

DIXIEME COLLOQUE SUR LE TRAITEMENT DU SIGNAL ET SES APPLICATIONS

1041



NICE du 20 au 24 MAI 1985

UNITE DE PRETRAITEMENT DE FORMES

R. MENGKO, G. BOUVIER

Laboratoire de Traitement d'Images et Reconnaissance de Formes
INP-ENSER, 46 avenue Félix Viallet 38031 - GRENOBLE-CEDEX (France)

RESUME

Nous avons testé une méthode de description de formes par approximation polygonale après filtrage adaptatif. Cette méthode donne des résultats particulièrement intéressants au point de vue réduction de l'information, l'invariance du codage vis à vis des variations du graphisme, l'insensibilité aux divers bruits.

Afin d'obtenir un ensemble de reconnaissance performant en vitesse, il a été réalisé une unité de prétraitement adaptée.

Cette unité effectue l'acquisition et le prétraitement d'une forme en 50 ms. Elle fonctionne en parallèle avec le calculateur et lui délivre une description optimale de la forme.

SUMMARY

We have tested a new shape description method, in which a polygonal approximation is made using an adaptive filter.

Some interesting results were obtained using this method. An information reduction along with an invariant coding in relation with graphic variations and the presence of different noises, is achieved.

To improve speed performance, a specialized preprocessing unit was built, where a total of about 30 ms is required for data acquisition and preprocessing operations.

This unit, working in parallel with a computer, gives a general description of a shape.



1. INTRODUCTION

Dans tout problème de reconnaissance de formes, on distingue généralement deux étapes :

- Perception : pré-traitement de la forme brute, codage.
- Interprétation : recherche d'indices visuels, classification.

Bon nombre d'algorithmes connus ont souvent négligé la première étape au bénéfice de la seconde ce qui conduit dans bien des cas à une impasse.

Nos efforts ont donc porté sur la mise en oeuvre d'algorithmes de perception performante qui donnent une description optimale des formes : non redondante mais contenant toutes les informations nécessaires pour l'interprétation de l'image.

Les tels algorithmes nécessitent une place importante en mémoire de calculateur et pénalisent notablement les vitesses de reconnaissance. Nous avons choisi de réaliser une unité de pré-traitement de formes qui décharge le calculateur de toutes les servitudes d'acquisition, élimine la redondance des formes brutes sans perte d'information sur la topologie tout en autorisant des vitesses de reconnaissance élevées.

2. METHODE PROPOSEE

Pour ne perdre aucune information sur la forme brute, la définition de l'échantillonnage, a priori, est grande voire même redondante. Pour une reconnaissance structurelle, cette redondance est réduite tout en conservant les propriétés topologiques de forme par des méthodes de squelettisation [6, 10, 11, 15, 16] ou d'extraction de contours.

Une segmentation en indices visuels fournit une description topologique de la forme traitée. Cette recherche peut être effectuée soit sur la forme brute, soit sur le squelette ou les contours [6, 8, 9, 10, 12]. Nous avons mis en oeuvre une démarche toute différente : un filtrage adaptatif échantillonne la forme brute. Ce filtrage est une application de la modélisation de certains réseaux nerveux [3, 7]. L'ensemble des nouveaux échantillons est particulièrement réduit par rapport à l'échantillonnage initial mais leur répartition s'adapte au graphisme. Tous les échantillons appartiennent au squelette de la forme et sont placés préférentiellement aux extrémités des segments ou sur les points anguleux.

La connectivité du graphisme est restituée par recherche des segments de droites limités par les nouveaux échantillons inclus dans le graphisme. Ces segments constituent l'approximation polygonale de la forme.

2.1. Filtrage adaptatif

Ce filtrage est une application des propriétés de modéliser le réseau nerveux [7]. Les cellules élémentaires du réseau (neurones) sont disposées dans un plan et réparties selon un maillage que nous avons pris carré. Chaque cellule reçoit des influx, soit :

- excitateurs, d'un domaine d'étendue i, j , du plan image noté $E(m, n)$ avec $a(i, j)$ la matrice des coefficients de pondération des excitations.
- inhibiteurs, d'un domaine d'étendue k, l , du plan de sortie noté $S(m, n)$ avec $\beta(k, l)$ la matrice de pondé-

ration des inhibitions.

Chaque cellule de gain A délivre un signal :

$$S(m, n) = A \left\{ \sum_{i=-a}^{+a} \sum_{j=-a}^a E(m+i, n+j) \alpha(i, j) - \sum_{k=-b}^{+b} \sum_{l=-b}^b S(m+k, n+l) \beta(k, l) \right\}$$

ou encore :

$$S(m, n) = E^*(m, n) - S^*(m, n)$$

$$\text{avec : } E^*(m, n) = \sum_i \sum_j E(m+i, n+j) \cdot A \cdot \alpha(i, j)$$

$$\text{et } S^*(m, n) = \sum_k \sum_l S(m+k, n+l) \cdot A \cdot \beta(k, l)$$

La réponse de chaque cellule est limitée conformément au comportement physiologique par un effet de seuil de sensibilité si $S(m, n) > S_{min}$ alors $S(m, n) = 0$, et par un effet de saturation si $S(m, n) > S_{max}$ alors $S(m, n) = S_{max}$.

Cette expression évoque la réponse d'un filtre spatial bidimensionnel récursif, non linéaire (pour A élevé la réponse est bornée). La réponse de ce type de réseau peut être oscillatoire dans l'espace des cellules sous certaines conditions :

- pour un gain A élevé, la réponse est limitée par S_{min} et S_{max} .
- les lieux des maximums des amplitudes des oscillations se stabilisent dans l'espace après quelques itérations.
- la phase des oscillations se synchronise sur les variations de l'image.
- la période des oscillations est liée à l'étendue du domaine d'inhibition.

Pour un filtrage adaptatif :

- le gain $A = 15$.
- le seuil de sensibilité $S_{min} = 0$.
- la répartition des pondérations des excitations est gaussienne sur un domaine limité aux 7 voisins.
- la répartition des pondérations des inhibitions est constante sur un domaine limité aux 9 voisins.

Dans ces conditions, le réseau est fortement instable. L'amplitude des oscillations est limitée par S_{min} et S_{max} , leur phase se synchronise sur les points singuliers de la forme (extrémités de segments, points anguleux). Les lieux des maximums des amplitudes se stabilisent après quelques itérations. La période des oscillations est de 9 cellules du réseau.

Nous présentons figure 1 (page 3) les résultats obtenus pour quelques formes. Les pixels notés sont les maximums des oscillations et constituent le nouvel échantillonnage $\{E\}$.

2.2. Recherche de connectivité, description polygonale

A chaque doublet d'échantillons $E_{0, f} = \{e_{0, f}, e_{0, f}^*\}$ peut être associé un segment de droite $\{L_{0, f}\}$. Seule une partition $\{L_{0, f}^*\}$ appartient au graphisme de la forme. $\{L_{0, f}^*\}$ définit l'ensemble $\{P\}$ des polygones de description. Si plusieurs polygones P définissent un même indice visuel seul est conservé le polygone qui minimise l'erreur d'approximation au squelette.

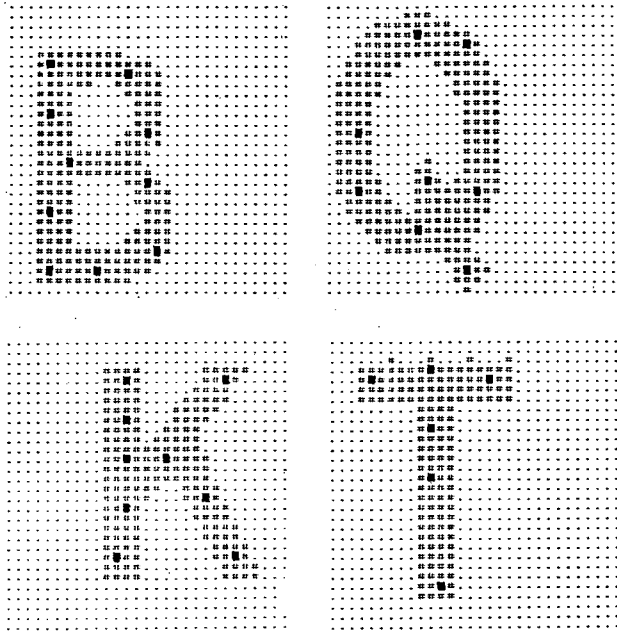


Figure 1 : Echantillonnage adaptatif. Matrice d'acquisition 30x30. Les pixels notés ■ constituent le nouvel échantillonnage.

Tous les échantillons appartiennent à la forme. On démontre que le polygone de longueur maximale :

$$L_{p\alpha} = \sum_{\alpha=0}^f \text{distance} \{F_0, E_f\}$$

donne la meilleure approximation de la forme. Une représentation des polygones est donnée figure 2 pour différentes formes.

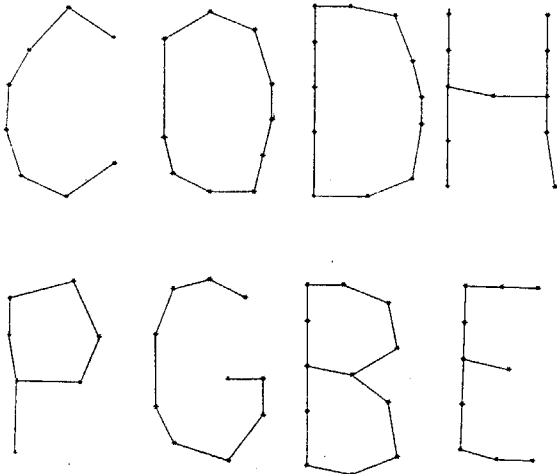


Figure 2 : Description polygonale pour les caractères C, O, D, H, P, G, B, E. Les échantillons {E} sont notés *.

3. APPAREILLAGE REALISE

Cette unité a été étudiée en vue de la lecture optique de documents imprimés multipolices. Ce domaine d'application permet une évaluation aisée des performances des algorithmes.

La structure la mieux adaptée aux contraintes de vitesse et de souplesse d'utilisation est une structure multiprocesseurs, multibus.

Un processeur d'exploitation programmable gère un automate d'acquisition et un processeur neuromimétique. Un bus de contrôle assure les liaisons multiprocesseur, un bus image autorise des transferts rapides d'informations. Une liaison vers un ordinateur HP 9816 S est prévue pour l'exploitation des résultats : recherche d'indices visuels ou classification des formes.

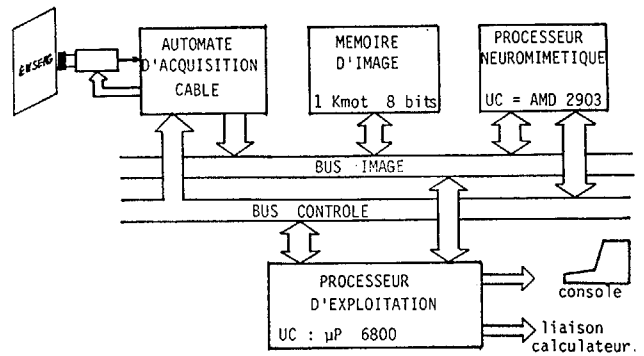


Figure 3 : Synoptique de l'appareillage réalisé

3.1. Automate d'acquisition : figure 4

Le capteur est une caméra TV vidicon 875 lignes. Seule une fraction de l'image est acquise (fig. 5). La fenêtre est échantillonnée en 120 pixels sur 60 lignes. Elle contient 2 caractères complets simultanément. La fréquence d'échantillonnage doit être de 22,4 MHz pour une maille carrée. La digitalisation du signal vidéo fréquence est sur 2 niveaux. Les pixels des deux consécutives sont mémorisés le temps du traitement. Un filtrage passe-bas, une réduction des dimensions de la matrice dans un rapport 2 (60x30) suivis d'un seuillage réduisent les bruits d'acquisition, d'impression ainsi que le nombre de pixels à traiter sans perte notable d'information.

Les transferts de l'image en mémoire s'effectuent pendant les retour-lignes. La mémoire d'image a une capacité de 1 Koctet et un temps d'accès de 100 ns. Les 8 pixels voisins sont mémorisés en mots de 8 bits (1 bit/pixel). Tous les paramètres nécessaires sont fournis par le processeur d'exploitation.

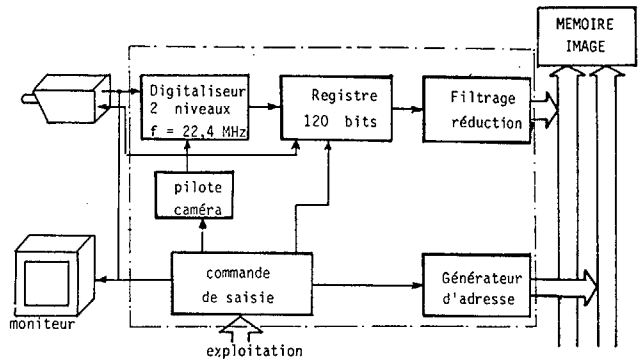


Figure 4 : Automate d'acquisition. Le capteur est une caméra vidicon. L'ensemble est en électronique câblée.

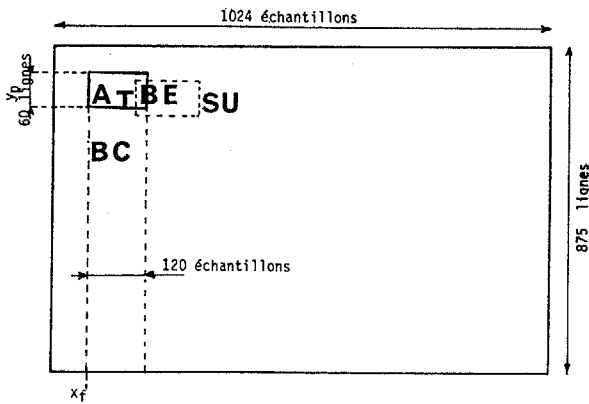


Figure 5 : Une fraction de l'image contenue dans une fenêtre (120 échantillons sur 60 lignes) est acquise à chaque balayage T.V. Le processeur d'exploitation détermine les coordonnées x_f, y_f pour l'analyse du texte.

3.2. Processeur neuromimétique : figure 6

L'image à traiter compte 1800 pixels. Le calcul de $S(m,n) = E^*(m,n) - S^*(m,n)$ avec $i, j, k, l \in \{-3,3\}$ nécessite 176400 adressages, produits et sommations par itération.

Plusieurs propriétés permettent de réduire le nombre d'opérations à effectuer.

- Calcul de $S(m,n)$

* Le gain A est inclu dans les coefficients de pondération :

$$\alpha'(i,j) = A.\alpha(i,j) \quad \beta'(k,l) = A.\beta(k,l)$$

* $E^*(m,n)$ est constant pour toutes les itérations. Cette image convoluée est calculée et mémorisée au cours de la première itération.

* $S(m,n)$ est bornée par S_{min}/S_{max} . $S^*(m,n)$ ne sera calculé que pour les pixels où $E^*(m,n) \neq 0$.

* Si en cours de sommation :

$$S^*(m,n) = \sum_i \sum_j S(m+i, n+j) \cdot \beta'(k,l) > E^*(m,n) \text{ le}$$

calcul peut être arrêté et $S(m,n) = 0$.

- Organisation mémoire

Les pixels sont regroupés dans la mémoire d'image en mots de 8 bits. Pour un mot de la ligne i, il est possible de calculer deux termes :

$$E_i^*(m,n) = \sum_{j=-3}^{j+3} E(m+i, n+j) \alpha'(i,j) \text{ et}$$

$$E_i^*(m,n_{+1}) = \sum_{j=-3}^{j+3} E(m+i, n_{+1} + j) \alpha'(i,j), \text{ puis sur}$$

7 lignes consécutives $i = -3$, à $i = +3$

$$E^*(m,n) = \sum_{i=-3}^{i+3} \sum_{j=-3}^{j+3} E(m+i, n+j) \alpha'(i,j)$$

$$E^*(m,n_{+1}) = \sum_{i=-3}^{i+3} \sum_{j=-3}^{j+3} E(m+i, n_{+1} + j) \alpha'(i,j)$$

Afin de réduire le nombre d'adressage, nous avons choisi d'effectuer les calculs sur 2 mots mais de ne déterminer simultanément que 8 termes $E(m, n-3) \dots$

$E(m, n+4)$ pour rester dans le format des mémoires.

L'algorithme de calcul de $S^*(m,n)$ est identique mais les calculs sont effectués en tenant compte des remarques simplificatrices précédentes. Le temps nécessaire au filtrage d'une forme est considérablement réduit par rapport au calcul direct et complet. Ce processeur prend en charge l'adressage des mémoires, ce qui nécessite 13 bits d'adresse. Quatre microprocesseur en tranche AMD 2903 sont nécessaires. Plusieurs mémoires sont associées à cette unité : mémoire de sortie $S(m,n)$, tableaux des coefficients $\alpha'(i,j)$ $\beta'(k,l)$. Après filtrage ce même processeur détermine les segments d'approximation polygonale de la forme.

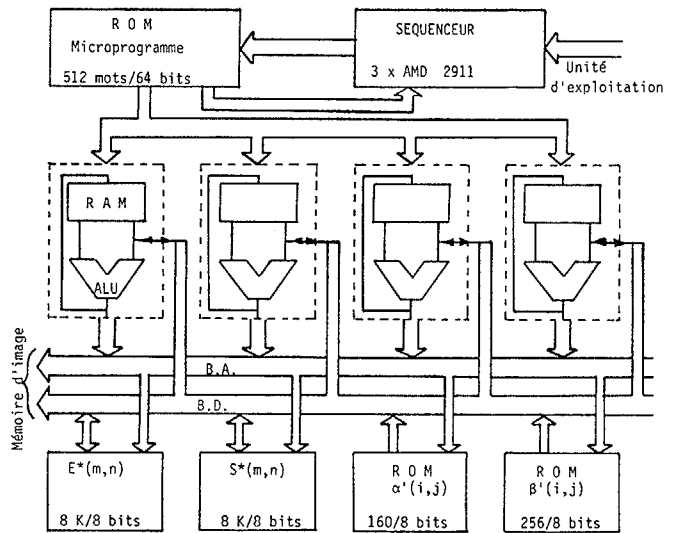


Figure 6 : Processeur neuromimétique.

3.3. Processeur d'exploitation

Ce processeur a pour tâche la gestion de l'automate d'acquisition, du processeur neuromimétique et des dialogues avec l'environnement. L'unité centrale adoptée est un microprocesseur 6800 auquel est adjoint 1 Koctet de mémoire RAM et 4Koctets de mémoire ROM, quatre interfaces parallèles pour les dialogues processeurs et un interface série pour la liaison avec le calculateur.

Ce processeur détermine les coordonnées de la fenêtre d'acquisition pour le suivi des lignes de texte ; le cadrage des caractères quelle que soit la fonction du document face à la caméra.

4. PERFORMANCES DE L'APPAREILLAGE

Deux caractères peuvent être acquis et filtrés simultanément. L'image est mémorisée dans la mémoire d'image avec deux lignes TV de retard par rapport au signal vidéo fréquence. Le filtrage spatial débute dès la fin de l'acquisition. Le temps nécessaire pour ce filtrage et l'approximation polygonale des caractères est de 50 ms. Comme ce processeur travaille en pipe-line avec le système d'acquisition par l'intermédiaire de la mémoire d'image, il est possible de traiter 20 acquisitions par seconde ce qui représente une moyenne de 30 caractères/seconde.

Les codes des formes se prêtent bien à la recherche d'indices visuels et à la classification. Nous présentons

tons figure 7 les résultats obtenus pour quelques caractères.

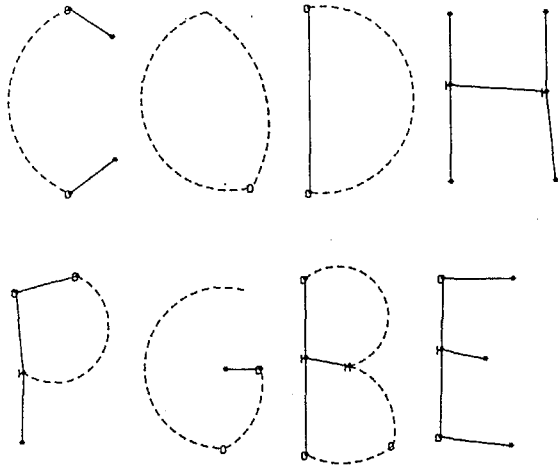


Figure 7 : Indices visuels.

courbes traits pointillés, segments traits pleins, H intersection de segment, o changement de direction, * extrémités de segments.

CONCLUSIONS

Par le choix des algorithmes mis en oeuvre et des techniques employées, cette unité donne des résultats intéressants tant du point de vue description des formes que sur le plan vitesse de traitement. L'approximation des formes se prête particulièrement bien à la segmentation. L'information est réduite mais toutes les propriétés topologiques sont conservées. Les descriptions sont très tolérantes vis à vis des variations de graphisme, de l'homothétie, de la translation, de la rotation et du bruit. Une vitesse de traitement de 30 caractères par seconde est largement suffisante pour beaucoup d'applications.

BIBLIOGRAPHIE

[1] M. BERTHOD
Expérimentation sur l'échantillonnage de tracés manuscrits en temps réel.
Process. 1er congrès AFCET-IRIA, Chatenay-Malabry, pp. 421-426, février 1978

[2] B. BRETAGNOLLE, C. RUBAT DU MERAC
ROMUALD : Un multiprocesseur interactif pour la saisie et le traitement.
Congrès AFCET-IRIA - Reconnaissance des Formes et Intelligence Artificielle, Toulouse, sept. 1979

[3] G. BOUVIER, J. HERAULT
Méthode de description de caractères par approximation polygonale.
Process. 8e colloque GRETSI, Nice, Tome 2, pp. 733-741, juin 1981

[4] S. CASTAN, J.Y. LATIL
Capteur d'images de Toulouse (CAPITOU)
Automatisme - tome XXI, pp. 90-93, mars-avril 1976

[5] J.P. CHIEZE, O. FAUGERAS
Un système interactif de traitement d'images.
Congrès AFCET-IRIA, Recon. des Formes et Int. Art., Châtenay-Malabry, février 1978

[6] B. DUPOUY, L. MICLET, S. TOHME
Compression d'image de l'écriture manuscrite.
Congrès AFCET-IRIA, Recon. des Formes et Intel. Art., Toulouse, septembre 1979

[7] J. HERAULT, G. BOUVIER, A. CHEHIKIAN
A new algorithm for image processing based on the properties of neural nets.
Journal de Physique-Lettres-41, pp. 75-78, 1980

[8] G. LORETTE
Algorithmes de représentation de lignes polygonales.
Process. 2e congrès AFCET-IRIA, Toulouse, pp. 158-165, septembre 1979

[9] W.W. LOY, T.D. LANDAU
An on-line procedure for recognition on hand-written alphanumeric characters.
Proc. of the 5th intl. Conf. on Pattern Recognition, Miami, ISA USA (1980)

[10] L. MICLET
Cours de reconnaissance de formes.
Ecole Nationale Supérieure des Télécommunications (1978)

[11] N.J. NACCACHE, R. SHINGHAL
SPTA : A proposed algorithm for thinning binary patterns
IEEE Transactions on Systems, Man, and Cybernetics, Vol. SMC 14, n° 3, mai-juin 1984

[12] M. PAVEL
Squelettes de catégories d'images.
Note Comptes-Rendus Acad. Sc., A-282, 751-754, (1976)

[13] J. PAVLIDIS
A review of algorithm for shape analysis. Computer graphics and image processing 7, 243-258, (1978)

[14] W.K. PRATT
Digital image processing.
Wiley Interscience, New-York, 1978

[15] Z.Q. WU, A. ROSENFELD
Filtered projection as an iad in corner detection.
Pattern recognition, vol. 16, n° 1, 31-38, (1983)

[16] S. TOHME
Prétraitement du chiffre manuscrit.
Process. 1er congrès AFCET-IRIA, Châtenay-Malabry, 568-576, février 1978

