

NEUVIEME COLLOQUE SUR LE TRAITEMENT DU SIGNAL ET SES APPLICATIONS

565



NICE du 16 au 20 MAI 1983

PROCESSEUR ADAPTATIF DE TRAITEMENT D'IMAGE BASÉ SUR UN MODELE DE RESEAU NERVEUX

J. HERAULT, G. BOUVIER, A. CHEHIKIAN

Laboratoire de Traitement d'Images et Reconnaissance des Formes
E.N.S.E.R. - 23 avenue des Martyrs 38031 - GRENOBLE-CEDEX

RESUME

Ce processeur est destiné au prétraitement de caractères alphanumériques non standards en vue de leur reconnaissance. Il s'applique à des images binaires sur-échantillonnées afin de conserver de manière sûre tous les détails graphiques riches en information.

La réduction d'information au minimum utile à la reconnaissance fait appel à un algorithme tiré des propriétés des réseaux nerveux à inhibitions latérales récurrentes : sous certaines conditions, ce type de réseau se comporte comme un échantillonneur spatial adaptatif et optimal. La maille d'échantillonnage n'est pas rigide, elle s'adapte automatiquement aux dimensions spatiales de la forme d'entrée et se cadre sur les discontinuités de celle-ci.

Le processeur câblé simule un réseau de 4096 neurones à 400 000 interconnexions et traite une image de 4096 pels en quelques millisecondes.

SUMMARY

This processor is intended to preprocessing of alphanumeric characters in the aim of their recognition. It applies to oversampled binary images in order to preserve every graphic detail where lies the most important information.

Information is reduced to a useful minimum for recognition by means of an algorithm issued from the properties of neural networks with "lateral backward inhibitions". Under some conditions, this kind of networks realizes an adaptative spatial sampling with optimal characteristics. The sampling mesh is not fixed, it adapts automatically to the spatial content of the input image and is centered on the discontinuities.

The hard-wired processor simulates a network at 4096 neurons with 400 000 interconnections and processes a 4096 pixel image within a few milliseconds.



PROCESSEUR ADAPTATIF DE TRAITEMENT D'IMAGE BASÉ SUR UN MODELE
DE RESEAU NERVEUX

La copie du monde biologique par les ingénieurs constitue un bon moyen d'améliorer les performances de matériels existants ou d'en créer de nouveaux. Une fois dépassé le stade des simplifications abusives typiques des dernières décennies, on voit actuellement se dessiner une ère prometteuse, grâce aux progrès tant de l'investigation physiologique que des moyens techniques dont disposent les ingénieurs. C'est probablement au niveau du concept de *capteur* et de *prétraitement* que le modèle biologique apporte une idée nouvelle la plus directement exploitable [1].

Dans le système nerveux, l'*opérateur élémentaire* est représenté par le neurone, peu fiable et peu précis et la *fonction* traitement de l'information est représentée par l'architecture des interconnexions entre les cellules, riche multiple et redondante.

Ce papier se compose de quatre parties. Les deux premières sont relatives à l'aspect neurophysiologique : l'opérateur neurone et les réseaux nerveux à "inhibitions latérales récurrentes" (I.L.R.), les deux autres montrent l'application au traitement d'images : le processeur câblé équivalent I.L.R. et les résultats de traitement de caractères alphanumériques.

I - L'OPÉRATEUR NEURONE

La morphologie du neurone est particulièrement bien adaptée au traitement et au transfert des informations [2, 3]. Le corps cellulaire (soma) est entouré d'une multitude de prolongements destinés à réaliser des interconnexions (synapses) avec les cellules voisines. Le motoneurone (fig. 1) en est le type.

Les influx afférents, captés au niveau des dendrites dans l'espace environnant produisent des variations locales du potentiel transmembranaire sommées dans le corps cellulaire. Cette sommation temporelle, pondérée par un effet de la distance synapse-soma, est algébrique, compte-tenu des natures excitatrices et inhibitrices des liaisons synaptiques mises en jeu. Lorsque le potentiel somatique dépasse un certain seuil, le neurone émet une série d'influx ou "potentiels d'action" (P.A.). Transmis par l'axone et l'arborisation terminale, ces P.A. sont distribués aux cellules voisines par l'intermédiaire d'un médiateur chimique qui assure le caractère excitateur ou inhibiteur des contacts synaptiques.

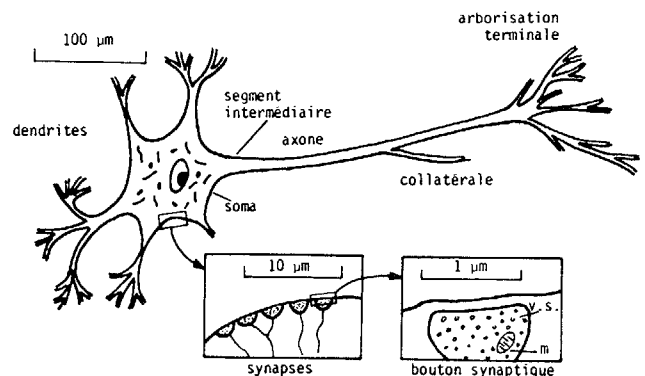


Fig. 1 : Aspect schématique d'un motoneurone.

La fréquence de répétition des influx est une fonction croissante du potentiel somatique, donc de l'intégration *spatiale* et *temporelle* des signaux afférents à l'arbre dendritique. Cette fonction présente globalement une allure logarithmique entre une zone de seuil et une zone de saturation (fig. 2.a.).

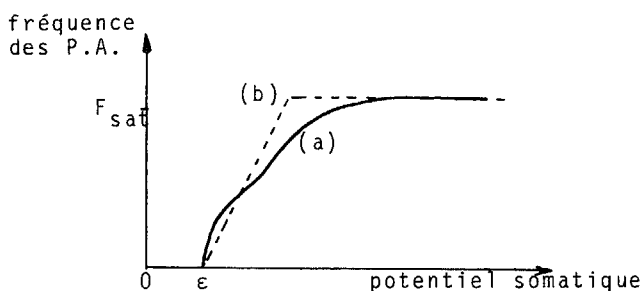


Figure 2.

- a) Caractéristique de transfert du neurone (unités arbitraires).
b) Caractéristique idéalisée souvent utilisée en simulation

ϵ : seuil d'excitation, F_{sat} : fréquence de saturation.

Dans les problèmes de simulation qui nous intéressent, nous prendrons l'approximation suivante :

- le neurone est un élément analogique à seuil.
- ses grandeurs d'entrée et de sortie sont des fréquences d'impulsions, c'est-à-dire des variables continues.

De plus, pour l'étude des propriétés purement spatiales des réseaux, ces grandeurs seront indépendantes du temps. Dans ces conditions, la fréquence S des influx en sortie sera considérée comme proportionnelle à la somme algébrique pondérée des fréquences d'influx e_i^+ des entrées excitatrices et e_k^- des entrées inhibitrices, entre une valeur de seuil ϵ et une valeur de saturation Sat , selon l'expression :

$$S = A \left\{ \sum_i \alpha_i e_i^+ - \sum_k \beta_k e_k^- - \epsilon \right\} \quad (1)$$

avec $S = 0$ si le second membre est ≤ 0 et $S = Sat$ si le second membre est $\geq Sat$. α_i et β_k sont des coefficients positifs de pondération. A et ϵ sont le gain et le seuil du neurone.

PROCESSEUR ADAPTATIF DE TRAITEMENT D'IMAGE BASÉ SUR UN MODELE DE RESEAU NERVEUX

II - RÉSEAUX A INHIBITIONS LATÉRALES RÉCURRENTES

Les réseaux nerveux à inhibitions latérales récurrentes (I.L.R.) sont très répandus dans le système nerveux. On les rencontre aussi bien dans les systèmes sensoriels, bulbe olfactif ou rétine [4], où ils effectuent un prétraitement au niveau de la fonction capteur [1] que dans les systèmes effecteurs (commande des muscles) [3]. Ils sont particulièrement abondants au niveau des cortex cérébral et cérébelleux [5] où leur fonction, plus élaborée, se traduit par une architecture plus complexe en plusieurs couches de cellules.

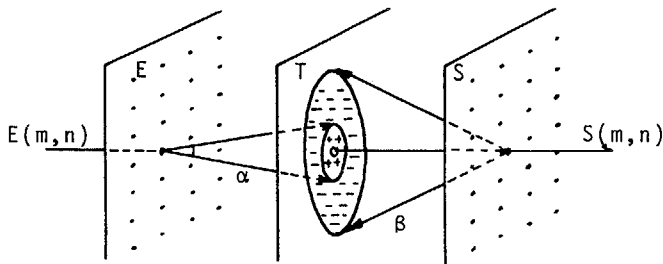


Figure 3 : Représentation schématique de la structure d'inhibitions latérales récurrentes (I.L.R.).

E : plan d'entrée avec un échantillonnage selon m, n .
 T : plan de traitement, zone des somas des neurones.
 S : plan de sortie.
 A chaque point (m,n) correspond un neurone, les voies directes (α) et récurrentes (β) se distribuent chacune sur un nombre élevé de neurones.

D'un point de vue formel, ces réseaux représentent des structures de filtres récursifs bidimensionnels d'ordre élevé comportant deux parties : à réponse impulsionnelle finie pour l'une, infinie pour l'autre. La différence essentielle avec les filtres récursifs réside dans la non-linéarité de l'élément neurone (les sorties sont comprises entre les deux valeurs 0 et Sat) et dans le fait que les voies de retour sont strictement affectées du même signe négatif (figure 3).

Dans l'approximation définie par la formule (1), la fonction de sortie $S(m,n)$ du réseau de la figure 4 s'exprime, en fonction du signal d'entrée $E(m,n)$ par :

$$S(m,n) = A \left\{ \sum_{i=-a}^{+a} \sum_{j=-a'}^{+a'} \alpha(i,j) \cdot E(m-i, n-j) - \sum_{k=-b}^{+b} \sum_{l=-b'}^{+b'} \beta(k,l) \cdot S(m-k, n-l) - \epsilon \right\} \quad (2)$$

avec pour restrictions : $0 \leq S(m,n) \leq \text{Sat}$.

a, a' : limites de définition des coefficients α positifs ou négatifs.
 b, b' : limites de définition des coefficients β positifs.

Cette expression amène trois remarques a priori sur les fonctions d'un tel réseau :

- amélioration du rapport signal/bruit en raison du lissage par la convolution $\alpha * E$ et du seuillage par ϵ .
- effet de filtrage passe-haut en raison de la rétroaction par le filtre passe-bas β (augmentation du contraste spatial).
- en outre, en raison de la non-linéarité de seuil, la convolution $\beta * S$ peut être considérée comme un seuil adaptatif en fonction du signal de sortie (donc de l'information extraite).

Dans certains cas, le système d'équations (2) est instable et la sortie $S(m,n)$ présente des oscillations spatiales (figure 4). L'amplitude de ces oscillations est limitée entre 0 et Sat, leur période dépend du gain A, des coefficients $\beta(k,l)$, des limites b et b' et des dimensions locales de la forme d'entrée $E(m,n)$.



Figure 4 : Comportement d'un réseau I.L.R. en présence d'une entrée E uniforme.
 S₁ : accentuation des contrastes. S₂ : début d'instabilité. S₃ : fonction échantillonnage.

A l'inverse des méthodes classiques de filtrage bidimensionnel [6], cette instabilité présente un intérêt majeur : elle constitue une sorte d'échantillonnage adaptatif du signal d'entrée dont les propriétés sont remarquables, spécialement dans le cas d'images d'entrée à structures de traits comme les caractères alphanumériques.

- i - Les "échantillons" sont des groupes de points où le signal $S(m,n)$ non nul est proportionnel à $E(m,n)$.
- ii - Ils sont centrés sur le squelette de la forme d'entrée.
- iii - Ils se placent en priorité sur les discontinuités du graphisme d'entrée : extrémités de segments, angles vifs.



PROCESSEUR ADAPTATIF DE TRAITEMENT D'IMAGE BASÉ SUR UN MODELE
DE RESEAU NERVEUX

- iv - Leur période, définie *approximativement* par les paramètres du réseau, *s'adapte* aux dimensions de la forme d'entrée.

Les premières études du comportement des réseaux I.L.R. ont été réalisées en simulation numérique. Elles ont permis de mettre en évidence une nouvelle méthode de traitement des images à structure de traits particulièrement intéressante du point de vue de la réduction d'information [7] : un caractère alphanumérique est décrit par un maximum de 15 échantillons toujours répartis d'une manière optimale mettant en évidence les extrémités de segment et les angles vifs. Cette dernière propriété permet de lever tous les doutes classiques en reconnaissance de formes, grands consommateurs de temps de calcul [8].

Pour toutes ces raisons, nous avons décidé de réaliser un processeur câblé capable de calculer le système d'équations (2) en un temps minimal, tâche toujours très longue avec un calculateur non spécifique.

III - LE PROCESSEUR CABLÉ

Principe : Les simulations numériques ont montré que le phénomène d'échantillonnage n'apparaissait qu'à partir d'une valeur suffisamment élevée du gain A des neurones et conservait les mêmes caractères pour des valeurs aussi grandes que l'on voulait. Cette propriété est très importante car la fonction amplification se transforme ainsi en une fonction comparaison, ce qui a pour effet de binariser le signal de sortie : $S(m,n)$ prend alors les valeurs 0 ou Sat selon que $(\alpha * E - \epsilon)$ est inférieur ou supérieur à $\beta * S$.

De la même manière, nous prendrons pour signal d'entrée un signal binaire $E(m,m)$, ce qui est compatible avec le type d'images que constituent les caractères alphanumériques : des ensembles de traits noirs sur fond blanc.

Cette simplification a deux conséquences très importantes dans le cas d'un processeur câblé. D'une part, la mise en mémoire des signaux d'entrée ou de sortie se réalise en affectant seulement 1 bit par point élémentaire (pel) d'image ; d'autre part, les produits de convolution discrète du genre :

$$\sum_{i,j} \alpha(i,j) \cdot E(m-i, n-j)$$

sont convertis en simples sommes des coefficients $\alpha(i,j)$ associés aux valeurs non nulles de $E(m-i, n-j)$.

La structure du processeur câblé est décrite par la figure 5 :

- * Les images d'entrée et de sortie sont stockées dans des mémoires circulantes constituées de 32 registres à décalages de 128 bits chacun, chaque image présente donc 4096 points élémentaires.
- * Les fenêtres de calcul des produits de convolution (coefficients α et β) ont une dimension de 7×7 pels.
- * Les coefficients $\alpha(i,j)$, $\beta(k,l)$, les pondérations A et B de leurs produits de convolution et le seuil ϵ sont contrôlables par l'expérimentateur.
- * Le signal de sortie du comparateur est réinjecté dans l'image de sortie $S(m,n)$ au centre de la fenêtre de traitement des coefficients.
- * Un séquenceur, non représenté, contrôle les diverses opérations de chargement et de lecture des registres, le nombre d'itérations du calcul et les divers signaux de service.

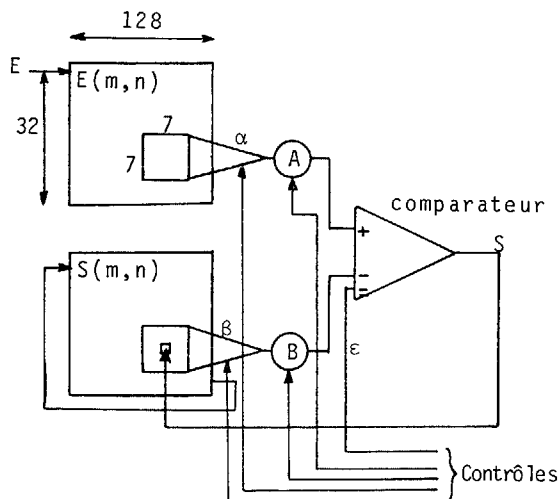


Figure 5 : Schéma de principe du processeur câblé.

Mémoires d'image et fenêtres de calcul : Dans la réalisation pratique, les images d'entrée et de sortie sont stockées dans 32 registres à décalage de 256 bits chacun, les cases paires sont affectées à l'entrée, les cases impaires à la sortie. La fréquence maximale de décalage de ces registres (circuits AM 2802 de 256×4 bits) étant de 10 MHz, la vitesse maximale de traitement est de 5 MHz, soit 200 ns par pel.

Sur les phases paires de l'horloge de décalage, on dispose de 32 sorties des registres qui constituent une colonne de 32 pels de l'image d'entrée, sur les phases impaires, on dispose de même, des 32 pels correspondants de l'image de sortie.

Les 49 points des fenêtres de calcul sont obtenus de la manière suivante : les sorties des 7 derniers registres (colonne de 7 pels) sont appliquées aux entrées de 7 registres à décalages de 7 cases à entrée série et sorties parallèles. Un groupe de 7 registres constitue la fenêtre des 49 valeurs $E(m-i, n-j)$, un

PROCESSEUR ADAPTATIF DE TRAITEMENT D'IMAGE BASÉ SUR UN MODELE
DE RESEAU NERVEUX

second constitue la fenêtre $S(m-k, n-l)$, i, j, k, l appartenant à la suite des entiers de -3 à $+3$. Le premier groupe de registre ne reçoit que les phases paires de l'horloge, le second, les phases impaires. On dispose ainsi, à chaque période d'horloge, des 49 points de chaque fenêtre en parallèle.

Opérateurs de convolution et calcul de $S(m,n)$: Chacune des 49 sorties des fenêtres de calcul est un signal binaire auquel est affecté un coefficient $\alpha(i,j)$ ou $\beta(k,l)$. Elle constitue un signal d'activation d'un générateur de courant $I(i,j)$ ou $I'(k,l)$ dont la valeur, définie de manière externe par l'utilisateur, correspond aux coefficients $\alpha(i,j)$ ou $\beta(k,l)$. Le produit de convolution discrète $\alpha * E$ est alors obtenu en effectuant la somme des courants $I(i,j)$ affectés des poids 1 ou 0 selon les valeurs de $E(m-i, n-j)$, de même pour $\beta * S$. Les courants I et I' obtenus débitent sur deux résistances identiques et les tensions qui en résultent sont appliquées aux entrées d'un comparateur rapide qui donne en sortie 1 ou 0 selon que le signe de $I - I'$ est positif ou négatif. Des coefficients A et B permettent d'ajuster les poids relatifs des ensembles $I(i,j)$ et $I'(k,l)$.

Cette technique de calcul par sommation de courants et comparateur rapide permet d'obtenir les 98 sommes de produits du second membre de l'équation (2) en 50 ns, ce qui permet d'obtenir un fonctionnement de l'opérateur à la vitesse maximale de décalage des registres.

Séquenceur : Le séquenceur contrôle les opérations de chargement de la mémoire d'entrée et de lecture de la mémoire de sortie, par un processeur hôte d'acquisition d'image par caméra TV [9].

Il contrôle bien évidemment les opérations de synchronisation des différentes tâches à réaliser : repérage du point courant (m,n) traité, prise en compte d'une garde de 3 pels en bordure de l'image, mise à jour du nombre d'itérations du calcul, etc...

De par la structure des mémoires d'entrée et de sortie (cases paires, cases impaires), il est facile de procéder à un nouveau traitement de l'image de sortie par le même processeur en échangeant les phases des horloges des deux registres constituant les fenêtres de calcul : les signaux $S(m,n)$ issus d'un premier calcul prennent alors la place de la précédente image d'entrée. Cette opération est gérée très facilement par le séquenceur.

Les performances obtenues par ce processeur câblé sont intéressantes du point de vue de la vitesse de calcul : une itération de calcul correspond à 4096 calculs de 98 sommes de produits soit 401408. Elle est réalisée en moins d'une milliseconde.

D'une manière pratique, l'effet d'échantillonnage spatial est obtenu en quelques itérations et le résultat est stable. Nous avons cependant remarqué que, pour certaines configurations des paramètres α et β , l'échantillonnage obtenu lors de la première itération, était largement suffisant pour permettre une exploitation directe des résultats.

IV - PERFORMANCES DU PROCESSEUR

Les images suivantes ont été obtenues de la manière suivante :

- coefficients α : lissage sur une fenêtre de 5x5 pels.
- coefficients β : valeur maximale dans une couronne de 2 pels sur le pourtour de la fenêtre 7x7, valeur nulle au centre sur une fenêtre 3x3.
- calcul en une seule itération.

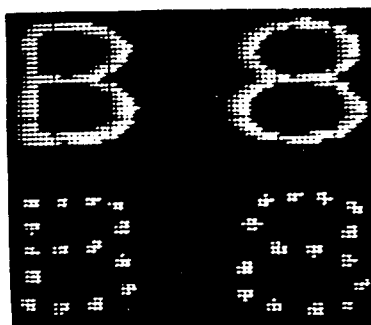


Figure 6 : Détail des signaux.

En haut : image d'entrée acquise par la caméra TV puis échantillonnée.

En bas : résultat du traitement.

Noter la position des groupes de points par rapport au signal d'entrée.

La figure 7 montre le type de traitement obtenu pour diverses formes de caractères ou du graphisme. Noter la qualité du traitement dans le cas des caractères "un" manuscrit où l'image sous la caméra TV a été volontairement très dégradée.

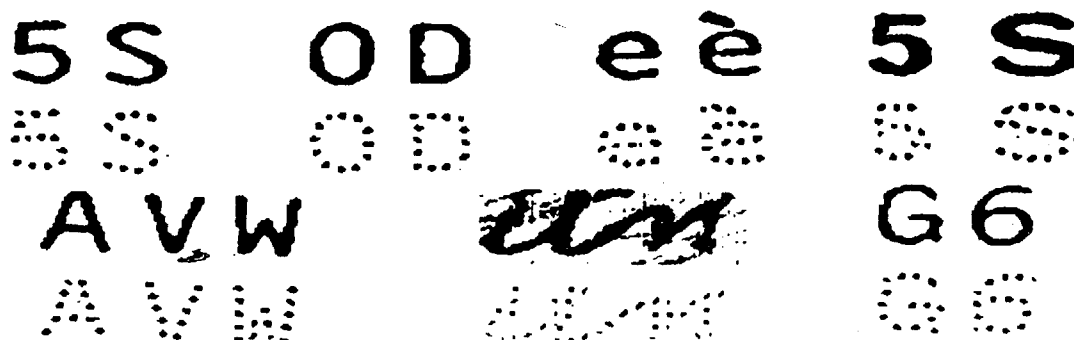


Figure 7 : Quelques exemples typiques du traitement obtenu.



PROCESSEUR ADAPTATIF DE TRAITEMENT D'IMAGE BASÉ SUR UN MODELE
DE RESEAU NERVEUX

V - CONCLUSIONS

Le processeur que nous avons décrit présente une structure et un mode de fonctionnement extrapolé à partir des principes généraux des réseaux nerveux à inhibitions latérales récurrentes. Il serait osé de prétendre que certains circuits du cerveau fonctionnent selon le même même mode, du moins en l'état actuel des connaissances physiologiques. Il n'en reste pas moins intéressant car sa propriété d'échantillonnage spatial adaptatif en fait un outil très appréciable en traitement des images à structure de traits.

Partant d'un échantillonnage initial important, il permet d'acquérir le maximum d'informations sur la position des discontinuités (angles vifs, extrémités de segments), puis il opère une réduction de la redondance en conservant précisément ces zones à contenu informationnel maximal. Utilisé en tant qu'organe de prétraitement dans une machine à reconnaissance des formes, il permet de lever un maximum de doutes en fournissant un nombre minimal de paramètres, ce qui doit conduire nécessairement à ces classifieurs de structure très simple.

BIBLIOGRAPHIE

- [1] J. HERAULT
Le capteur biologique, un modèle de performances et de fiabilité à la portée des technologies actuelles.
Le Vide, les couches minces, 211, mars-avril 1982
- [2] P. LAGET
Biologie et physiologie des éléments nerveux
Masson & Cie, Paris (1970)
- [3] P. BUSER, M. IMBERT
Neurophysiologie fonctionnelle
Hermann, Paris (1975)
- [4] H.K. HARTLINE, F. RATLIFF
Spatial and temporal aspects of retinal inhibitory interactions
in : Studies on excitation and inhibition in the retina Chapman-Hall, London (1974)
- [5] M. HASSUL, P.D. DANIELS
Cerebellar dynamics ; the mossy fiber input
IEEE Trans. BME 24, 449-455 (1977)
- [6] R. MERSEREAU, D. DUDDGEON
Two dimensional digital filtering
Proc. IEEE 63, 610-623 (1975)
- [7] J. HERAULT, G. BOUVIER, A. CHEHIKIAN
A new algorithm for image processing based on the properties of neural nets
J. Physique - Lettres 41 (1980)
- [8] Z.Q. WU, A. ROSENFELD
Filtered projections as an aid in corner detection
Pattern recognition, vol. 16, n° 1, 31-35 (1983)
- [9] A. CHEHIKIAN, G. BOUVIER
Appareil d'acquisition de textes pour machine à reconnaissance des formes
Brevet Anvar n° 73.46.285 (1973)