

Binarisation d'images

Deux solutions à ce problème

Image binarization: two solutions for this problem

**Alain CHEHIKIAN**

Laboratoire de Traitement d'Images et Reconnaissance de Formes, LTIRF/INPG, 46, avenue Félix-Viallet, 38031 GRENOBLE CEDEX.

Professeur à l'Université Joseph Fourier de Grenoble, directeur du laboratoire TIRF de l'ENSERG/INPG. Domaines de recherches: Lecture Optique (Thèse de Doctorat ès Sciences, 1977), Traitement d'Image, Vision par ordinateur, Architectures de machines de Vision.

RÉSUMÉ

Nous nous intéressons à une classe d'images représentatives d'objets sombres disposés sur un fond clair ou inversement. L'image analogique délivrée par un capteur traduit généralement très mal cette propriété: défaut d'éclairage, inhomogénéité de réponse du capteur, nature même de l'image, et un procédé classique de décision par comparaison à un seuil ne peut convenir dans l'immense majorité des cas. Nous proposons deux algorithmes performants pour résoudre ce problème. Le premier procède par une analyse multifréquence pour définir si un pixel appartient ou non à un objet. Le second procède par apprentissage d'un seuil local et classification des pixels sur la base de ce seuil.

MOTS CLÉS

Binarisation d'image, représentation multiéchelle, représentation multifréquence, apprentissage du seuil.

SUMMARY

We are concerned with images representing dark objects located on a clear background or reciprocally. The analog image issued from the scanning hardware does not exhibit such a property because of non uniform illumination, non ideal resolution characteristics of the scanner and also because of the nature of the image itself. Simply thresholding such an image cannot solve the problem. We propose two approaches to solve the image binarization problem. The first one uses a multispectral analysis to define those pixels lying into a dark object. The second one uses a local threshold learning process to classify pixels belonging or not to a dark object.

KEY WORDS

Image binarization, multiscale representation, multispectral representation, threshold learning.

Introduction

Nous nous intéressons dans ce qui suit à une classe d'images représentatives d'objets de faible réflectance disposés sur un fond de réflectance plus élevé (ou inversement). Cette classe d'images se rencontre en lecture optique où les caractères sont «noirs» et le fond habituellement «blanc». On la rencontre également dans des scènes industrielles où grâce à un éclairage artificiel, on cherche à distinguer un objet particulier parmi d'autres en exploitant ses caractéristiques de réflectance. On notera cependant que ces

deux classes d'images diffèrent par le fait que les caractères alphanumériques sont des objets à structure de trait: une des dimensions est très inférieure à la dimension dans une direction perpendiculaire, ce qui n'est généralement pas le cas des objets industriels. En outre, lorsqu'on traite des caractères alphanumériques, l'épaisseur des traits ainsi que leur longueur maximale est connue *a priori*, au moins approximativement; cette information peut être utile, et est utilisée, dans certains algorithmes de binarisation qui ne sont donc pas applicables aux images représentatives de scènes industrielles. Les solutions que nous proposerons par la suite au problème de la binarisation

n'utiliseront pas ce type d'information, leur application sera donc *a priori* possible tant aux images à structure de trait: caractères alphanumériques, dessins, empreintes digitales..., qu'aux images d'objets de forme quelconque, à la condition toutefois de respecter des contraintes minimales concernant le contraste.

1. Le problème de la binarisation

Binariser une image c'est convertir une image (signal) analogique (le fait que le signal soit échantillonné et numérisé ne modifie pas de manière essentielle sa nature vis-à-vis du problème qui nous concerne) non idéale délivrée par un capteur en une image binaire idéale. Le problème serait simple si le niveau de gris associé au fond était uniforme, le niveau de gris associé aux objets l'était également et si ces niveaux de gris étaient suffisamment différents pour que, par comparaison à un seuil supposé connu, on puisse étiqueter « blanc » tous les pixels de niveaux de gris supérieur ou égal à ce seuil et « noir » tous les pixels de niveau de gris inférieur à ce même seuil. Dans la pratique, cette situation idéale ne se rencontre que très rarement. Les niveaux de gris associés au fond et aux objets diffèrent d'un point à l'autre de l'image soit de manière inhérente au support du fond et des objets, soit par défaut d'éclairage, soit par la nature même du capteur dont la réponse n'est pas homogène. A cela, il faut ajouter le bruit introduit par le capteur d'une part, l'électronique associée d'autre part.

Considérons la figure 1 où on a représenté une ligne d'ordonnée y_0 d'une image analogique $I(x, y)$, où I représente le niveau de luminosité ou niveau de gris du pel de coordonnée (x, y) .

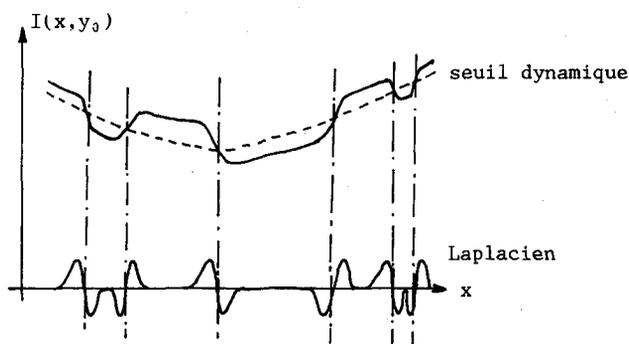


Fig. 1. — Ligne d'image et laplacien.

Cette ligne traverse trois objets sombres par rapport à un environnement clair. Il est bien visible sur cet exemple qui n'a rien d'académique (sauf que par soucis de simplicité on a considéré l'image exempte de bruit), qu'il n'existe pas une valeur de seuil (au sens habituel du terme: seuil constant) permettant de séparer les objets noirs du fond blanc.

On notera que le problème ainsi posé est un problème classique de vision et plus précisément de segmentation d'images en région [1]. Néanmoins les algorithmes utilisés dans ces applications sont de type

itératif et donc difficilement implémentables en temps réel. Or, qu'il s'agisse d'application en lecture optique, ou qu'il s'agisse de vision industrielle, le temps réel s'avère une caractéristique indispensable.

Une façon plus classique de résoudre ce problème consiste à calculer de manière dynamique un seuil local, comme étant une moyenne du niveau de gris sur un environnement particulier autour de chaque pixel. Sur la figure 1, la ligne pointillée représente un seuil dynamique ainsi calculé. On trouvera dans [2] un algorithme basé sur cette méthode. Néanmoins, ces méthodes supposent qu'on définisse la taille de l'environnement sur lequel le calcul de la moyenne sera effectué, c'est-à-dire que l'on dispose de connaissances *a priori* sur la taille des objets noirs. C'est le cas des caractères alphanumériques à la condition toutefois qu'il existe une certaine uniformité dans les tailles, ce ne sera par exemple pas le cas des documents commerciaux ou même de la page de titre de cet article.

White et Rohrer [2] proposent une alternative qui consiste à utiliser le signe d'un laplacien pour marquer les points d'entrée et de sortie d'un objet noir au cours d'un balayage selon x de l'image. Un point d'entrée est reconnu par le passage du laplacien d'une valeur positive à une valeur négative, un point de sortie sera caractérisé par la situation inverse. Sur la figure 1, on a représenté l'évolution du laplacien $L(x, y_0)$. L'algorithme est évidemment très intéressant dans la mesure où il permet une implémentation matérielle en temps réel pour un coût modéré. Cependant, les performances sont altérées en présence de bruit susceptible de perturber le laplacien qui comme toutes les opérations de dérivation est sensible aux bruits de haute fréquence. L'algorithme peut alors oublier une séquence d'entrée dans un objet noir ou au contraire une séquence de sortie d'un objet noir. On peut remédier à ce défaut en prenant en compte l'information *a priori* sur la dimension des objets rencontrés au cours d'un balayage selon x de l'image. On perd alors le caractère de généralité propre à cet algorithme.

Nous proposons dans ce qui suit deux solutions au problème de la binarisation qui n'impliquent pas que l'on fasse d'hypothèse sur la taille des objets. Le premier de ces algorithmes que nous appellerons Binarisation par Analyse Multifréquence à la caractéristique de procéder sans comparaison à un seuil, le second qui dérive du précédent procédera par un Apprentissage Local d'un Seuil.

2. Binarisation par analyse multifréquence

Considérons la photo de la figure 2, elle représente des caractères alphanumériques de bonne qualité. Intéressons-nous au laplacien-gaussien de cette image calculé selon le schéma-bloc de la figure 3.

Dans ce schéma-bloc:

- $I(j, k)$ est l'image délivrée par un capteur (ici une caméra vidéo), représentée par un tableau $n \times p$ (64×256) de pixels codés sur 256 niveaux de gris.
- $I_1(j, k)$ est l'image obtenue par convolution de

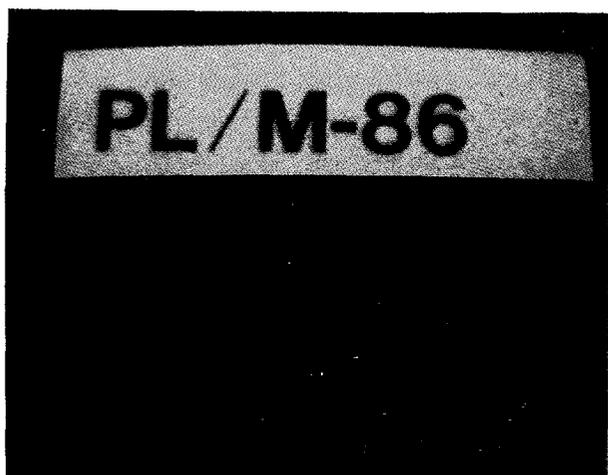


Fig. 2. — Image de caractères alphanumériques

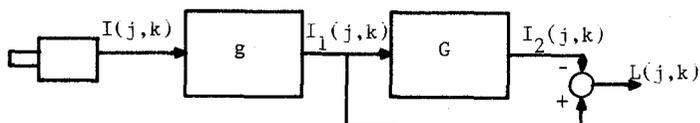


Fig. 3. — Calcul du laplacien.

$I(j, k)$ avec un filtre de réponse impulsionnelle gaussienne:

$$I_1(j, k) = I * g = I(j, k) * \begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{pmatrix} \times 1/16$$



Fig. 4. — Signe du laplacien de l'image de la figure 2.

où le masque de convolution est une approximation de la fonction:

$$g(j, k) = \frac{1}{2\pi \cdot \sigma_1^2} e^{-(j^2 + k^2 / 2\sigma_1^2)}$$

avec $\sigma_1^2 = 0,5$.

— $I_2(j, k)$ est l'image obtenue par convolution de $I_1(j, k)$ avec le masque $G(j, k)$ de taille 7×7 :

$$G(j, k) = g * g * g$$

est une approximation de la fonction

$$G(j, k) = \frac{1}{2\pi \cdot \sigma_2^2} e^{-(j^2 + k^2 / 2\sigma_2^2)}$$

avec $\sigma_2^2 = 3$, $\sigma_1^2 = 1,5$.

— $L(j, k)$ est le laplacien-gaussien de $I_1(j, k)$, calculé sur un environnement 7×7 :

$$L(j, k) = I_1(j, k) - I_2(j, k).$$

On notera que si f_1 est la fréquence maximale contenue dans l'image I_1 :

$$f_1 = \sqrt{v_x^2 + v_y^2} = \frac{c}{\sqrt{2\pi^2 \cdot \sigma_1^2}}$$

où C est une constante qui dépend de la façon dont on estime la fréquence maximale d'un signal ($C=1$ si f_1 est la fréquence de coupure du filtre passe-bas gaussien définie par une atténuation de 1 Np), la fréquence maximale contenue dans l'image $I_2(j, k)$ sera f_2 :

$$f_2 = \frac{c}{\sqrt{2\pi^2(\sigma_1^2 + \sigma_2^2)}} = \frac{f_1}{2}$$

Ainsi $L(j, k)$ représente la réponse à l'image $I_1(j, k)$ d'un filtre d'octave passe-bande. La figure 4 représente le signe du laplacien-gaussien de l'image de la figure 2, codé « + » si le laplacien est positif, « - » si le laplacien est négatif, « ○ » sinon.

On peut constater sur cette figure que les pixels situés à l'intérieur des caractères et au voisinage du contour sont caractérisés par un laplacien négatif, ceux situés à l'extérieur des caractères et au voisinage du contour sont caractérisés par un laplacien positif, la notion de voisinage est, ici, relative à l'étendue du domaine sur lequel est calculé le laplacien, dans notre cas un environnement 7×7 centré sur chaque pixel. Inversement, un pixel caractérisé par un laplacien nul peut appartenir aussi bien au fond qu'à un caractère. On lèvera facilement cette ambiguïté en étendant le domaine sur lequel on calcule le laplacien, ou, ce qui revient au même dans notre cas, en utilisant un filtre passe-bande de largeur supérieure à une octave. Cependant, en procédant ainsi on réduit la résolution sur les contours de l'image et, de plus, il devient nécessaire d'émettre une hypothèse sur la taille minimale des objets (largeur des traits). Une représentation multifréquence de l'image va nous permettre de résoudre ce problème.

2.1. LA REPRÉSENTATION MULTIFRÉQUENCE

Cette technique de représentation d'une image par une pyramide de laplacien est employée en vision [4, 5]. Nous en rappelons brièvement le principe. Comme nous l'avons vu précédemment, une séquence de convolution par un filtre gaussien permet de réaliser de manière contrôlée une réduction de la fréquence maximale contenue dans une image. Lorsque cette fréquence maximale a été réduite dans un rapport 2, on peut procéder à une décimation (sous-échantillonnage) tout en restant en accord avec le théorème de l'échantillonnage. Une séquence de filtrage-décimation permet d'obtenir à partir d'une image ($n \times p$), une suite d'images :

$$(n/2 \times p/2), (n/4 \times p/4), (n/8 \times p/8), \dots$$

et il est clair que cette suite est limitée. L'ensemble des images ainsi obtenues constitue une «représentation

multiéchelle» ou «pyramide multirésolution» permettant d'avoir d'une même scène un ensemble de «vues», depuis la plus globale (image du sommet de la pyramide) jusqu'à la plus détaillée (image de la base de la pyramide). En procédant à une remise à l'échelle $n \times p$ de chacune de ces images par une expansion-interpolation, on transforme la pyramide multirésolution en une représentation multifréquence. L'organigramme de la figure 5 explicite le mode de calcul de cette représentation: l'ensemble des laplaciens l_i constitue la pyramide multirésolution, l'ensemble des laplaciens L_i constitue la représentation multifréquence: réponse d'un ensemble de filtres d'octaves adjacents passe-bandes.

La photo de la figure 6 représente de haut en bas le signe des laplaciens L_1, L_2, L_3, L_4 relatifs à l'image de la photo, fig 2, codé selon :

- blanc si $L_i > 0$;
- gris si $L_i = 0$;
- noir si $L_i < 0$.

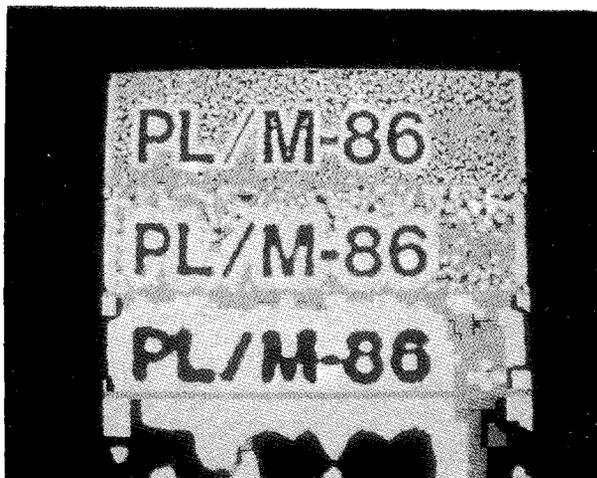


Fig. 6. — Signe des laplaciens multifréquence.

On constate que le niveau de gris du fond étant affecté de variations pourtant peu perceptibles à l'œil, les laplaciens traduisent cette situation par une varia-

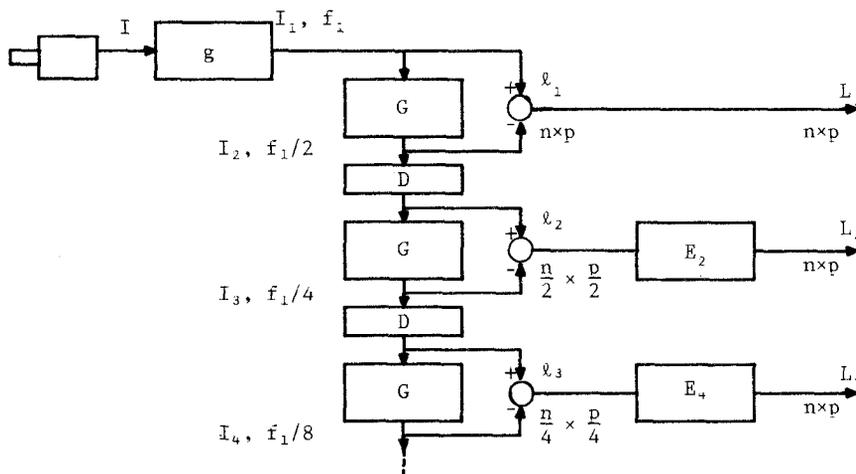


Fig. 5. — Organigramme de calcul de laplacien multifréquences.

tion erratique de leur signe. On limitera cet effet, en codant le signe par :

blanc si $L_1 > +s$;

noir si $L_1 < -s$;

gris sinon.

Expérimentalement, une valeur de $s=2$ s'est avérée convenable. La photo de la figure 7 explicite le codage obtenu.

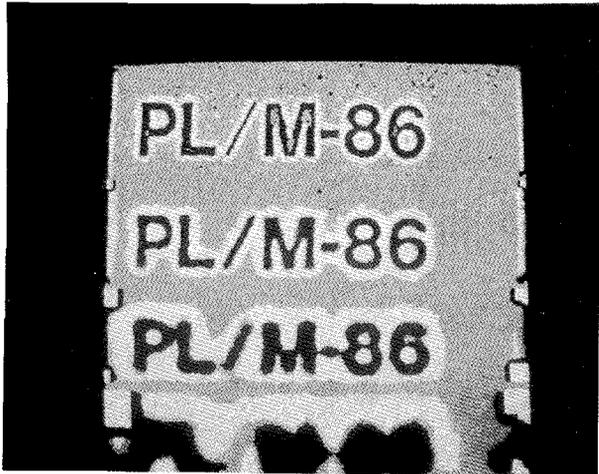


Fig. 7. — Signe des laplaciens multifréquence.

On notera que le laplacien haute fréquence L_1 repère par ses changements de signe les pixels frontières des objets, les points intérieurs trop éloignés des contours ne sont pas repérés (cf. P, M, 6). Le laplacien L_2 repère les pixels intérieurs, L_4 repère un caractère ou un groupe de caractère (cf. 86), les laplaciens de plus basse fréquence repèreront de même un mot, une ligne de texte, une zone de texte dans une page, etc.

2.2 BINARISATION

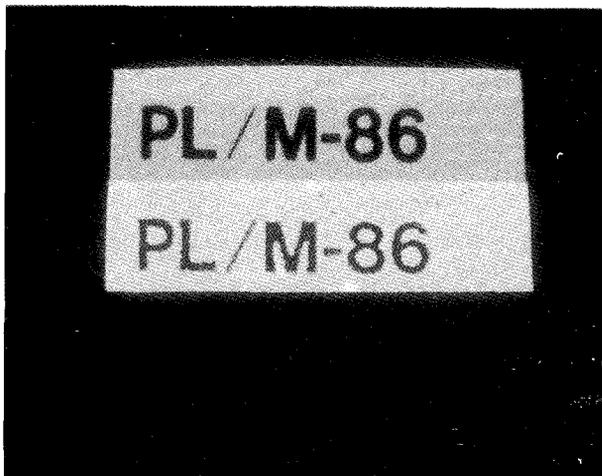
A partir de la représentation multifréquence, chaque pixel (j, k) sera étiqueté « blanc » ou « noir » selon la règle :

```

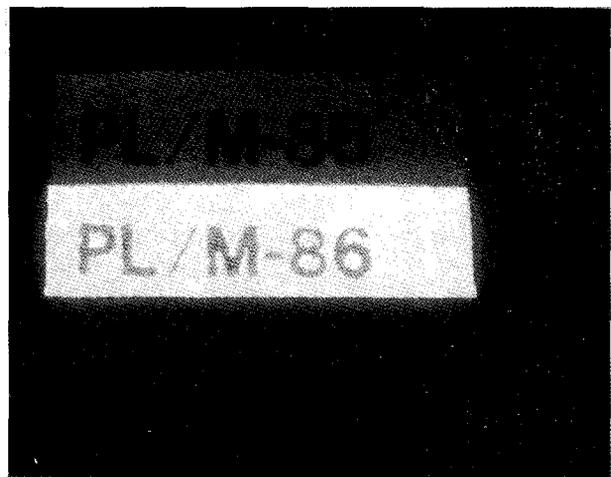
Si  $L_1(j, k) < -s$  alors noir
sinon si  $L_1(j, k) > +s$  alors blanc
  sinon si  $L_2(j, k) < -s$  alors noir
    sinon si  $L_2(j, k) > +s$  alors blanc
      sinon si.....
        .....
          sinon si  $L_4(j, k) < -s$  alors noir
            sinon blanc
    
```

La fin de la règle de décision privilégie le label « blanc » en cas d'ambiguïté.

Cette règle de décision, examinant en premier le signe du laplacien de haute fréquence assure que les contours des objets seront préservés, les laplaciens de



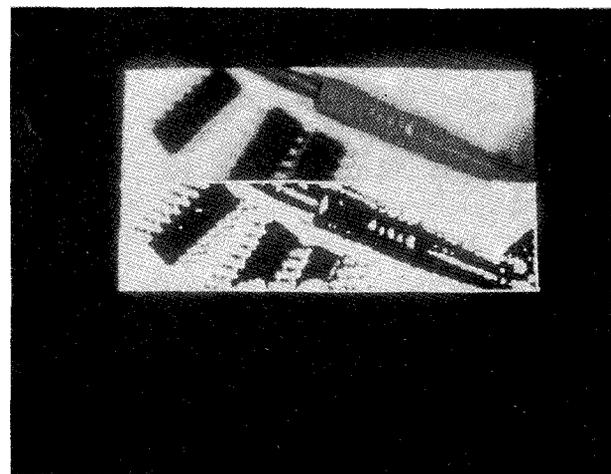
(a)



(b)



(c)



(d)

Fig. 8. — Exemples de binarisation.

plus basse fréquence sont utilisés pour «remplir» l'intérieur des objets.

Les photos de la figure 8 représentent les résultats obtenus dans divers cas de figure à partir d'une représentation multifréquence limitée aux quatre premiers octaves. En (a) l'image est de bonne qualité (fond et objets uniformes), en (b), l'image a les mêmes caractéristiques mais le contraste est sensiblement réduit, en (c), l'image est de qualité médiocre (stencil tiré sur un papier de mauvaise qualité), en (d), l'image représente des objets manufacturés n'ayant pas une structure de trait, les parties sombres n'ont pas un niveau de gris uniforme (pattes des circuits intégrés plus clairs que les circuits intégrés, corps de la fiche banane plus clair que les circuits intégrés) les parties claires ne sont également pas uniformes (motifs de la fiche banane plus sombres que le fond).

2.3. DISCUSSION

L'algorithme de binarisation multifréquence que nous avons décrit présente évidemment avantages et inconvénients.

L'avantage essentiel réside dans le fait qu'à aucun moment un seuil quelconque n'est utilisé (il faut cependant décider, en présence de bruit dû en particulier à la numérisation, du signe d'un laplacien), l'étiquetage des pixels fait appel à la notion d'intérieur des objets, elle-même définie à partir d'un contraste local. Grâce à la représentation multifréquence, la connaissance *a priori* des dimensions maximales des objets n'est pas nécessaire.

Quant aux inconvénients ils sont de deux ordres. En ce qui concerne l'implémentation de l'algorithme, bien que le choix du laplacien-gaussien assure la faisabilité en électronique câblée des diverses opérations, il est peu réaliste, bien que théoriquement possible, de réaliser ces opérations à la volée sur le signal vidéo issu du capteur. Ce type d'algorithme s'implémente bien sur un processeur câblé associant une mémoire d'image et un opérateur de décimation-convolution-soustraction-expansion. Une telle structure suppose que la taille maximale de l'image soit prédéfinie. Une maquette réalisée au laboratoire a permis d'évaluer le temps de traitement d'une image 256×256 pixels pour laquelle 5 laplaciens-gaussiens sont calculés: ce temps de traitement est d'environ 22 ms. Ce format d'image et le temps de traitement associé peuvent convenir lorsqu'il s'agit de binariser l'image d'un objet industriel mais ils ne conviendront pas lorsqu'on voudra traiter des images de textes issus d'un scanner. Un autre inconvénient lié à cet algorithme résulte de sa capacité à binariser des images à très faibles contrastes. A la limite, il binarise de la même façon des objets réels et des objets fictifs dus à des variations du niveau de gris du fond. Il est toujours possible de contrôler la binarisation en contrôlant le codage des laplaciens négatifs au moyen d'un seuil, il est néanmoins très difficile de lier une différence de niveau de gris et la valeur d'une dérivée seconde faisant intervenir à la fois une différence de niveau de gris et une étendue sur laquelle cette différence est évaluée.

Ce sont ces raisons qui nous ont amené à proposer l'algorithme suivant qui procède par Apprentissage Local d'un Seuil.

3. Binarisation par apprentissage local d'un seuil

Dans l'algorithme précédent, nous avons utilisé le fait qu'un laplacien de haute fréquence nous permet d'étiqueter les pixels intérieurs et extérieurs proches du contour des objets, les laplaciens de plus basse fréquence sont utilisés pour remplir l'intérieur des objets par couches successives, ils nous servent donc à établir un lien entre pixels intérieurs situés au voisinage du contour et pixels intérieurs éloignés du contour. Une autre façon d'établir ce lien consiste à dire que tous les pixels intérieurs aux objets sont caractérisés par un niveau de gris différent du niveau de gris des pixels extérieurs situés au voisinage (c'est ce que traduit le laplacien). Il est donc possible d'étiqueter les pixels par comparaison à un seuil à la condition de connaître les niveaux de gris caractéristiques de l'intérieur et de l'extérieur des objets. Le laplacien de haute fréquence va nous permettre de connaître ces niveaux de gris. Considérons la figure 9 où on a représenté une ligne d'image analogique $I(x, y_0)$ et signe $[L_1(x, y_0)]$.

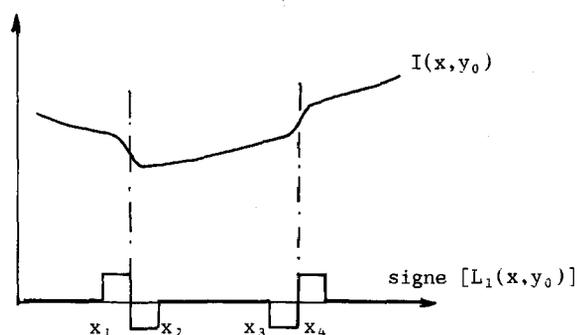


Fig. 9. — Ligne d'image et signe du laplacien.

Soit $B(x, y_0)$ et $N(x, y_0)$ respectivement les niveaux du blanc et du noir dans cette ligne.

Ils seront déterminés par les règles:

$$B(x, y_0) = \begin{cases} I(x, y_0) & \text{si } L_i > 0 \\ B(x-dx, y_0) & \text{sinon:} \\ & \text{conserve la valeur précédente} \end{cases}$$

$$N(x, y_0) = \begin{cases} I(x, y_0) & \text{si } L_i < 0 \\ N(x-dx, y_0) & \text{sinon} \end{cases}$$

$B(x, y_0)$ sera donc réactualisé au cours d'un balayage selon les x croissants chaque fois qu'on approche d'un caractère: on connaîtra donc le niveau local du blanc.

$N(x, y_0)$ sera réactualisé chaque fois qu'on entre ou qu'on sort du caractère: on connaîtra donc le niveau local du noir.

En l'absence de connaissance sur la probabilité *a priori* des pixels blancs ou noirs, on choisira un seuil $S(x, y_0)$:

$$S(x, y_0) = [B(x, y_0) + N(x, y_0)]/2$$

qui sera donc actualisé à chaque changement de signe du laplacien. La seule contrainte à respecter est que

APPLICATIONS

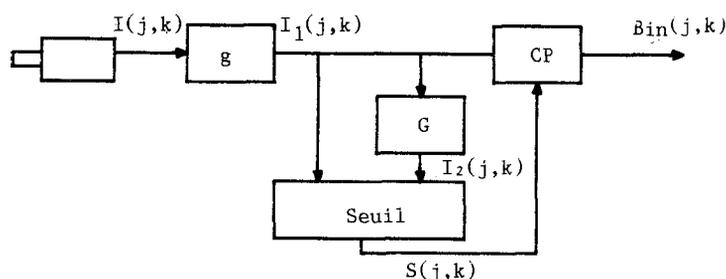


Fig. 10. — Organigramme de l'algorithme de binarisation par apprentissage local d'un seuil.

dans un intervalle $[x_2, x_3]$ par exemple, où $L_1(x)=0$ (cf. fig. 9), le niveau de gris ne varie pas de manière significative par rapport à $S(x)$. Dans notre exemple, il faudra que :

$$I(x) < [I(x_1) + I(x_2)]/2$$

soit

$$\delta I = I(x) - I(x_2) < [I(x_1) - I(x_2)]/2$$

Contrainte d'autant plus facile à respecter que le contraste sur le bord d'entrée de l'objet est important. La première partie de l'algorithme se déroule comme dans l'algorithme précédent et permet d'obtenir l'image filtrée passe-bas $I_2(j, k)$. Un opérateur d'apprentissage du seuil et un comparateur permettent d'obtenir directement l'image binarisée. La figure 10 explicite l'organigramme des opérations.

Dans cet organigramme :

— CP est un comparateur qui délivre l'image binarisée $\text{Bin}(j, k)$ telle que

$$\text{Bin}(j, k) = \begin{cases} 1 & \text{si } I_1(j, k) < S(j, k) \\ 0 & \text{sinon} \end{cases}$$

— l'opérateur d'apprentissage de seuil obéit aux règles :

$$(1) \quad \begin{cases} B(j, 0) = I_2(j, 0) \\ N(j, 0) = 0 \end{cases}$$

$$(2) \quad \text{si } I_1(j, k) > I_2(j, k) \quad \text{alors } B(j, k) = I_2(j, k)$$

et si

$$N(j, k) > I_2(j, k) \quad \text{alors } N(j, k) = I_2(j, k)$$

$$(3) \quad \text{si } I_1(j, k) < I_2(j, k) \quad \text{alors } N(j, k) = I_2(j, k)$$

et si

$$B(j, k) < I_2(j, k) \quad \text{alors } B(j, k) = I_2(j, k)$$

$$(4) \quad S(j, k) = [B(j, k) + N(j, k)]/2.$$

La règle (1) assure l'initialisation des niveaux de blanc et de noir en début de ligne. Telle qu'elle est écrite, elle résulte de l'hypothèse qu'en début de ligne l'image représente le fond.

Les règles (2) et (3) assurent l'actualisation des niveaux du blanc et du noir respectivement. On remarquera que les niveaux réactualisés ne correspondent pas à l'image I_1 mais à l'image filtrée I_2 , on

réduit ainsi fortement les variations erratiques de ces niveaux dues au bruit de fond. La deuxième partie de ces règles assure qu'à la suite d'une réactualisation, on ne peut se trouver dans la situation où un niveau de noir serait supérieur à un niveau de blanc ou inversement, situations qui pourraient se produire en présence de variations lentes des niveaux de gris.

La photo de la figure 11 visualise de haut en bas : l'image de bonne qualité I_1 , l'image binaire obtenue BIN, le signe du laplacien $L_1 = I_1 - I_2$, l'image du seuil $S(j, k)$.

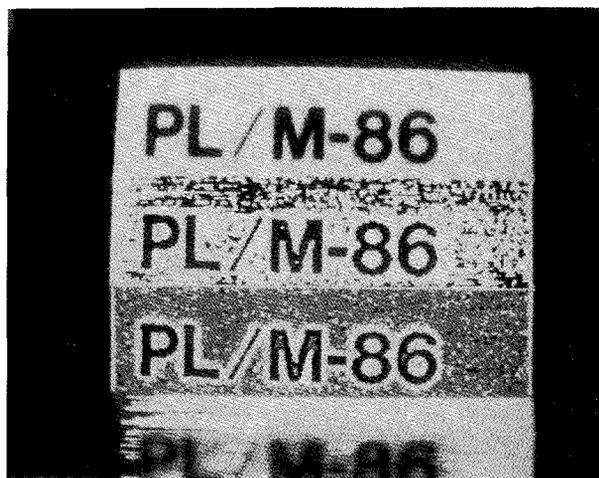


Fig. 11.

On constate sur cette photo que l'algorithme fonctionne bien dans les zones où les objets sont présents : le seuil est convenablement réactualisé, à la condition que à l'intérieur des objets le laplacien ne devienne jamais positif (ce qui n'est pas le cas, par deux fois à l'intérieur du M). Par contre, dans les zones où uniquement le fond est présent, le signe du laplacien utilisé comme indicateur d'intérieur et d'extérieur d'objet s'avère déficient (comment pourrait-il en être autrement).

Ceci nous a amené à modifier les règles (2) et (3) de réactualisation des niveaux du blanc et du noir selon :

$$(2') \quad \text{si } [I_1(j, k) > I_2(j, k)] \quad \text{et } [I_2(j, k) > S(j, k)]$$

alors

$$B(j, k) = I_2(j, k)$$

Cette règle assure la réactualisation du niveau du blanc si le laplacien est non négatif et à la condition

que le pixel soit situé à l'extérieur des objets. Ainsi, le niveau du blanc est contraint à suivre le niveau du fond.

(3') si

$$[I_1(j, k) < I_2(j, k)] \text{ et } [I_2(j, k) < B(j, k) - CT]$$

alors

$$N(j, k) = I_2(j, k).$$

Cette règle assure la réactualisation du niveau du noir si le laplacien est négatif : le pixel est présumé situé à l'intérieur d'un objet. On distingue un objet véritable d'un bruit au moyen d'un contraste local minimal CT. Dans ce qui suit le contraste local minimal est 1/32 de la pleine échelle.

La photo de la figure 12 visualise le résultat obtenu en utilisant ces nouvelles règles. On notera sur l'image du seuil que celui-ci reste bas dans les zones où aucun objet n'est présent. On notera également qu'il varie en fonction du contraste local dans les zones où les objets sont présents.

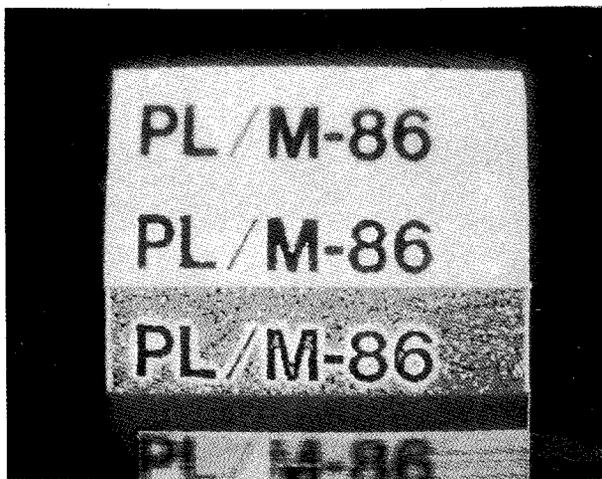


Fig. 12. - Règles de réactualisation du seuil modifiées.

Les photos des figures 13 et 14 visualisent les résultats sur des images faiblement contrastées. Les images du seuil sont peu visibles, celui-ci restant toujours faible.

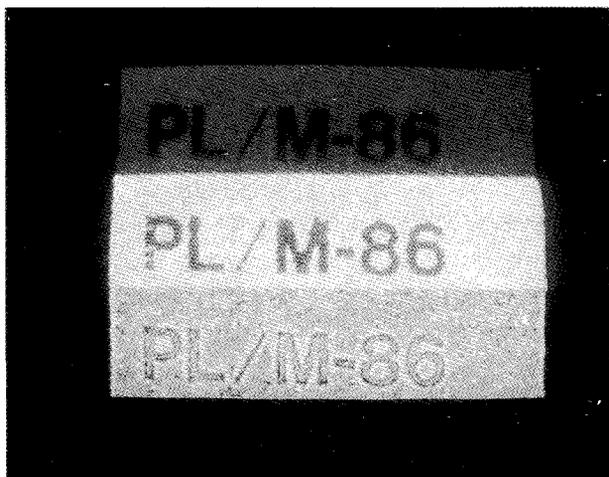


Fig. 13.

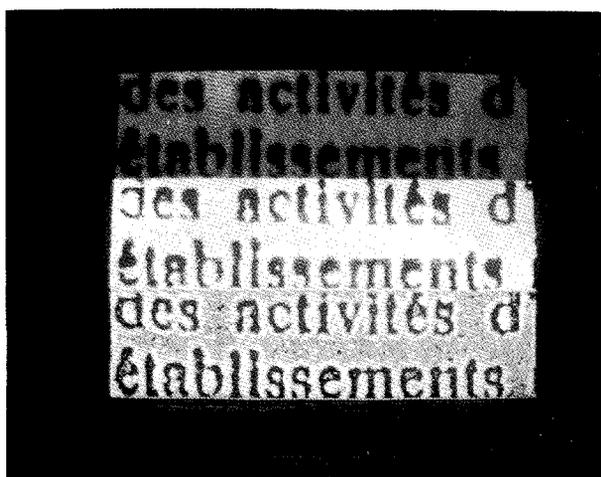


Fig. 14.

Les graphiques de la figure 15 montrent une ligne d'image (ramenée à 64 niveaux de gris) correspondant aux photos des figures 11, 12, 13, 14. En 15a, on observe une réactualisation intempestive (repérée par une flèche) du seuil à l'intérieur de l'objet, en 15b le problème a été éliminé par l'utilisation de la règle (2').

La photo de la figure 16 illustre le fonctionnement de l'algorithme lorsque les objets sont clairs et le fond sombre. Dans cette application, les règles (1'), (2'), (3') et (4) ont été modifiées compte tenu de la nature de l'image. Ainsi la règle (1) est devenue :

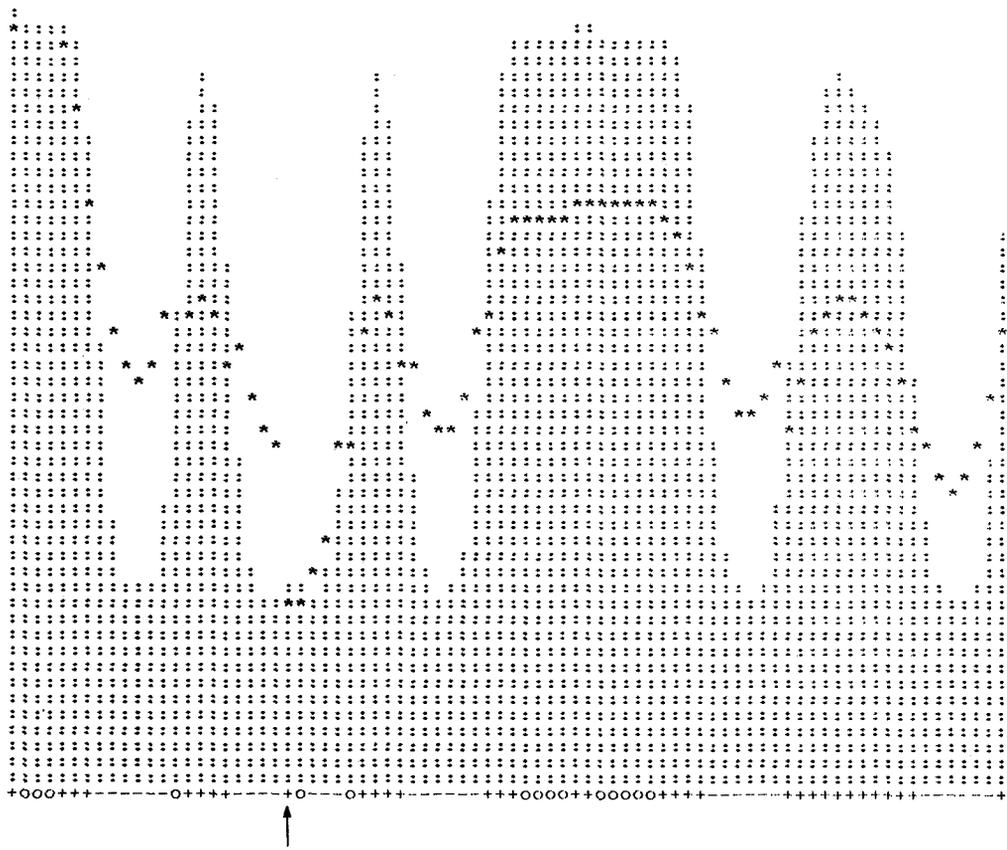
$$(1') \quad \begin{cases} B(j, 0) = 255 \\ N(j, 0) = I_2(j, 0) \end{cases}$$

compte tenu de l'hypothèse que le début de l'image représente le fond sombre. De la même façon, les autres règles sont modifiées pour être symétriques des précédentes.

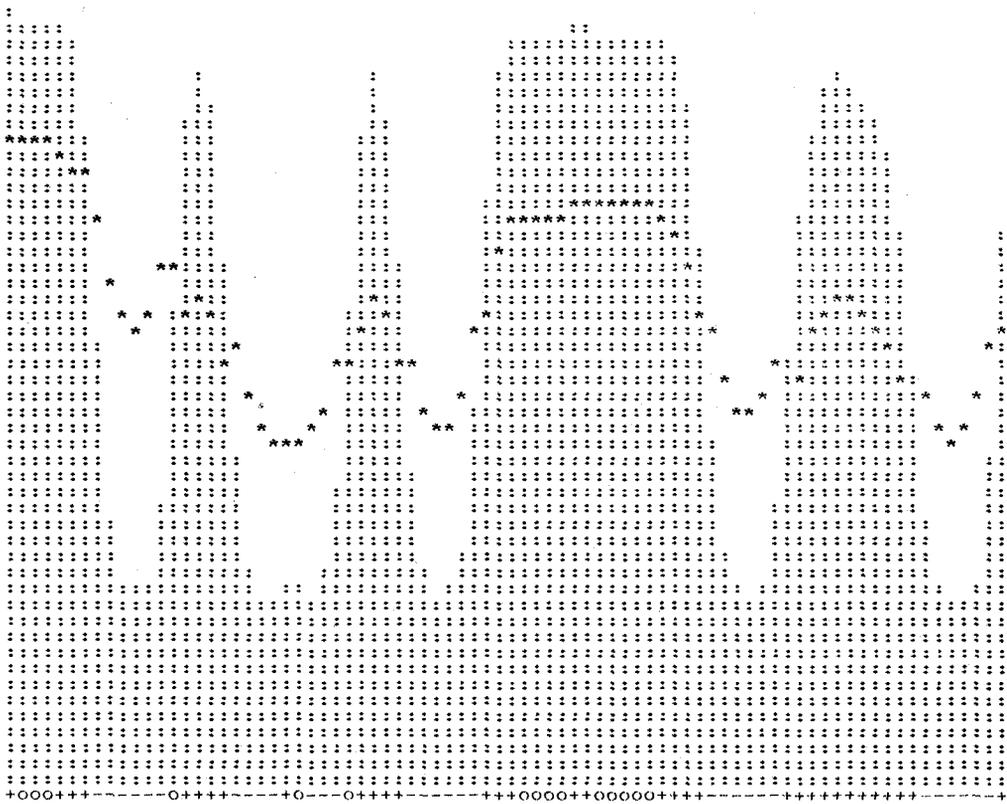
4. Conclusion

Nous avons présenté dans cet article deux solutions au problème de la binarisation des images. Sur le plan conceptuel, la première, que nous avons appelée Binarisation Multifréquence, à notre préférence. Elle n'impose aucune hypothèse sur la nature des objets, mais seulement l'hypothèse que l'image ne fait apparaître que deux classes de pixels, ceux associés au fond et ceux associés aux objets et qu'on peut définir une frontière entre objets et fond à partir d'une mesure de contraste. En outre, elle met en œuvre une analyse multifréquence que l'on retrouve dans certains processus de vision chez les mammifères, et que cherche-t-on à réaliser dans un système de vision si ce n'est un substitut (ne serait-ce que très partiel) au système de vision humain. Sur le plan de la réalisation, cet algorithme présente un inconvénient majeur qui tient, paradoxalement, à sa qualité d'être totalement parallélisable : on peut imaginer son implémentation sur un processeur multicouches, chaque couche

APPLICATIONS



(a)



(b)

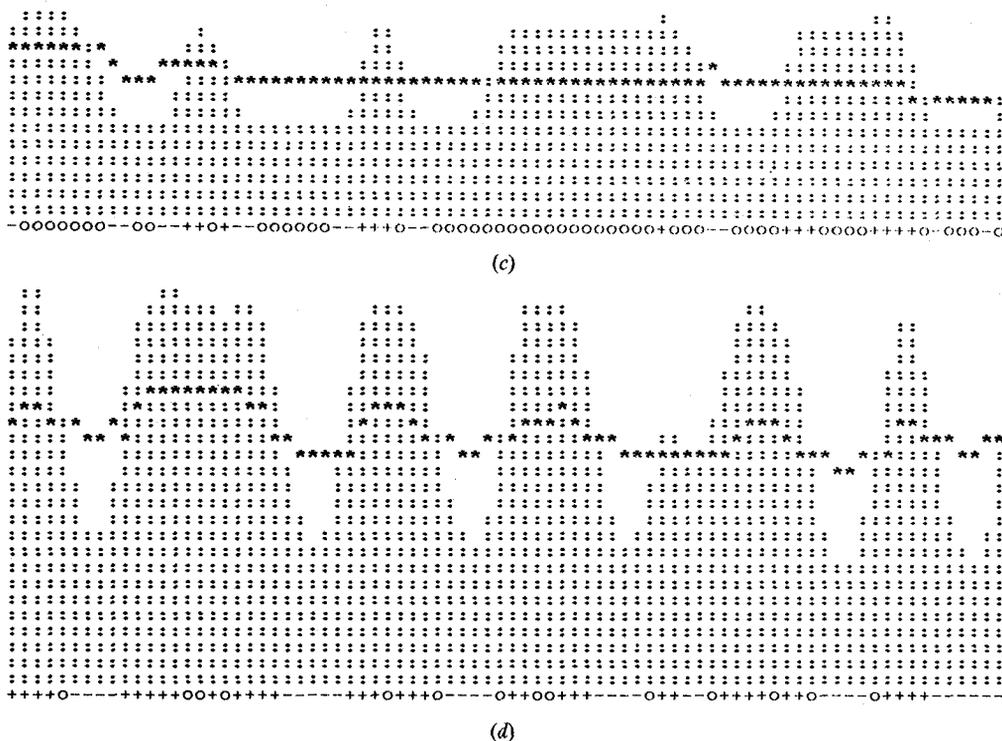


Fig. 15. — Lignes d'images (c) et seuil local (*), correspondant aux photos des figures 11: (a), 12: (b), 13: (c), 14: (d).

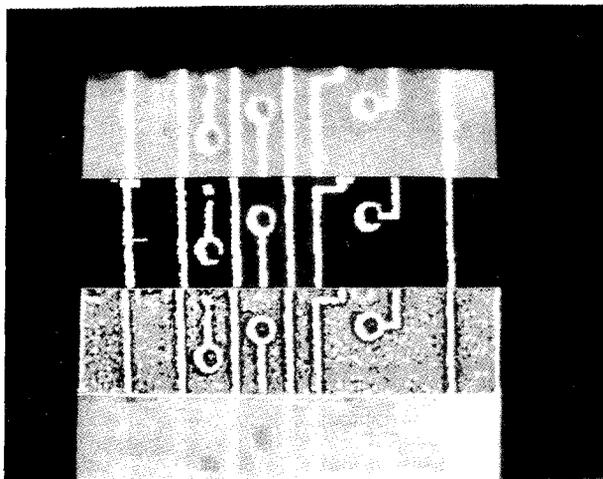


Fig. 16. — Objets clairs sur fond sombre.

correspondant à un niveau de convolution et une couche au niveau de décision, chaque couche comprenant autant de processeurs qu'il y a de pixels à traiter. Malheureusement, la technique actuelle ne permet pas la réalisation aisée de tels processeurs. Une architecture pipe-line peut apporter une solution à ce problème, néanmoins elle conduit à une réalisation matérielle lourde, hors de propos avec les applications visées. Nous avons donc proposé la deuxième solution qui opère par Apprentissage Local d'un Seuil et qui s'est avérée très efficace. Sur le plan de la réalisation pratique, cet algorithme est d'implémentation peu coûteuse, et grâce à une architecture pipe-line, il peut fonctionner à la volée sur le signal vidéo numérisé délivré par le capteur. Par rapport au précédent, il nécessite qu'on choisisse un contraste minimal, mais

dans beaucoup d'applications, binarisation d'images de documents, par exemple, ceci serait un avantage.

Remerciements

Ce travail bénéficie du soutien de la Direction des Études et Recherches d'EDF [6, 7].

Manuscrit reçu le 22 janvier 1988.

BIBLIOGRAPHIE

- [1] R. M. HARALICK et L. G. SHAPIRO, Survey: Image segmentation techniques, *Computer vision, Graphics and Image Processing*, 29, 1986, p. 100-132.
- [2] J. M. WHITE et G. D. ROHRER, Image thresholding for Optical Character Recognition and other applications requiring character image extraction, *IBM J. Res. Development*, 27, 1986, p. 400-411.
- [3] A. R. HANSON et E. M. RISEMAN, *Computer Vision Systems*, Academic Press, New York, 1978, p. 758-768.
- [4] P. J. BURT, Fast filter transforms for image processing, *Computer graphics and Image Processing*, 16, 1981, p. 20-51.
- [5] J. L. CROWLEY et A. C. PARKER, *A representation for shape on peaks and ridges in the difference of low-pass transform*, The Robotics Institute, Carnegie Mellon University, CMI-RI-TR 83-4.
- [6] A. CHEHIKIAN et P. LEFEVRE, *Adaptation de Gris*, Brevet n° 8900444.
- [7] A. CHEHIKIAN et P. LEFEVRE, *Lecture Optique*, Brevet n° 8900445.