Region growing : childhood and adolescence", Computer Graphics and Image Processing 5, pp. 382-399, 1976
- [9] J.L. MUERLE, D.C. ALLEN : "Experimental evaluation of technique for automatic segmentation of objects in a complex scene", in Pictorial Pattern Recognition (G.C. CHEN et al. Eds), pp. 3-13, Thomson, Washington, 1968
- [10] C. BRICE, C. FENNEMA : "Scene analysis using regions", Artificial Intelligence 1, pp. 205-226, 1970
- [11] S.L. HOROWITZ, T. PAVLIDIS : "Picture segmentation by a direct split and merge procedure", Proc. 2nd International Joint Conf. on Pattern Recognition, pp. 424-433, 1974
- [12] T. PONG et al. : "Experiments in segmentation using a facet model region grower", Computer graphics and image processing 25, pp. 1-23, 1984
- [13] R.M. HARALICK, WATSON : "A facet model for image data", Computer Graphics and Image Processing 15, pp. 113-129, 1981
- [14] J.P. GAMBOTTO, O. MONGA : "A parallel and hierarchical algorithm for region growing", Proc. IEEE Conf. on Computer Vision and Pattern Recognition, San-Francisco, Juin 1985
- [15] R.O. DUDA, P.E. HART : Pattern Classification and Scene Analysis, Wiley, New-York, 1973, pp. 228-237.

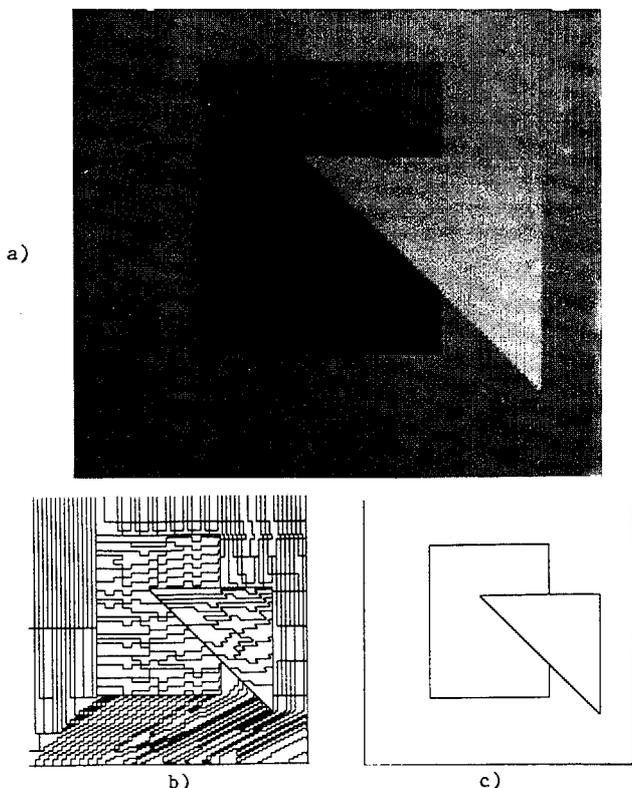


Figure 2 : Image synthétique comprenant : un fond (niveau 128), un rectangle (niveau de 130 à 94 avec une pente égale à  $-.5$ ), un triangle (niveau de 155 à 183 avec une pente égale à  $.5$ )

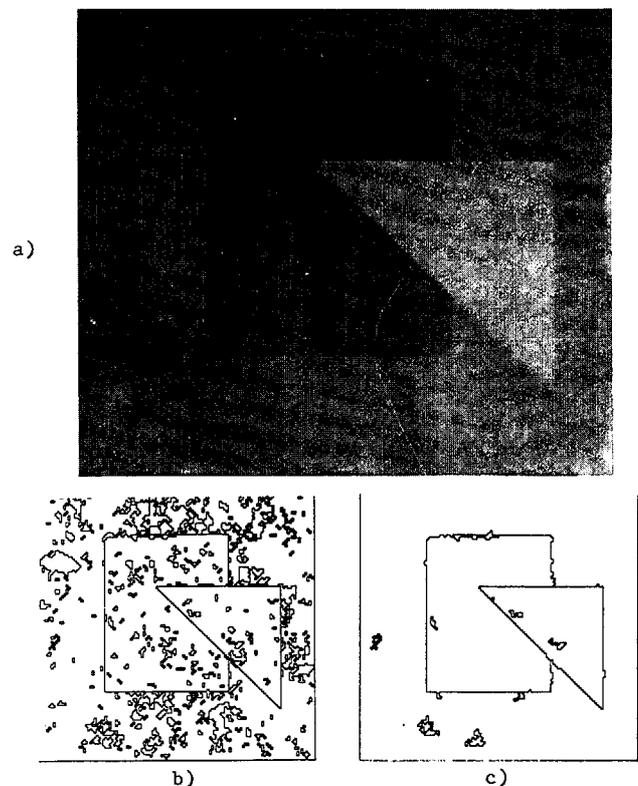


Figure 3 : Image synthétique comprenant 3 régions de niveau constant (130, 150, 170) auquel on a ajouté un bruit blanc gaussien



ESTIMATION RECURSIVE DE LA MOYENNE DE SIGNAUX BIDIMENSIONNELS :  
VERS UNE APPROCHE PARALLELE DE LA SEGMENTATION D'IMAGES

