

50376  
1986  
91

# THÈSE

présentée à

L'UNIVERSITE DES SCIENCES ET TECHNIQUES DE LILLE FLANDRES ARTOIS

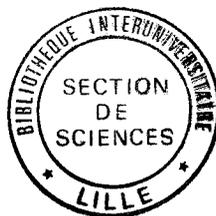
pour obtenir le titre de

**DOCTEUR DE 3<sup>ème</sup> CYCLE**

par

Mohammed BENJELLOUN

## **ETUDE THEORIQUE ET EXPERIMENTALE DU FILTRAGE NUMERIQUE DE L'IMAGE D'UN TEXTE EN RELIEF BRAILLE POUR SA TRANSCRIPTION EN TEXTE NOIR**



Soutenue le 17 Juin 1986 devant la Commission d'Examen

Membres du Jury :

G.	SEGUIER,	Président
J.P.	DUBUS,	Rapporteur
Y.	CROSNIER,	Examineur
M.	DELBOS,	Examineur
A.	LEBRUN,	Examineur
Y.	MOSCHETTO,	Examineur

à Sylvie, ma femme

à ma Famille,

## R E M E R C I E M E N T S

- : - : - : - : - : - : - : - : - : - : - : - : - : - : - : - : -

*Cette étude a été effectuée au sein de l'équipe "Modélisation et Instrumentation Électronique" du Centre de Recherche Sciences des Matériaux & Techniques de Construction de l'Université des Sciences et Techniques de Lille 1.*

*Je tiens à exprimer tout d'abord ma profonde reconnaissance envers Monsieur le Professeur J.P. DUBUS qui a assuré la direction de cette étude et pour son aide constante et efficace à la réalisation de ce travail. Je le remercie de la confiance qu'il a investi en moi, du soutien qu'il m'a apporté.*

*Je suis particulièrement reconnaissant envers Monsieur le Professeur G. SEQUIER pour l'honneur qu'il me fait de présider la commission d'examen, de juger mon travail, ainsi que pour le soutien matériel qu'il m'a apporté.*

*Je remercie Monsieur le Professeur A. LEBRUN, Directeur et Père du CREMAT pour lequel il est resté, en quelque sorte, un directeur spirituel. Son dynamisme et son intérêt intarissable pour la recherche scientifique sont pour moi un réel exemple et je suis très fier de le compter parmi les examinateurs.*

*Je tiens à exprimer ma reconnaissance à Monsieur le Professeur Y. CROSNIER pour les discussions que nous avons eues et les conseils avisés qu'il m'a prodigués.*

*Mes remerciements vont également à Monsieur M. DELBOS, Directeur du Laboratoire Téléopération et Robotique de l'E.D.F. Chatou et Monsieur MOSCHETTO, Directeur du CTB INSERM de Lille, pour l'intérêt qu'ils ont bien voulu porter à ce travail et pour leur participation à la commission d'examen.*

*Ma sympathie va à Monsieur R. LEGAI, Directeur du Service Braille du CRDP et à son équipe, qui par leur collaboration efficace, m'ont encouragé pour l'aboutissement de ce travail.*

*Qu'il me soit permis d'associer à ces remerciements Messieurs C. DRUON, F. WATTRELOT, P. ONANGA, V. DEVLAMINCK, F. WAUQUIER et J.M. POLLET pour la sympathie qu'ils m'ont témoignée.*

*Madame R. CASTEGNIER a eu la tâche délicate d'assurer la dactylographie de ce manuscrit. Je lui en suis reconnaissant et la remercie pour sa patience et son amabilité.*

*J'exprime enfin ma reconnaissance à tous les membres du laboratoire ainsi qu'à l'ensemble du personnel technique et administratif, notamment Monsieur J.P. DEHORTE qui a assuré l'impression de ce manuscrit.*

**-Squelettisation et détection de contour par des règles de réécriture -**

Dans une première étape, on procède à une analyse monodimensionnelle de l'image. Cette analyse effectue une segmentation de chaque ligne, guidée par une grammaire d'attributs. La grammaire dépend du modèle d'objet que l'on cherche à reconnaître: suivant le cas, on effectuera une squelettisation ou une détection de contour. La représentation à l'issue de cette segmentation est double:

- ✓ - Un segment, correspondant à l'épaisseur du contour ou de l'objet sur cette ligne,
- Un point, représentatif de l'ensemble des points de ce segment.

Les points représentatifs extraits sont chaînés d'une ligne sur l'autre pour constituer un graphe. Deux points seront chaînés si les segments les contenant sont connexes. La plupart du temps, ce chaînage est réalisé sans ambiguïté: le lien est bijectif. Cependant, il arrive qu'un segment soit connectable à plusieurs autres segments d'une ligne adjacente: on se trouve alors en présence d'un embranchement, il y a ambiguïté du chaînage et le lien n'est plus bijectif.

Le traitement des embranchements est résolu par la création d'un nouveau niveau de reconnaissance, caractérisant les points d'embranchement, considérés comme des primitives de reconnaissance. On ne résoud à ce niveau que les cas ambigus liés à un contexte restreint, la correction est effectuée par l'application de règles associées à chaque type d'ambiguïté détectée dans la configuration des primitives.

L'analyse monodimensionnelle décrite ci-dessus ne permet pas une détection satisfaisante des objets dont la direction est proche de la direction d'analyse, aussi procède-t-on à la même analyse, mais cette fois selon un balayage colonne par colonne.

Les résultats obtenus par ces deux opérateurs ne coïncident pas toujours, ils entrent même en conflit, en particulier pour les objets dont la direction est proche de la diagonale. Il est donc nécessaire de mettre en place un protocole de communication entre les graphes obtenus lors de chacune de ces deux analyses. Celui-ci s'appuie sur les points pour lesquels les deux analyses donnent des résultats cohérents: les points germes. Le graphe final résulte de la "fusion" des deux graphes et permet d'éliminer les redondances de représentation et les arcs parasites au moyen de règles de réécriture.

La segmentation n'utilise que de l'arithmétique entière, et la construction des graphes se fait par interclassement de listes ordonnées. Seule l'algorithmique de fusion des graphes dépend de la complexité de l'image, mais son coût reste cependant raisonnable, car un nombre très restreint de points est pris en considération à ce niveau.

Une dernière étape permet de situer les résultats de la méthode par rapport à ceux obtenus par des méthodes classiques de détection de contour.

S O M M A I R E  
-:-:-:-:-:-:-:-:-:-:-

	<u>N° page</u>
INTRODUCTION .....	1
CHAPITRE I : CLASSIFICATION DES CARACTÈRES BRAILLE EN FONCTION DE LA QUALITÉ D'EMBOSSAGE DU RELIEF COMPTE TENU DES CONDITIONS DE MESURES.	
I - INTRODUCTION .....	3
II - RAPPELS SUR LE PRINCIPE DE L'ECRITURE EN RELIEF .....	3
III - POSITION DU PROBLEME .....	6
IV - IDENTIFICATION DU SIGNAL LUMINANCE REPRESENTANT LE RELIEF BRAILLE .....	8
IV-1- Conditions de mesures.....	8
IV-2- Image numérique en relief Braille .....	10
IV-3- Etude statistique préliminaire.....	18
V - CONCLUSION .....	22
CHAPITRE II : PRÉSENTATION DE DIFFÉRENTES MÉTHODES POUVANT FOURNIR UNE SOLUTION THÉORIQUE AU PROBLEME DE DÉTECTION DU RELIEF BRAILLE.	
I - INTRODUCTION .....	23
II - NUMERISATION .....	23
III - DEFINITION DU VOISINAGE D'UN POINT .....	24
IV - TRAITEMENTS GLOBAUX .....	25
IV-1- Lissage d'images .....	25
IV-1-1- Définition de la moyenne.....	26
IV-1-2- Définition de la médiane.....	26
IV-2- Renforcement de contour.....	28
IV-3- Méthodes fréquentielles : Transformée de Fourier.....	29
IV-4- Méthodes récursives.....	31
IV-5- Squelettisation .....	32

	<u>N° page</u>
V - TRAITEMENTS LOCAUX .....	36
V-1- Suivi de contour .....	36
V-1-1- Suivi de bord .....	37
V-1-2- Suivi direct de frontière.....	38
V-2- Approximation d'une courbe par un polygone .....	38
V-2-1- Approximation d'un ensemble par une droite .....	39
V-2-2- Approximation polygonale d'une courbe.....	40
V-2-2-2-1- Approximation itérative .....	40
V-2-2-2-2- Approximation directe par un polygone.....	41
V-3- Codage de Freeman .....	42
VI - LES K FORMES, DEFINITION ET APPLICATIONS NOUVELLES EN TRAITEMENT D'IMAGES.	45
VI-1- Image : Tableau et graphe.....	45
VI-2- Codification des K-formes (maillage rectangulaire droit).....	46
VI-3- Application des K-formes pour la détection de contours.....	48
VII - EROSION ET DILATATION D'UNE IMAGE BINAIRE .....	48
VII-1- Introduction .....	48
VII-2- L'érosion .....	49
VII-3- La dilatation.....	50
VII-4- Lien entre érosion et dilatation.....	51
VII-5- L'érosion et la dilatation face à l'union et l'intersection .....	52
VII-5-1- L'érosion.....	52
VII-5-2- La dilatation.....	52
VII-6- Ouverture et fermeture .....	52
VIII - CONCLUSION .....	54

### CHAPITRE III : DÉTECTION DU RELIEF BRAILLE

I - INTRODUCTION .....	55
------------------------	----

#### PREMIÈRE PARTIE : PREMIÈRE MÉTHODE DE DÉTECTION DU RELIEF

II - LE PRETRAITEMENT DE L'IMAGE EN RELIEF BRAILLE .....	57
III - SEGMENTATION D'IMAGES .....	59
IV - EXTRACTION D'ATTRIBUTS DU RELIEF BRAILLE .....	61
IV-1- Renforcement linéaire des contours du relief Braille .....	61
IV-1-1- Renforcement linéaire des contours dans les directions orthogonales ...	62
IV-1-2- Renforcement linéaire des contours par les compas directionnels .....	63

.../...

	<u>N° page</u>
V - LA BINARISATION .....	65
V-1- Le seuillage .....	65
V-2- Le filtrage binaire .....	67
V-3- Le lissage de l'image binaire .....	67
VI - DISCUSSION .....	72
 DEUXIÈME PARTIE : APPROXIMATION DE LA LUMINANCE DU RELIEF BRAILLE PAR UN MODÈLE POLYNOMIAL À L'AIDE DE RÉGRESSION LINÉAIRE,  	
VII - DETECTION ET MISE EN VALEUR DU RELIEF BRAILLE EN APPRIMANT LA LUMINANCE PAR UN MODELE POLYNOMIAL .....	74
VII-1- Lissage du relief Braille par approximation polynomiale .....	74
VII-1-1- Détermination des masques de lissage du relief Braille par la méthode des moindres carrés .....	74
VII-2- Détection du relief Braille .....	81
VII-3- Partie expérimentale .....	82
VII-4- Robustesse des masques fixes face au bruit.....	102
VIII - CONCLUSION .....	103
 CHAPITRE IV : PARTICULARITÉS DE L'ARCHITECTURE DU NOUVEAU TRANS- CRIPTEUR BRAILLE/NOIR ,  	
I - INTRODUCTION .....	105
II - ETUDE DE LA FONCTION FENETRAGE ET SYNCHRONISATION.....	106
III - ETUDE ET REALISATION DE LA MEMOIRE IMAGE .....	107
III-1- Analyse des particularités de la mémoire image .....	112
III-1-1- Réalisation du signal chargement .....	112
III-1-2- Problèmes liés au temps d'accès des mémoires par rapport à la fréquence d'arrivée des pixels .....	112
III-1-3- Transformation série/parallèle des pixels.....	113
III-1-4- Etude et réalisation de la visualisation de la mémoire image.....	117
IV - CONCLUSION .....	120
CONCLUSION .....	121
BIBLIOGRAPHIE .....	123

## I N T R O D U C T I O N

---

Lorsque le traitement numérique des images s'est développé, au début des années 60, l'amélioration et la restauration d'images recueillirent une grande attention. Deux raisons distinctes expliquent ce phénomène. D'une part, la généralisation au cas des images de techniques bien connues pour le traitement des signaux temporels devenait possible grâce au développement des moyens informatiques. D'autre part, les résultats obtenus étaient spectaculaires y compris pour le public non initié. L'oeil humain est en effet souvent le juge final de l'efficacité de tels traitements.

Au fil du temps, d'autres problèmes placés en aval ont été abordés et ont, peu à peu, attiré une grande attention dans le domaine de la recherche. L'extraction des primitives s'est alors développée d'une façon considérable. De nombreux algorithmes de détection de contours ont été proposés et beaucoup d'efforts ont été consacrés à l'analyse et à la synthèse de textures. En utilisant ces paramètres extraits des images, la classification et la segmentation sont devenues des disciplines de base du traitement d'images. En prenant en compte enfin des modèles d'images, l'intelligence artificielle a trouvé dans les données d'images, matière à de nombreux développements. C'est ainsi qu'est née l'analyse de scène.

L'image numérique d'un relief Braille produite par une caméra peut être considérée comme une fonction bi-dimensionnelle représentant la luminance de points élémentaires. La reconnaissance du relief Braille en vue de sa traduction fait appel aux méthodes et techniques de filtrage numérique d'images.

.../...

Cependant, le thème de cette thèse se rapporte à la définition d'une méthode de filtrage du relief Braille dont la mise en oeuvre numérique soit aisée pour permettre sa reconnaissance dans le but de réaliser un nouveau système de transcription du texte Braille en texte noir. Ce procédé, appelé LECTOBRAILLE, trouve des applications dans le domaine de l'éducation (correction d'épreuves d'examen par des enseignants voyants) et dans la restitution d'ouvrages Braille anciens et altérés.

Dans ce travail, nous nous intéressons aux prétraitements de l'image en relief Braille, à l'extraction de ses attributs, à sa segmentation, à la détection du relief des points Braille et à la description de l'architecture du nouveau système de transcription Braille-noir.

Quatre axes d'études sont abordés :

Dans le premier chapitre, nous rappelons la position du problème ainsi que les conditions de mesures. Nous identifions le signal luminance représentant le relief Braille et nous essayons de mettre en évidence les différents problèmes posés en vue de sa reconnaissance.

Nous présentons, dans le deuxième chapitre, quelques méthodes numériques de traitement d'images qui peuvent fournir une solution théorique au problème de détection du relief Braille.

Dans la troisième partie, nous nous intéressons à la reconnaissance du relief Braille. Nous mettons en oeuvre ainsi une méthode de filtrage numérique qui s'avère efficace pour la détection du relief.

Enfin, dans un quatrième chapitre, nous décrivons les particularités de l'architecture du système de transcription Braille/noir.

CHAPITRE I : CLASSIFICATION DES CARACTÈRES BRAILLE EN FONCTION  
DE LA QUALITÉ D'EMBOSSAGE DU RELIEF COMPTE TENU  
DES CONDITIONS DE MESURES.

## I - INTRODUCTION

Après avoir brièvement rappelé le principe de l'écriture en relief, la position du problème, ainsi que les conditions de mesures, nous identifions le signal luminance représentant le relief Braille et nous essayons de mettre en évidence les différents problèmes posés en vue de sa reconnaissance. Dans ce but, nous avons effectué une étude statistique préliminaire permettant une classification des caractères Braille en fonction du contraste de l'objet (point Braille) et les défauts d'alignement.

Nous terminons ce chapitre par l'exposé des différents types de classes retenus et par une discussion sur la facilité ou la difficulté d'extraction d'attributs permettant la reconnaissance du relief Braille.

## II - RAPPELS SUR LE PRINCIPE DE L'ECRITURE EN RELIEF

Nous rappelons ici le principe de l'écriture en relief pour permettre de définir le type de signal que nous avons à traiter.

Le moyen le plus ancien qui a permis aux déficients visuels l'accès à la connaissance est la lecture du texte par un voyant<sup>[1][2]</sup>. On imagine sans mal les contraintes d'une telle pratique : disponibilité du lecteur, interprétation du texte etc. de sorte qu'historiquement il a été complété par la communication écrite directe en relief. La chaîne de communication écrite directe est composée d'un organe source, soit manuel (membre preneur comme la main, le pied, la bouche par exemple), soit mécanique (machine à embosser), et un instrument graveur (poinçon), d'un canal support de l'information matérialisé à l'origine et encore souvent aujourd'hui par du papier relief et d'un récepteur constitué par un organe tactile (peau du doigt ou de la main en général).

Le système universellement adopté est celui proposé par Louis BRAILLE. En 1925, ce système a eu quelques difficultés au début pour être adopté, mais sa logique et la facilité de la reconnaissance de la forme l'a rendu universel. Il a donné aux déficients visuels la possibilité de lire des textes à une vitesse à peu près égale à celle des voyants.

Un caractère Braille est constitué d'une matrice à six points pouvant être mis en relief (figure I-1-a). La fabrication de ce relief peut se faire soit en déposant une matière sur un support papier plan (thermogravure), soit en poinçonnant un creux au verso du support papier (poinçonnement ou embossage), soit en déformant une feuille de matière thermodurcissable (thermoformage). Dans le cas de fabrication manuelle par poinçonnement par les aveugles pour leur communication écrite individuelle, l'écriture d'un texte relief en Braille se fait dans le sens inverse de la lecture du même texte. C'est-à-dire, que l'on écrit en creux de droite à gauche sur une face et on lit le relief de gauche à droite sur l'autre face. Mais dans ces conditions, le caractère Braille apparaît lui-même inversé en écriture et en lecture. Il en résulte une convention de repérage numérique des points d'un caractère qui se composent dans le sens écriture de la manière représentée figure I-1-a.

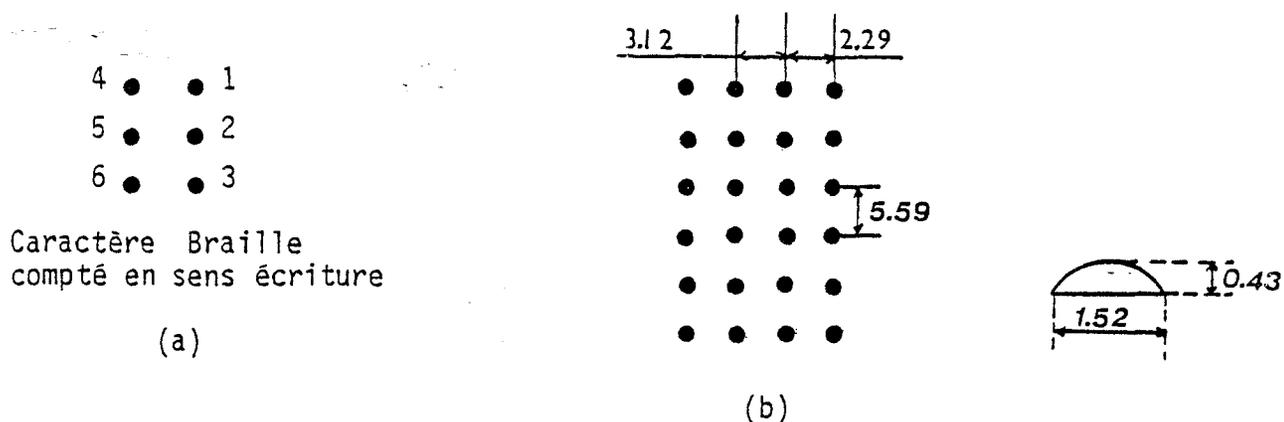


Figure I-1

Le nombre de caractères possibles pouvant être générés par ce système est donc de  $2^6$ , soit 64 caractères différents (espaces compris).

La matrice à 6 points est optimale car avec un point en moins le nombre de caractères générés par le système serait insuffisant (32) et avec un point en plus, les dimensions d'un caractère dépasseraient la zone la plus sensible du doigt : la pulpe.

Ce système, actuellement le plus adapté à la psychophysiologie des excitations tactiles qui sont du type dynamique avec une constante de temps  $< 1$  s, est du genre tout ou rien et ne demande qu'une faible quantité d'informations à appréhender au toucher.

En ce qui concerne les dimensions et les caractéristiques physiques d'un caractère Braille, il n'y a pas vraiment de normalisation entre les différents pays du monde<sup>[1]</sup>, cependant, une étude expérimentale menée par Meyers, Ethington et Ashcroft a montré que les dimensions données sur la figure I-1-b sont les meilleures pour la lecture.

Les 64 caractères de l'alphabet Braille peuvent représenter seuls les lettres alphabétiques, accentuées ou non, et les caractères de ponctuation. C'est le cas du Braille intégral<sup>[3][4]</sup> ou Braille I. Ils peuvent représenter aussi des groupements de lettres conditionnellement à leurs positions dans un mot (en terminaison, en début, devant des voyelles, devant des consonnes), soit primitif, soit d'une zone de préfixes, d'une racine, d'un suffixe au pluriel ou féminin pluriel. C'est le cas du Braille abrégé ou Braille II.

Dans le cadre de l'étude d'un système de transcription automatique de relief Braille en noir, le problème consiste à acquérir un signal mémorisé présentant l'amplitude de chaque point élémentaire du relief, de le prétraiter, d'extraire ses attributs, de le segmenter, de reconnaître le relief des points Braille et les graphèmes, de coder chaque graphème, de transcrire les chaînes codées obtenues en noir et visualiser ou imprimer le texte noir obtenu.

Dans le cadre de sa thèse de 3<sup>e</sup> cycle, Monsieur MORTREUX a étudié et mis en oeuvre une méthode d'élaboration des algorithmes de transcription Braille-noir en vue de leur implantation<sup>[3]</sup> sous forme logicielle adaptée aussi au nouveau transcrip-teur Braille-noir qui est en cours de réalisation.

Dans notre thèse, nous nous intéressons aux prétraitements de l'image en relief Braille, à l'extraction de ses attributs, à sa segmentation, à la reconnaissance du relief des points Braille et à la réalisation matérielle d'un nouveau système de transcription Braille-noir.

### III - POSITION DU PROBLEME

Des travaux ont été menés depuis plusieurs années au laboratoire de Mesures Automatiques sur la communication aveugle-voyant. Ils ont conduit à la réalisation d'un lecteur optique de relief Braille (figure I-2) avec transcription en texte noir<sup>[3][4][5]</sup>.

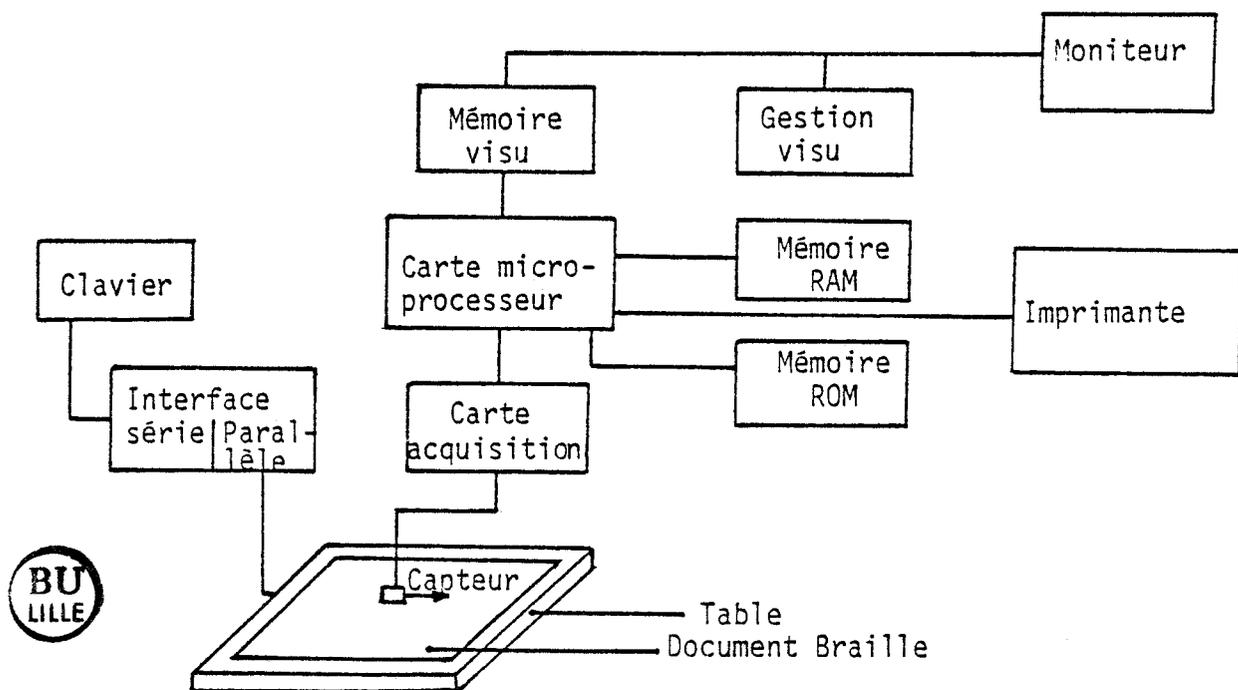


Figure I-2 : Synoptique matériel du transcrip-teur Braille/noir

Ce système est muni d'un éditeur de texte et d'un noyau logiciel capable de restituer après acquisition, filtrage, reconnaissance et transcription Braille-noir, les textes noir et Braille, soit sur un support visuel, soit sur un support courant.

Ce procédé, appelé Lectobracille, trouve des applications dans le domaine de l'éducation (correction d'épreuves d'examen par des enseignants voyants) et dans le domaine de la restitution d'ouvrages Braille anciens et altérés.

Ce système de reconnaissance Braille, utilisant un capteur optique ponctuel monté à la place de la plume d'une table traçante a comme inconvénients d'avoir un balayage mécanique long et d'utiliser une table traçante précise de coût élevé.

Pour ces raisons et l'évolution rapide de la technologie, il a été envisagé de remplacer la table traçante par un capteur optique matriciel assimilable à une rétine.

Dans le but de réaliser un nouveau transcritteur Braille/noir s'adaptant aux caractéristiques du capteur CCD, un système a été étudié (figure I-3). La partie matérielle réalisée de ce système est décrite au chapitre IV.

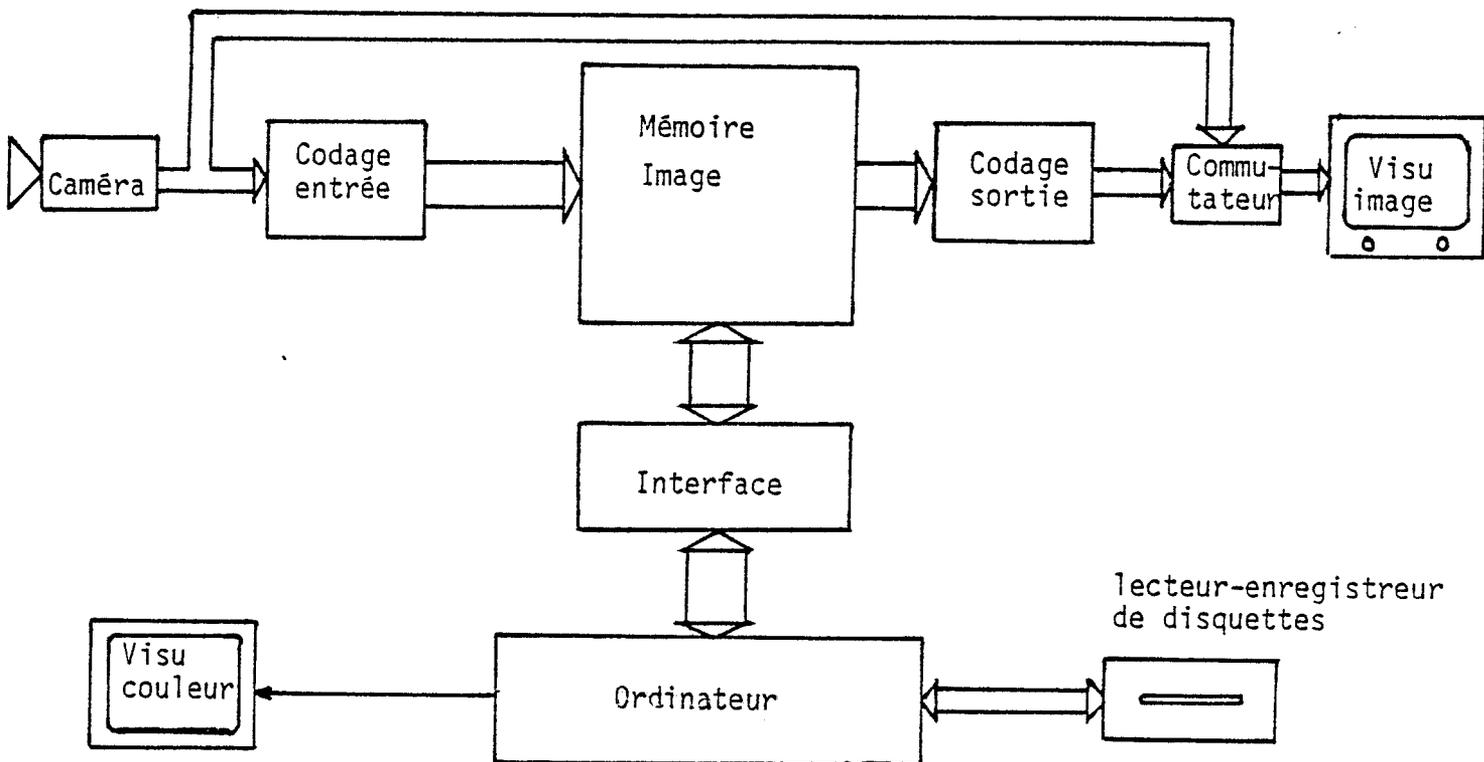
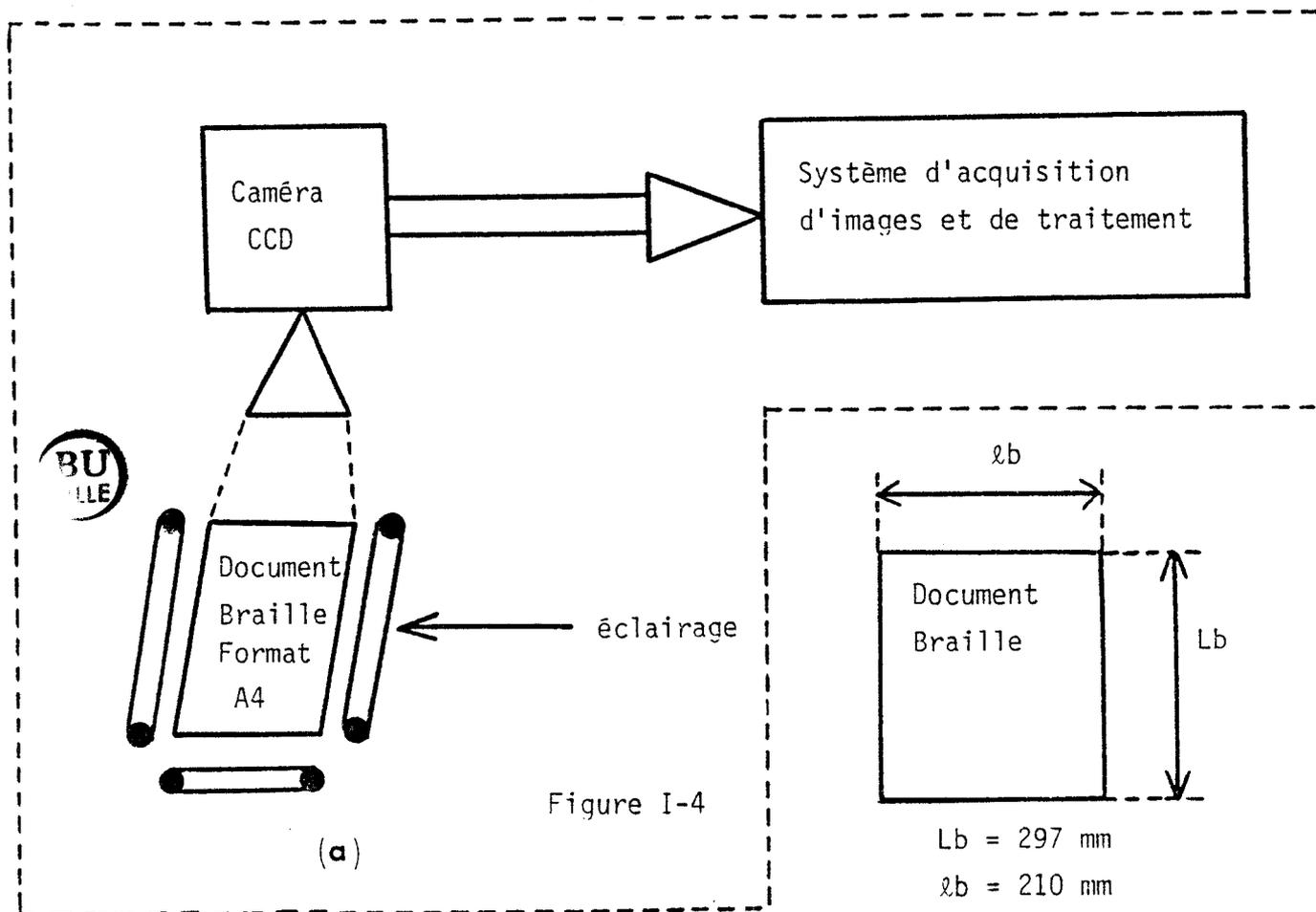


Figure I-3- Synoptique matériel du nouveau transcritteur Braille/noir.

#### IV - IDENTIFICATION DU SIGNAL LUMINANCE REPRESENTANT LE RELIEF BRAILLE

##### IV-1- CONDITIONS DE MESURES

Dans notre application, le sujet est constitué par des documents en relief Braille soumis à un éclairage rasant comme le montre la figure I-4-a.



La caméra à cellule CCD observe le document Braille sur une surface rectangulaire ayant un rapport entre la longueur et la largeur égal à  $\frac{L_c}{\ell_c} \approx 1,33$  (figure I-5-a).

Le capteur CCD matriciel restitue un signal analogique échantillonné (image vidéo) de 576 lignes chacune comportant 384 points élémentaires (pel ou pixel) (dimensions d'un pel figure I-5-b).

Si on observe le sujet en faisant correspondre les longueurs et les largeurs du document et la cellule CCD, un point Braille est représenté par  $2,74 \times 1,3$  points élémentaires, ce qui ne vérifie pas le théorème de Schannon. Nous avons donc choisi d'observer le document Braille en faisant correspondre sa largeur avec la longueur de la cellule. Un point Braille est alors observé par  $1,82 \times 3,67$  pixels suivant la largeur et la longueur du document.

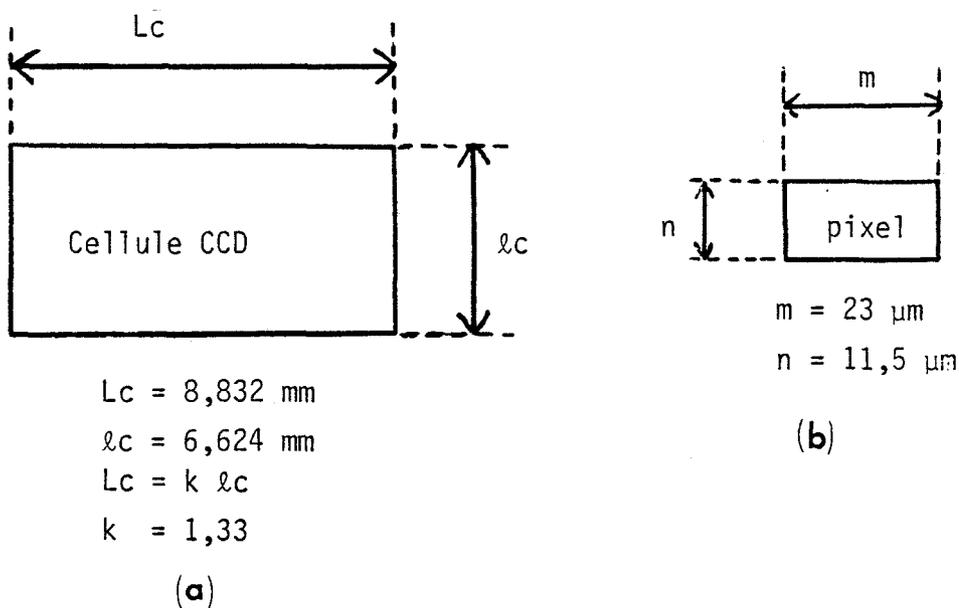


Figure I-5

Dans ces conditions de mesures, la page Braille sera analysée en deux fois correspondant à des surfaces qui se chevauchent sur une bande de largeur supérieure au pas des lignes Braille. La figure I-6 représente les surfaces observées par la caméra sur le document Braille.

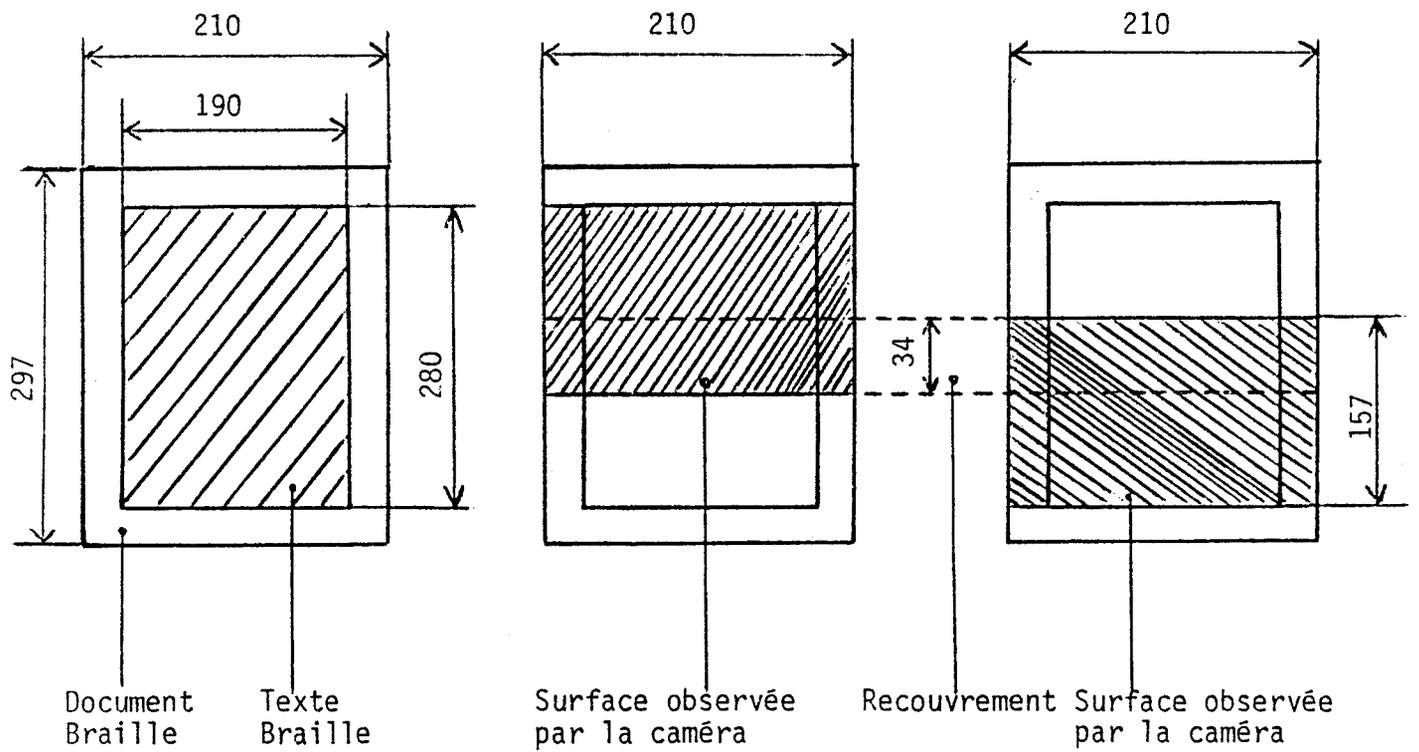


Figure I-6

Cependant, le calcul du nombre de pixels par point Braille précédent montre que l'on dépasse encore la limite de Shannon. L'éclairage rasant du relief provoque un effet d'élargissement des dimensions apparentes des points en relief vues par le capteur matriciel, ce qui permet de restituer les conditions de Shannon dans la pratique.

#### IV-2- IMAGE NUMERIQUE EN RELIEF BRAILLE

Par une conversion analogique digitale, l'image en relief Braille est numérisée (la numérisation d'images est décrite au chapitre II). Elle peut être considérée comme une fonction bi-dimensionnelle  $f(x,y)$  représentant l'intensité lumineuse (la grandeur physique est la luminance) de points élémentaires (pels, pixels, échantillons). Chacun d'eux est une valeur entière, positive, quantifiée et codée sur 6 bits (64 niveaux de gris) dont les valeurs 0 et 63 correspondent respectivement au noir et au blanc.

Cet échantillon de luminance de l'image discrète est déterminé par la conjugaison de plusieurs facteurs physiques au point correspondant de la scène, incluant notamment les propriétés photométriques de la source lumineuse incidente, les caractéristiques de transmission ou de réflexion de la surface et ses orientations relatives du capteur ou de la source. On peut y ajouter également l'influence des paramètres géométriques et photométriques de la cellule utilisée.

L'image est donc le mélange de toutes ces informations, lesquelles sont reflétées par une valeur unique attribuée au point.

Nous donnons comme exemple d'acquisition d'images en relief, un graphème Braille se composant d'un seul point (figure I-7) (le document est de couleur blanche et le procédé de fabrication est du type "embossage mécanique").

$x \backslash y$	-3	-2	-1	0	1	2	3
-3	20	20	20	20	21	20	20
-2	20	20	19	22	22	21	19
-1	20	20	20	20	28	22	18
0	20	19	19	21	25	20	18
1	20	20	19	19	23	22	18
2	20	20	19	19	23	19	19
3	20	20	20	19	19	19	19

$(x,y)$  : coordonnées de l'amplitude de la luminance

$f(x,y)$  : l'intensité lumineuse en niveaux de gris

Par exemple, pour  $x = 0$  et  $y = 0$ , on a  $f(x,y) = 21$ .

Figure I-7 : Exemple d'un relevé mémorisé (7x7) de l'intensité lumineuse d'un relief représentant un point Braille.

Une image "mieux interprétable" est souvent une image améliorée en ce sens qu'elle sera débarrassée des défauts, des bruits qui en masquent l'interprétation. Les défauts peuvent être dus aux conditions de la prise d'image ou à un phénomène propre à l'objet mesuré. Flous à enlever, améliorations de contrastes locaux sont des améliorations fréquentes.

Par une image "mieux interprétable" il faut entendre également une image où l'information cherchée est mise en valeur.

L'extraction d'une information qualitative ou quantitative d'une image est la seconde étape du traitement d'image. Extraire une information qualitative rapproche de l'idée de "mieux interpréter" une image.

Elle en diffère en cela qu'elle propose une transformation de l'image qui peut mener à un résultat très différent de l'image originale.

Un exemple simple est notre application. En cas de présence de l'objet "point Braille élément d'un graphème", seul celui-ci est conservé dans l'image. En cas d'absence, l'image n'a plus de raison d'être. Ce type d'information qualitative est illustré par la reconnaissance de graphème "caractère Braille" en vue de sa traduction.

Le premier calcul effectué sur une image binaire est l'histogramme de ses niveaux de gris. En abscisse, les niveaux de gris (0 à 63), en ordonnée, le nombre de points de l'image ayant le niveau de gris correspondant.

On donne comme exemple, l'histogramme d'une image en relief Braille de type "embossage mécanique" représenté par la figure I-8.

L'histogramme obtenu n'est pas bimodal (2 pics). Ce qui ne traduit pas la présence de deux objets de l'image : le point Braille et le fond.

Les conditions d'éclairage, de prise de photo, de nature de l'objet étudié etc. conduisent le plus souvent à des histogrammes dont les modes sont plus rapprochés et plus évasés à la base comme dans notre cas.

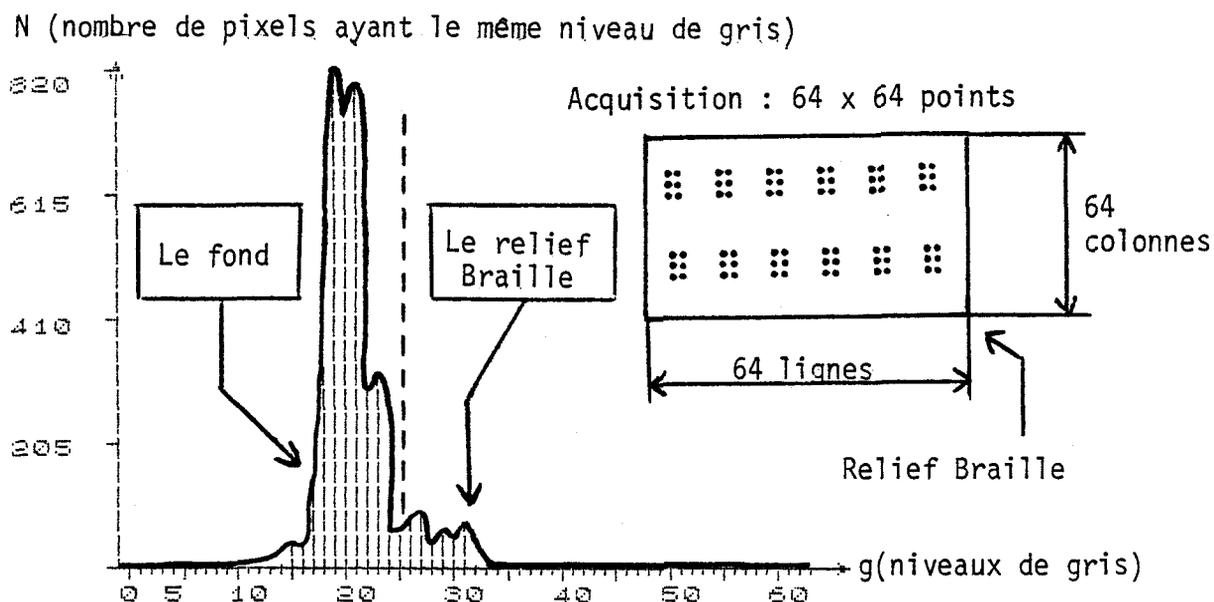


Figure I-8 : Histogramme des niveaux de gris

Le choix du seuil ne pourra pas alors se faire de manière interactive (suites d'essais et arbitrage de celui qui analyse). La séparation de l'objet du fond et l'élimination du bruit aléatoire pour l'extraction d'attributs ne peut se faire qu'en traitant l'image. Cependant, pour donner une idée générale des problèmes à résoudre, nous décrivons quelques exemples d'acquisition d'images numériques en relief dont le procédé de fabrication est du type "embossage mécanique ou manuel" et de couleur blanche.

Ces images discrètes se composent de 11 lignes chacune comportant 17 points élémentaires. A chaque point de coordonnées  $(x,y)$ , on associe une valeur de l'intensité lumineuse  $f(x,y)$  en niveau de gris.

Les échantillons sont positionnés selon les singularités du graphisme (angles vifs, extrémités de segments) et selon les propriétés spatiales du capteur CCD.

L'examen des figures représentant quelques échantillons de relief nous permet de faire les remarques suivantes :

- la présence de l'objet s'identifie par une augmentation du contraste d'un certain nombre de pixels dont la taille est en fonction des dimensions du relief ;
- Un point Braille embossé correctement de type "embossage mécanique ou manuel" présentera une surface plus grande et une amplitude plus forte qu'un point dont le

relief est mal embossé comme le montre les figures I-9, I-10, I-11 et I-12 ( $\Delta g$  est la différence des niveaux de digitalisation de l'objet et du fond) ;

- Les objets sont de petite taille (figure I-9) et faiblement contrastés (6 à 10 niveaux de digitalisation) ou fins (figure I-12) et très faiblement contrastés (3 à 5 niveaux de digitalisation). Mais il existe aussi des objets dont le relief est fort écrasé qui s'identifie comme un défaut du papier ou un bruit aléatoire.

La figure I-14 représente un relief écrasé d'un document embossé manuellement par un non-voyant (elle peut être une copie d'examen du baccalauréat). On peut constater que l'objet et le bruit se confondent ;

- Les objets sont proches. Ce qui peut se traduire par des interférences de luminosité. La périodicité du relief et la résolution spatiale du capteur sont fixes. Pour éviter le problème des interférences, le seul moyen est le réglage optimal de la luminance qui dépend de la quantification du signal et les conditions d'éclairage.

La figure I-10 montre que deux points voisins appartenant au même graphème peuvent avoir une influence mutuelle l'un sur l'autre, mais seulement horizontalement ("a" représente la distance du voisinage de 2 points Braille appartenant au même graphème) ;

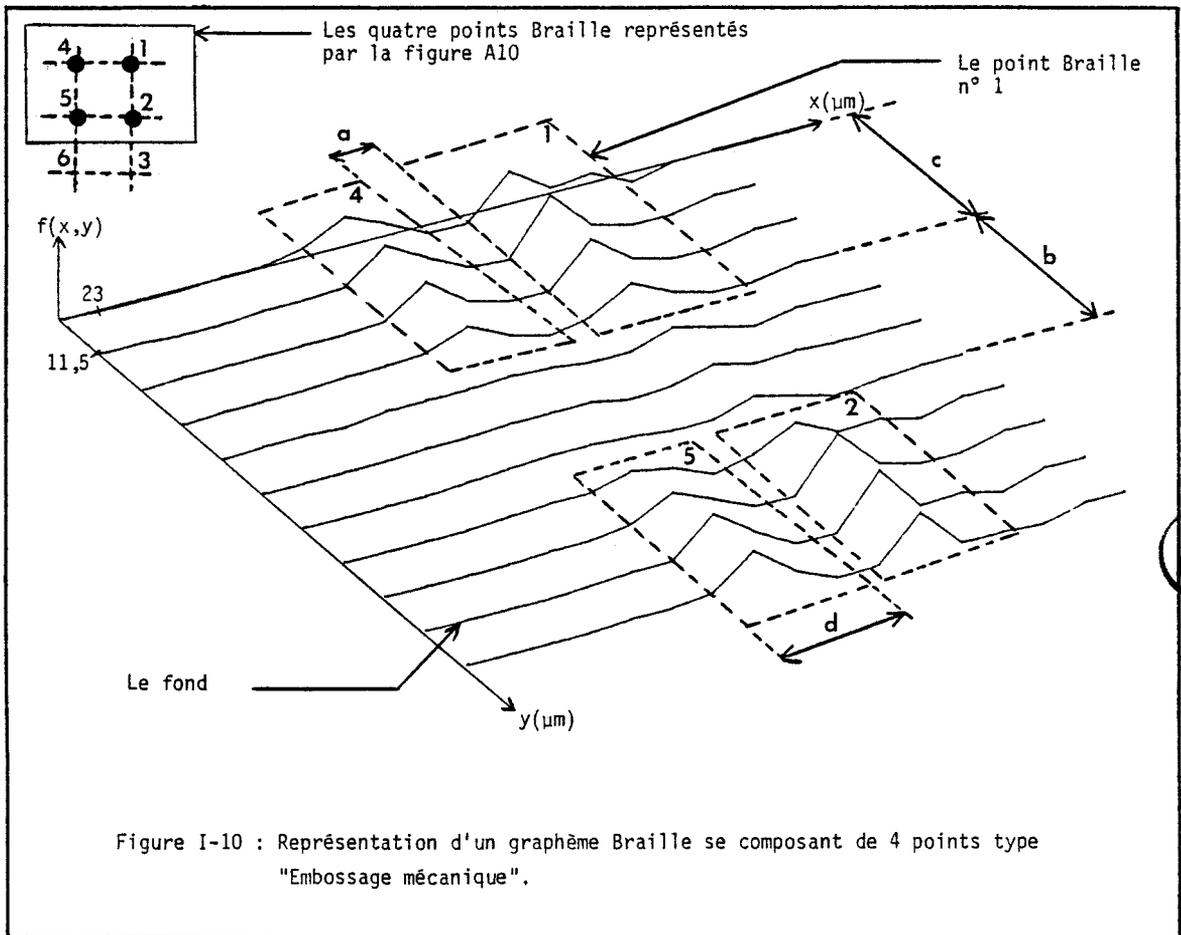
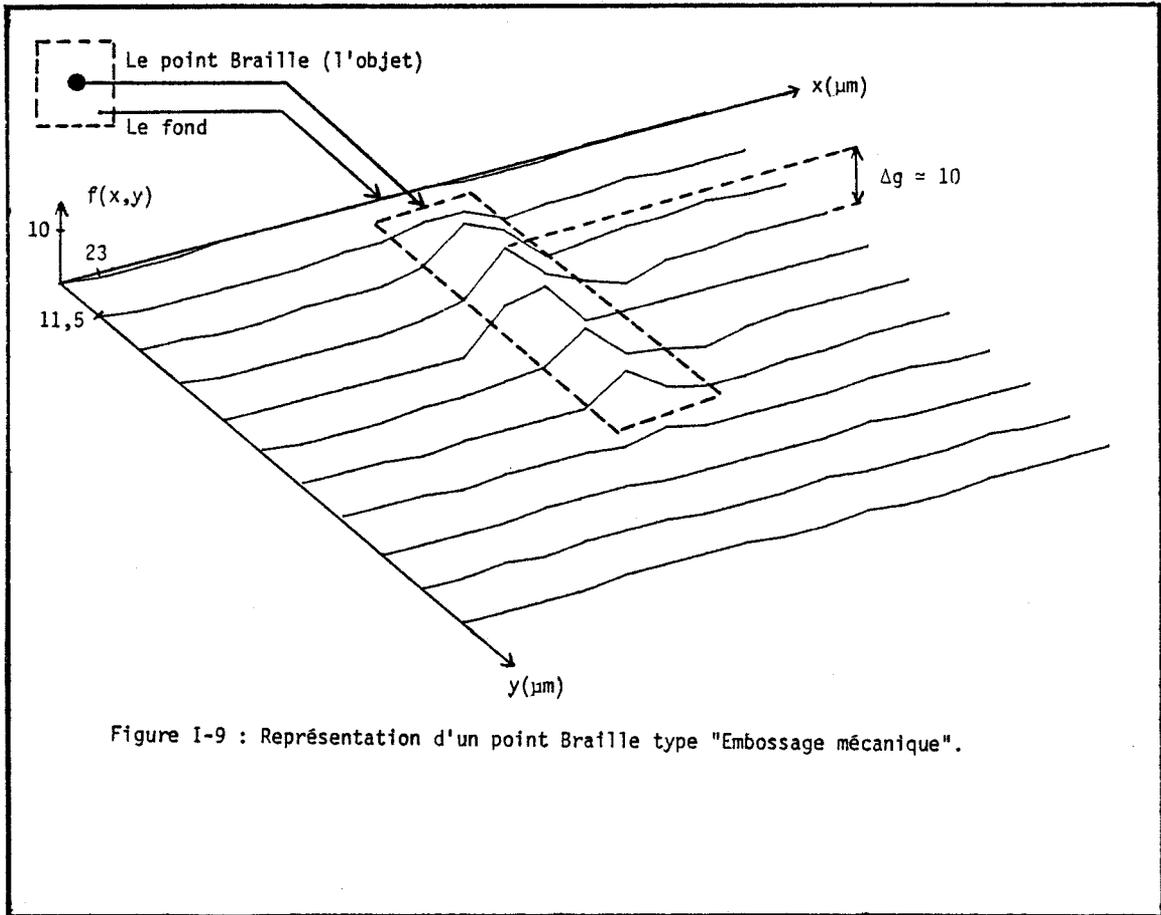
- Des objets de formes complexes.

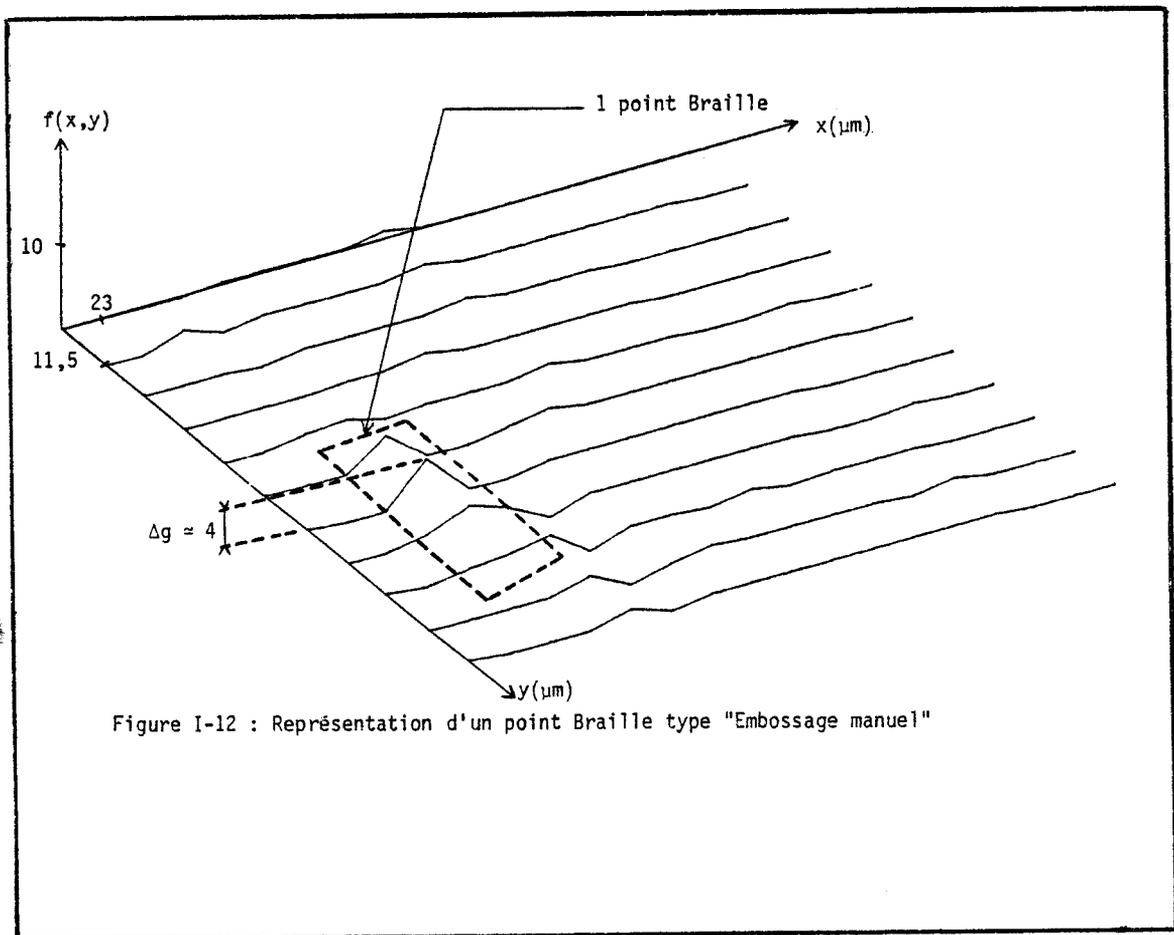
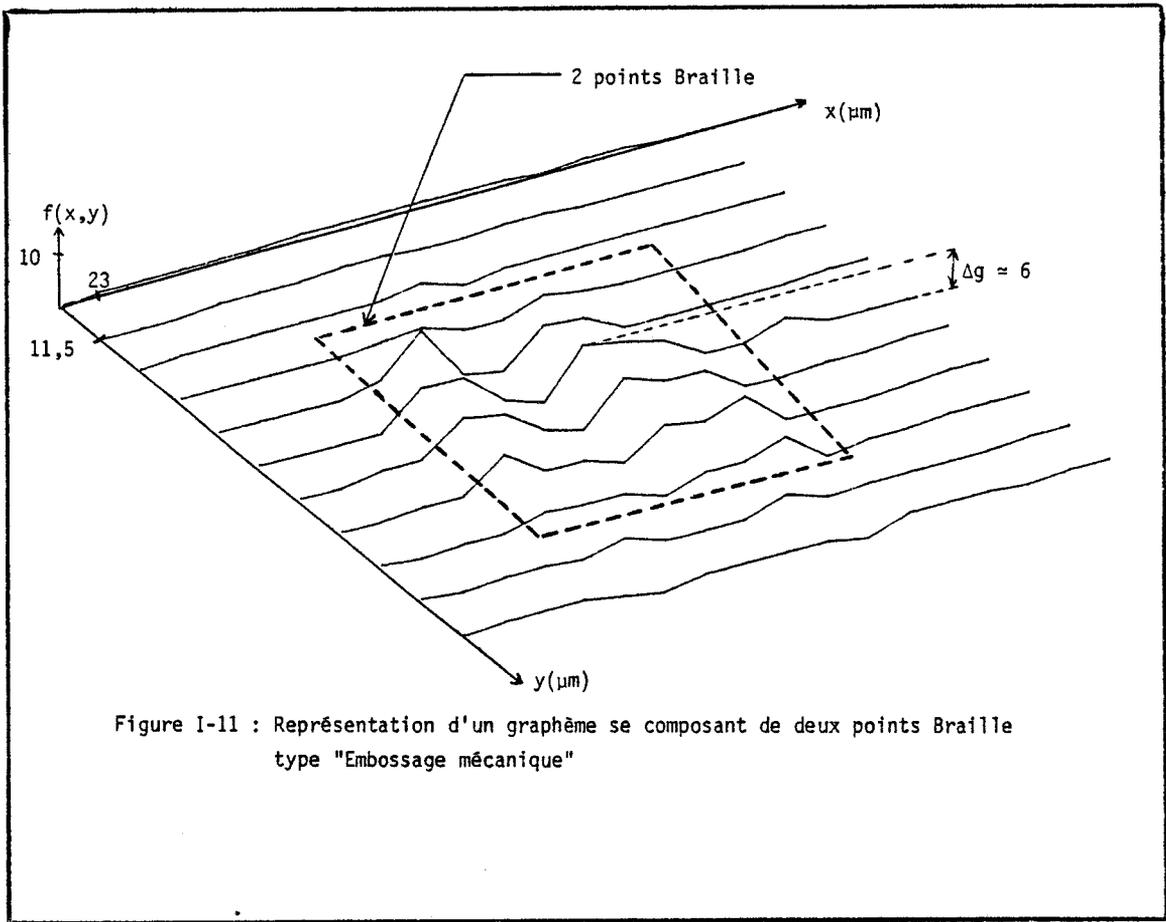
La figure I-13 est un exemple simple présentant un point Braille dont la variation de luminosité est importante suivant les lignes.

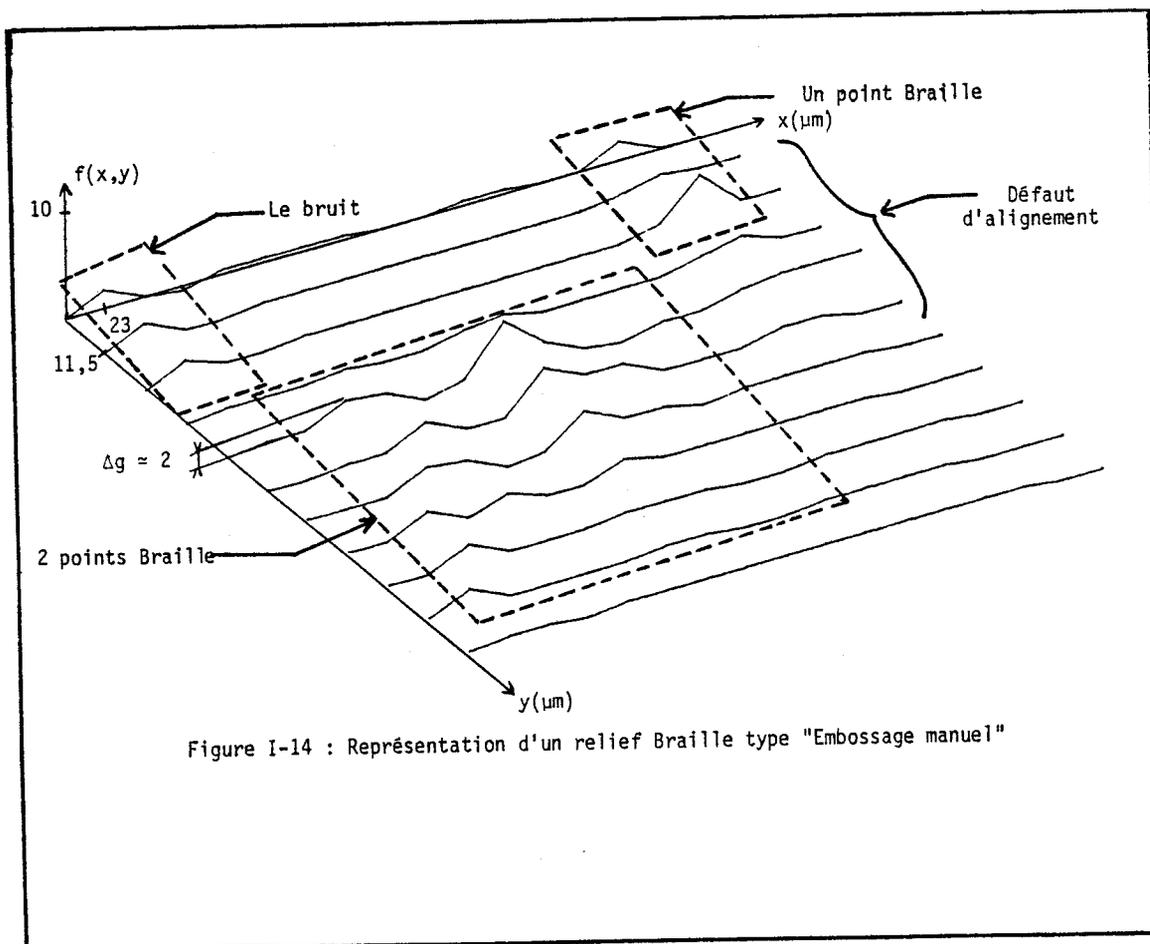
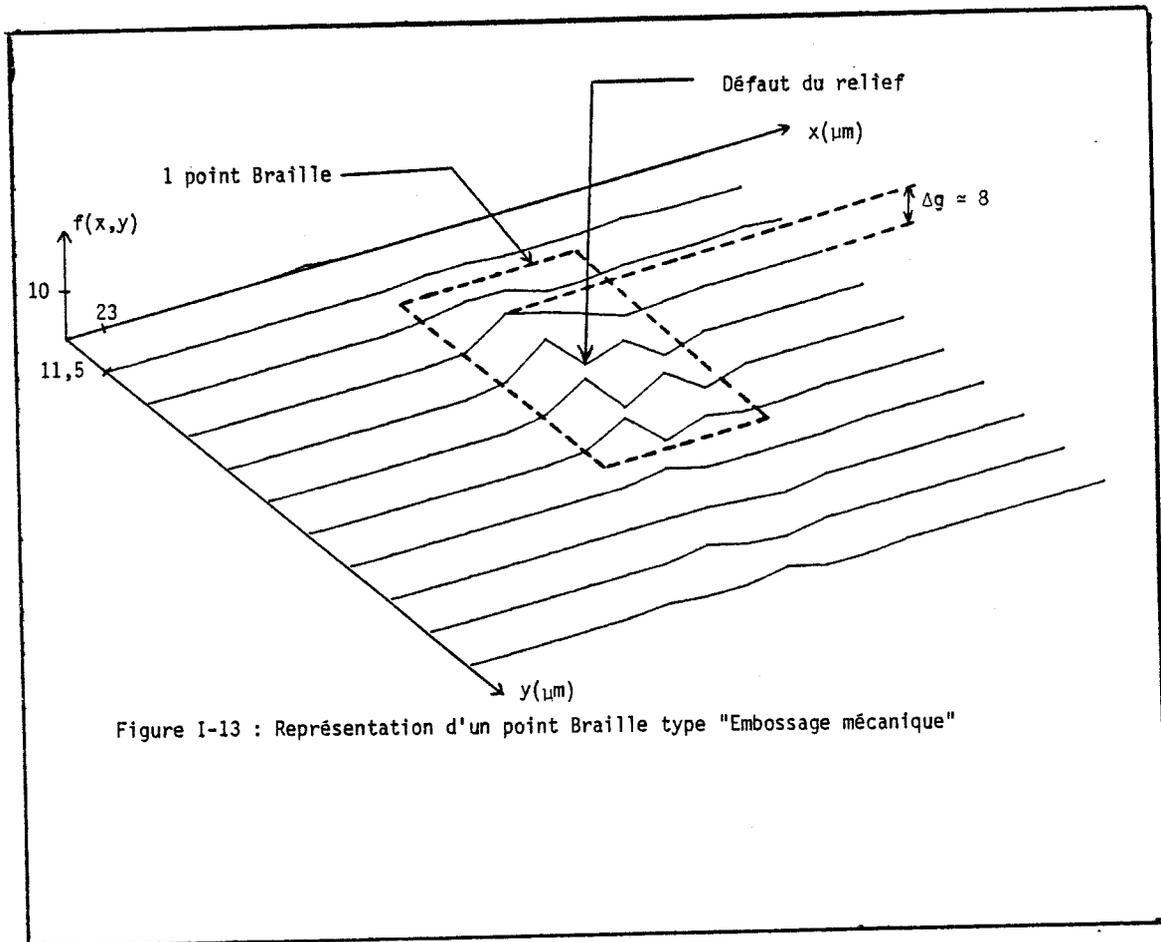
Ces discontinuités peuvent être dues à un défaut du papier, un embossage incorrect ou autre ;

- Le défaut d'alignement. Un point Braille n'occupe pas la position géométrique correspondant à sa place dans la matrice 6 points définissant le caractère Braille (figure I-14).

Ce défaut apparaît surtout dans le cas d'embossage manuel des documents. Le même point Braille peut être aussi embossé incorrectement ou écrasé.







#### IV-3- ETUDE STATISTIQUE PRELIMINAIRE

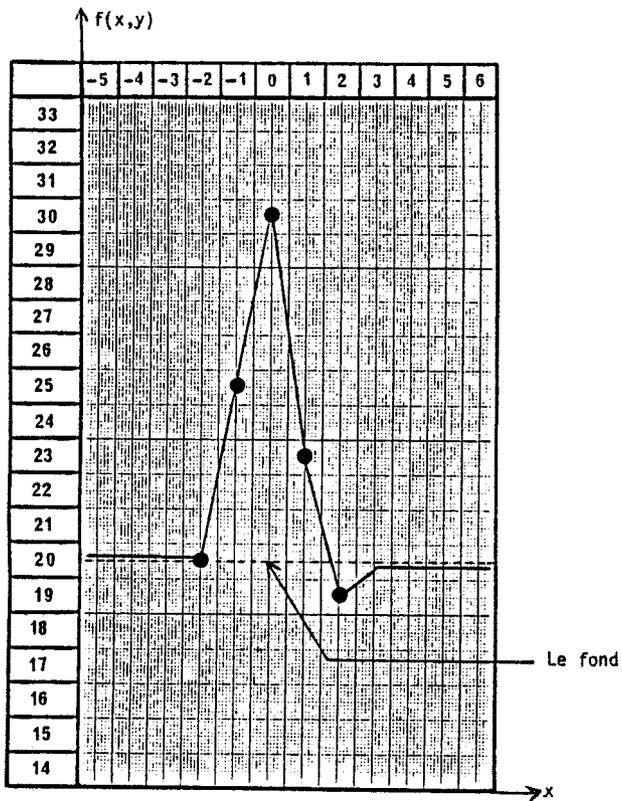
Les conditions citées au paragraphe précédent nous permettent de dénombrer les premières difficultés posées afin d'analyser une image en relief Braille. Cependant, le nombre important de données à traiter, de l'ordre de 500.000 points environ par page Braille, nous a amené à effectuer une étude statistique préliminaire permettant une classification des caractères Braille en fonction du contraste de l'objet et des défauts d'alignement.

Cette étude statistique a été menée dans les conditions suivantes :

- l'échantillon élément d'une population appartenant à une classe bien définie est le caractère Braille (graphème) ;
- Un caractère est élément de la classe A, si tous les points de celui-ci sont faiblement contrastés (6 à 10 niveaux de gris. Exemple : figure I-15-a et I-15-d) ;
- Un caractère est élément de la classe B si un point au moins de celui-ci est très faiblement contrasté (3 à 5 niveaux de gris. Exemple : figure I-15-b, I-15-c et I-15-e). Les autres points étant de classe A ;
- Un caractère est élément de la classe C, si un point au moins de celui-ci n'occupe pas sa place dans la matrice 6 points. Il n'est pas aligné avec ses voisins (Exemple : figure I-14). Les autres points appartiennent à la classe A ou/et B ;
- Un caractère est élément de la classe D, si un point au moins de celui-ci est écrasé (Exemple : figure I-15g) ou de forme complexe (Exemple simple : figure I-15f). Les autres points appartiennent à la classe A ou/et B ou/et C.

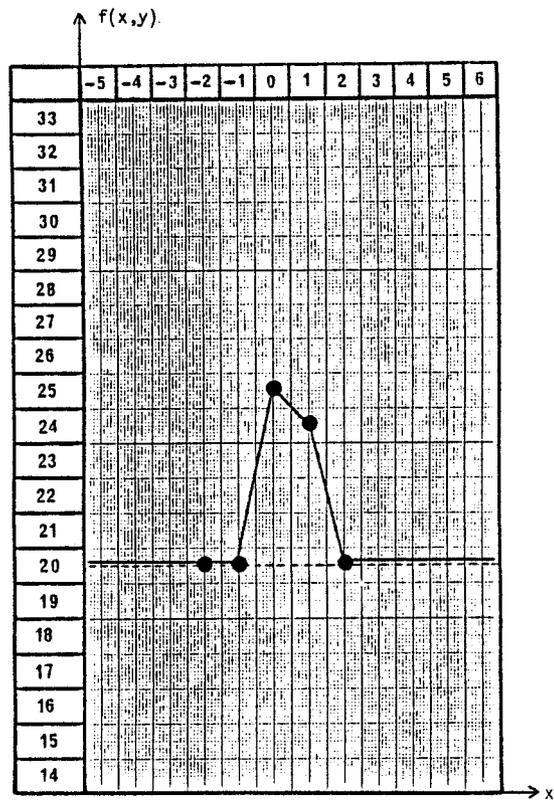
Cette étude sommaire a été effectuée dans nos conditions de mesures pour deux procédés de fabrication qui sont le relief du type "embossage mécanique" et le relief du type "embossage manuel" (les documents sont de couleur blanche). Le but est aussi de pouvoir estimer après traitement un taux approximatif de réussite d'extraction d'attributs du relief Braille.

Le nombre d'échantillons représentatifs et les fréquences d'apparition de chaque classe sont donnés par le tableau 1.



(a)

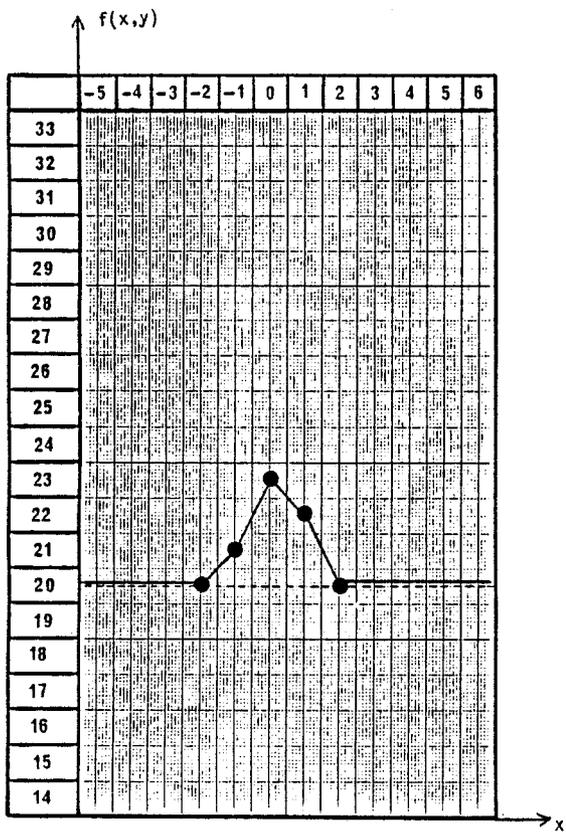
Variation de la luminosité suivant une ligne d'un point Braille "classe A"



(b)

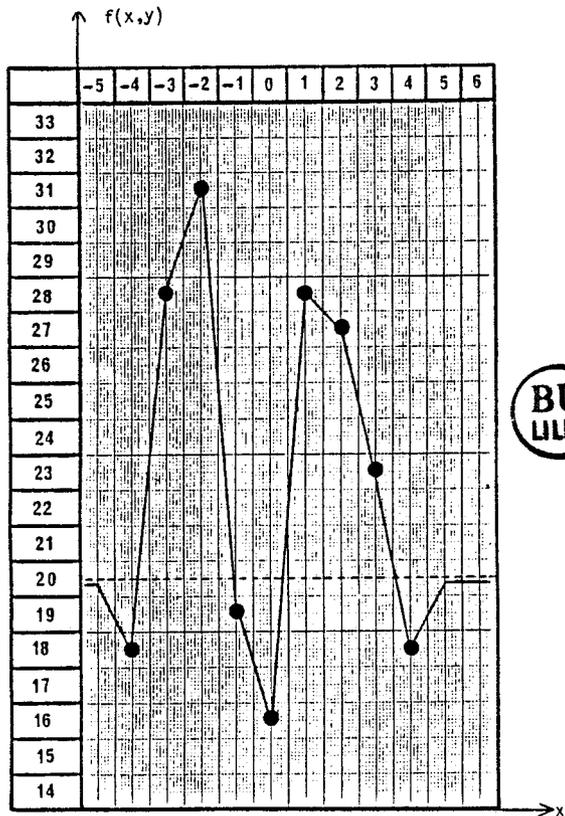
Variation de la luminosité suivant une ligne d'un point Braille "classe B"

Figure I-15



(c)

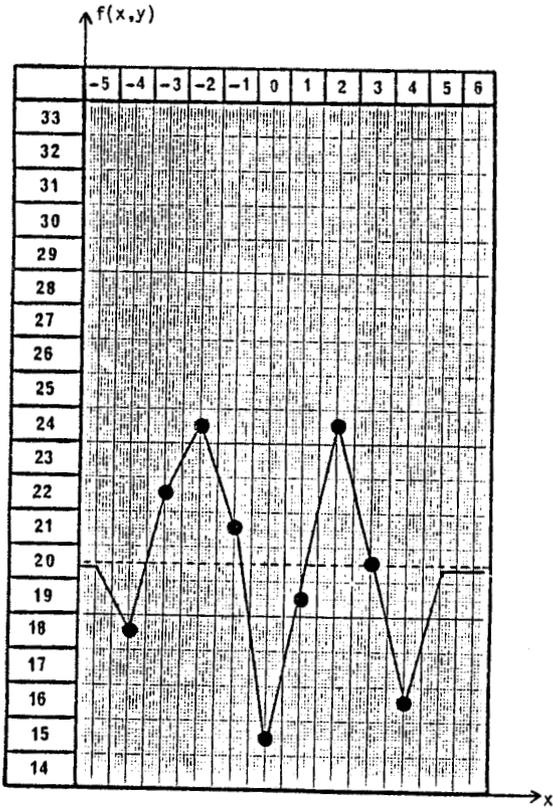
Variation de la luminosité suivant une ligne d'un point Braille "classe B"



(d)

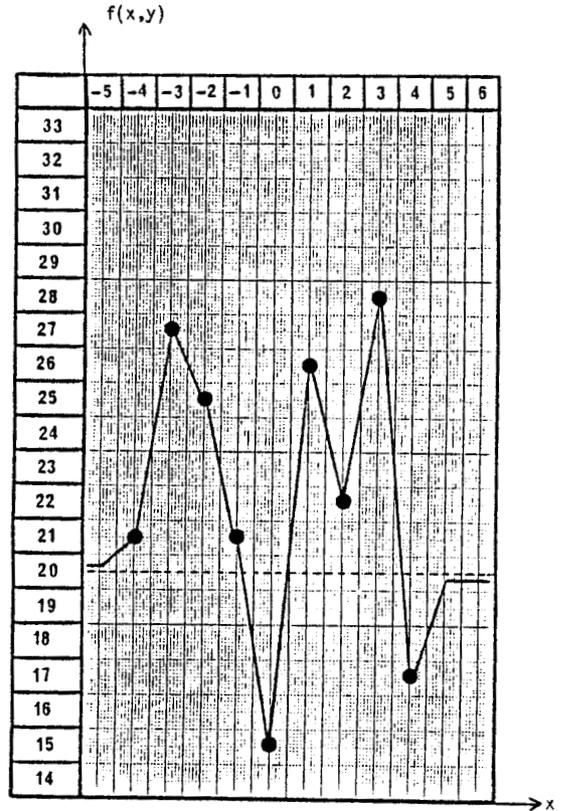
Variation de la luminosité suivant une ligne de 2 points Braille "classe A"





(e)

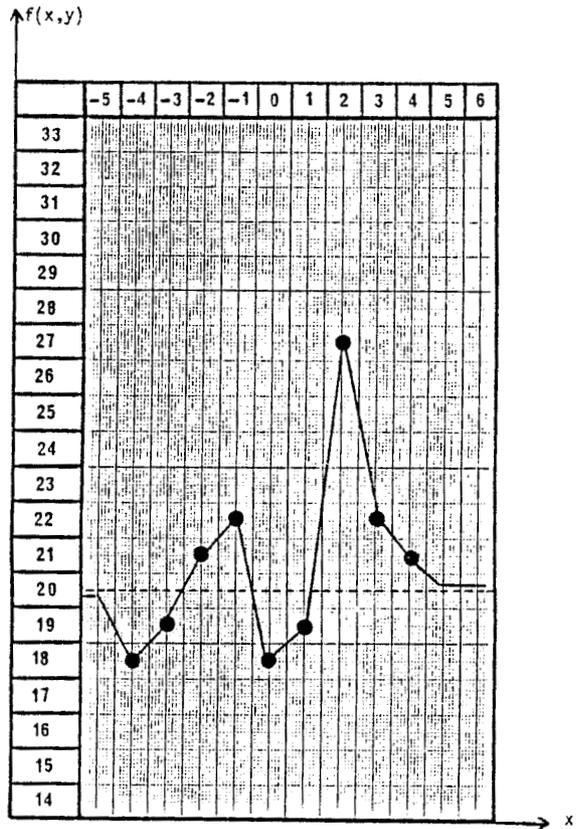
Variation de la luminosité suivant une ligne de 2 points Braille "classe 8"



(f)

Variation de la luminosité suivant une ligne de 2 points Braille "classe 0"

Figure I-15



(g)

Variation de la luminosité suivant une ligne de 2 points Braille "classe 0"



		Fréquences d'apparition			
		Classe A	Classe B	Classe C	Classe D
Procédé de fabrication du relief	Nombre total d'échantillons (graphème)	$f_A = \frac{N_A}{N_T}$	$f_B = \frac{N_B}{N_T}$	$f_C = \frac{N_C}{N_T}$	$f_D = \frac{N_D}{N_T}$
"Embossage mécanique"	$N_T = 2834$	86,74 %	10,76 %	0,84 %	1,66 %
"Embossage manuel"	$N_T = 3256$	69,30 %	21,34 %	2,17 %	7,19 %

Tableau 1 = Fréquences d'apparition de différentes classes du relief Braille.

Nous pouvons déduire de cette analyse que nous sommes contraints de mettre à la même valeur les objets appartenant aux graphèmes de la classe A, B et une partie de la classe D qu'on ne peut estimer avant toute segmentation de l'image.

Il reste à résoudre le problème de la classe C qui est indépendant du filtrage numérique.

Si on se place dans l'hypothèse qu'au minimum les graphèmes appartenant aux classes A et B seront extraites, alors le taux de réussite d'extraction de caractères en relief Braille sera compris entre :

$97,5 < T_R < 1$  pour un document embossé mécaniquement

$90,64 < T_R < 1$  pour un document embossé manuellement.

La reconnaissance de relief ne peut être assurée à 100 %, même dans le cas des documents embossés mécaniquement. Il faut donc prévoir une interactivité dans le système de transcription permettant une correction du texte noir.

## V - CONCLUSION

Dans ce chapitre, nous avons essayé de mettre en évidence de façon générale la plupart des difficultés posées pour extraire l'information relief. Mais, l'objectif majeur est celui de pouvoir extraire totalement les caractères de la classe B et récupérer le maximum de ceux de la classe D si possible.

C'est dans ce but que nous allons introduire au chapitre II de cette thèse quelques méthodes de traitement d'images qui peuvent fournir une solution théorique au problème.

Dans ces conditions, si le filtrage permet d'atteindre l'objectif prévu, on pourra obtenir une extraction voisine de 95 %. En espérant extraire 95 % de l'information, il est probable qu'après transcription en noir, l'analyse du contexte permette de restituer les 100 % du document.

CHAPITRE II : PRÉSENTATION DE DIFFÉRENTES MÉTHODES POUVANT FOURNIR  
UNE SOLUTION THEORIQUE AU PROBLEME DE DETECTION DU  
RELIEF BRAILLE.

## I - INTRODUCTION

Une immense quantité d'informations transite par le canal visuel humain afin d'être traitée par le cerveau, et c'est un des objectifs importants de l'informatique que de pouvoir remplacer l'homme dans ce type de tâche. Cela suppose naturellement que, d'une part, des moyens matériels permettent aux unités centrales un accès aisé à une information bi-dimensionnelle, et d'autre part, que soit développé un immense logiciel de traitement de ces images numériques. Les techniques de reconnaissance des formes ont naturellement un grand rôle à jouer dans ce domaine, à la condition que l'on puisse définir "les formes" d'images.

Ce chapitre a pour but d'aborder les problèmes de la numérisation des images, et de présenter les techniques de traitement (d'une part, globales : renforcement de contraste, lissage, extraction de "squelette" ; d'autre part, locale : suivi de frontières, recherche d'éléments descriptifs, et un rappel de deux opérations élémentaires de la morphologie mathématique). Ces deux derniers points ont pour but de fournir à partir des images numériques considérées comme des "formes" les paramètres qui serviront à leur reconnaissance.

## II - NUMERISATION

Les considérations théoriques qui président à la numérisation des signaux bidimensionnels sont moins simples que dans le cas monodimensionnel, où le pas d'échantillonnage est directement lié à la qualité du codage du spectre du signal : la surface  $S$  où le spectre du signal en un point  $(x,y)$  n'est pas nul ne permet pas en général de trouver une bonne discrétisation du plan de l'image. On se contentera donc en général de "paver" le plan soit par des carrés, soit par des hexagones, et d'affecter à chacun de ces pavés une valeur correspondant à l'énergie ou à la densité optique du signal sur cette surface élémentaire.

On démontre que le meilleur pavage régulier est en fait celui à maille hexagonale ; il a également des avantages sur le plan topologique. Cependant, les systèmes pratiques de numérisation des images utilisent tous presque sans exception le pavage carré. Nous supposons donc dans toute la suite que les images étudiées sont numérisées de cette façon<sup>[6]</sup>.

Un autre problème posé à la numérisation est celui de la quantification du signal : il faut choisir en effet le nombre de bits sur lequel coder les valeurs en chacun des éléments du pavage, et savoir à quel intervalle de valeurs on affectera quel niveau quantifié. On peut, pour perdre le minimum d'informations, calculer ces paramètres à partir des propriétés statistiques du signal : si celui-ci est en moyenne très lumineux, on affectera plus de niveaux de quantification aux valeurs "claires" qu'aux valeurs "sombres". Ceci permet de réaliser une quantification adaptée au mieux au signal étudié. On peut faire divers calculs permettant de réaliser cet ajustage de façon optimale<sup>[7]</sup>.

En général, le nombre de niveaux retenu pour la quantification du signal est une puissance de 2. Ceci pour optimiser le codage en éléments binaires. On dira alors que l'image est codée en  $2^n$  "niveaux de gris", les valeurs 0 et  $2^n$  correspondant l'une au noir, l'autre au blanc.

Un cas particulier important est la quantification binaire des images : les points sont alors codés sur deux niveaux seulement, indiquant la présence ou l'absence de signal sur le pavé correspondant.

### III - DEFINITION DU VOISINAGE D'UN POINT

Une image à niveaux de gris est définie par L lignes et C colonnes. Nous noterons  $P(I,J)$  le point de la I<sup>ème</sup> ligne et de la J<sup>ème</sup> colonne. Considérons une image à niveaux de gris, on définit le voisinage  $V(m,n)$  du point  $P(I,J)$  comme étant la fenêtre de taille  $(m \times n)$ , m lignes et n colonnes, centrée sur  $P(I,J)$ . Dans ces conditions, le point  $P(I,J)$  est au centre d'une fenêtre de taille  $m \times n$  si, et seulement, si m et n sont impairs.

Nous noterons  $V(m,n) = \{a_i\}$ ,  $i$  variant de 1 à  $m \times n$ .

Exemple : La fenêtre  $(m \times n) = (5,3)$  constitue le voisinage  $V(5,3)$  de  $a_8$

			1	2	3		
			4	5	6		
			7	8	9		
			10	11	12		
			13	14	15		

Figure II-1

#### IV - TRAITEMENTS GLOBAUX

On s'intéresse dans ce paragraphe aux divers traitements que l'on peut faire subir à une image numérique d'une façon globale, dans le but de faciliter l'extraction de paramètres (qui sont forcément locaux). On aura ainsi, de la même façon que dans les signaux temporels, des techniques de lissage, d'extraction de caractéristiques fréquentielles, etc...

Ces traitements auront pour objet de préparer l'image à une analyse plus fine qui conduira à leur reconnaissance.

##### IV-1- LISSAGE D'IMAGES

Le lissage a pour but d'éliminer le bruit de l'image, en particulier, le bruit de quantification, afin de régulariser les frontières.

IV-1-1- Définition de la moyenne<sup>[8]</sup>

Considérons une image à niveaux de gris, on calcule la somme S des niveaux de gris des éléments du voisinage V(m,n) associé à un point P(I,J). Cette somme S peut être ramenée pour chacun des éléments de V(m,n) à une valeur moyenne  $\bar{a}$ .

$$\bar{a} = \frac{S}{m.n} = \frac{1}{m.n} \sum_{i=1}^{i=m} \sum_{j=1}^{j=n} a_{i,j} = \frac{1}{m.n} \sum_{j=J-p}^{j=J+p} \sum_{i=I-q}^{i=I+q} P(i,j)$$

où m = 2p+1 et n = 2q+1

La moyenne  $\bar{a}$  est arrondie à l'entier le plus proche.

Le lissage par "filtre moyennneur" est un filtre passe-bas bidimensionnel.

IV-1-2- Définition de la médiane

On définit la valeur M de la médiane au point P(I,J) sur le voisinage V(m,n) comme étant le Kème élément de la suite ordonnée des points de V(m,n).

Le nombre K est l'indice de l'élément  $a_K$  représentant le point P(I,J) dans le voisinage V(m,n).

Posons : m = 2p+1 ; n = 2q+1

(car m et n impairs)

et

$$K = (nm+1)/2 = 2pq + p + q + 1$$

Exemple : n = m = 3

$$K = (nm+1)/2 = 5$$

La valeur M de la médiane au point P(4,4) de l'image I(7,7) sur le voisinage V(3,3) est le cinquième élément de la suite ordonnée des éléments de V(3,3).

La suite des ordonnées des éléments de  $V(3,3)$  est la suivante :

9, 17, 18, 18, 19, 19, 23, 34, 41

Le cinquième élément de cette suite est 19 et par conséquent  $M = 19$ .

22	13	17	29	34	52	48
19	18	17	33	31	27	45
30	17	17	19	23	26	36
20	16	19	18	41	31	33
21	20	18	34	9	52	47
22	28	25	54	43	38	36
18	17	19	17	18	31	29

Figure II-2

Le filtrage par "médiane" est une technique très puissante de lissage d'image, n'affectant pas les contours des objets s'y trouvant. C'est une opération non linéaire<sup>[9][10]</sup>.

Le filtrage par "médiane" s'avère très efficace dans la réduction du bruit, notamment le bruit dit "poivre et sel". Ce résultat, pour être atteint, passe par la recherche de la taille optimale de la fenêtre de traitement. En effet, une fenêtre de petite taille donne des résultats insuffisants, alors qu'une taille trop grande cause une perte d'informations. Cette recherche peut également mener à une combinaison en cascade de filtres médian de taille et de largeur fixes et/ou variables.

Enfin, une forme inchangée par une première application d'un filtre médian, le restera pour d'autres passages ultérieurs.

#### IV-2- RENFORCEMENT DE CONTOUR

Si l'opération de lissage peut s'interpréter comme un filtrage passe-bas de l'image, on peut aussi la voir comme une intégration numérique : chaque point est remplacé par l'intégrale de la valeur du signal sur une fenêtre centrée en ce point.

Il est clair que l'opération inverse, celle de dérivation numérique, aura tendance, elle, à renforcer les contrastes locaux dans l'image, et donc à extraire les contours des objets représentés. Il peut paraître paradoxal de présenter à la suite deux techniques de traitement qui semblent s'annuler, le lissage et le renforcement de contour. Cependant, le premier a pour but de nettoyer quelque peu l'image ; l'autre, d'en extraire des contours, qui sont des informations souvent essentielles pour le traitement. De plus, une image transformée d'abord par lissage, puis par un gradient, un laplacien (gradient et Laplacien de la fonction intensité lumineuse) ou autre ne sera pas identique à l'image source. Les paramètres des opérations peuvent être en effet réglés de façon à ce que leur suite réalise un bon traitement. Dans une méthode de gradient, on remplace donc la valeur  $g(m,n)$  de l'image discrète au point de coordonnées  $(m,n)$  par la valeur :

$$g'(m,n) = \left( [g(m,n) - g(m+1,n+1)]^2 + [g(m+1,n) - g(m,n+1)]^2 \right)^{1/2}$$

ou par

$$g''(m,n) = \left| g(m,n) - g(m+1,n+1) \right| + \left| g(m+1,n) - g(m,n+1) \right|$$

On peut utiliser des "gradients" plus complexes dans le but de réduire la sensibilité au bruit. On a proposé par exemple sur une fenêtre  $(3 \times 3)$ <sup>[11]</sup>:

$$g'(m,n) = (A^2 + B^2)^{1/2} \quad \text{ou} \quad g'(m,n) = |A| + |B|$$

avec

$$A = [g(m+1,n+1) + 2g(m+1,n) + g(m-1,n-1)]$$

$$- [g(m-1,n+1) + 2g(m-1,n) + g(m-1,n-1)]$$

$$B = [g(m-1,n-1) + 2g(m,n-1) + g(m+1,n-1)]$$

$$- [g(m+1,n-1) + 2g(m+1,n) + g(m+1,n-1)]$$

ou encore, sur une fenêtre  $(5 \times 5)$  [12].

Avec la convention :

$S_0^a$  = somme des valeurs des cellules marquées "a" dans le tableau correspondant à 0.

et  $0^\circ \begin{pmatrix} b & b & b & b & b \\ o & o & o & o & o \\ a & a & a & a & a \\ o & o & o & o & o \\ c & c & c & c & c \end{pmatrix}$

$45^\circ \begin{pmatrix} o & b & b & o & a \\ b & b & o & a & o \\ b & o & a & o & c \\ o & a & o & c & c \\ a & o & c & c & c \end{pmatrix}$

$90^\circ \begin{pmatrix} b & o & a & o & c \\ b & o & a & o & c \\ b & o & a & o & c \\ b & o & a & o & c \\ b & o & a & o & c \end{pmatrix}$

$135^\circ \begin{pmatrix} a & o & c & c & o \\ o & a & o & c & c \\ b & o & a & o & c \\ b & b & o & a & o \\ o & b & b & o & a \end{pmatrix}$

$$S_0 = 2 S_0^a - S_0^b - S_0^c$$

$$g'(m,n) = g(m,n) + \begin{cases} \text{Max} (S_0, S_{45}, S_{90}, S_{135}) & \text{si } \geq 0 \\ 0 & \text{si non} \end{cases}$$

On voit que ces méthodes d'une façon générale, s'appuient sur la recherche de la direction de l'éventuelle frontière par le point traité, et renforcent sélectivement la valeur des points critiques.

#### IV-3- METHODES FREQUENTIELLES : TRANSFORMEE DE FOURIER [13]

L'extension de la transformée de Fourier à deux dimensions est une des approches purement globales que l'on peut appliquer aux images. Si l'on note l'image analogique  $g(x,y)$  en lui supposant les propriétés de continuité nécessaires, sa transformée sera :

$$G(f_x, f_y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x,y) \exp [-2 \pi i (f_x x + f_y y)] dx dy$$

L'interprétation physique est que si la fonction  $G$  possède un maximum marqué pour un couple de fréquences  $(u,v)$ , alors l'image origine présentera une ondulation sinusoïdale importante de période spatiale

$$L = \frac{1}{\sqrt{u^2 + v^2}}$$

disposées selon la figure suivante :

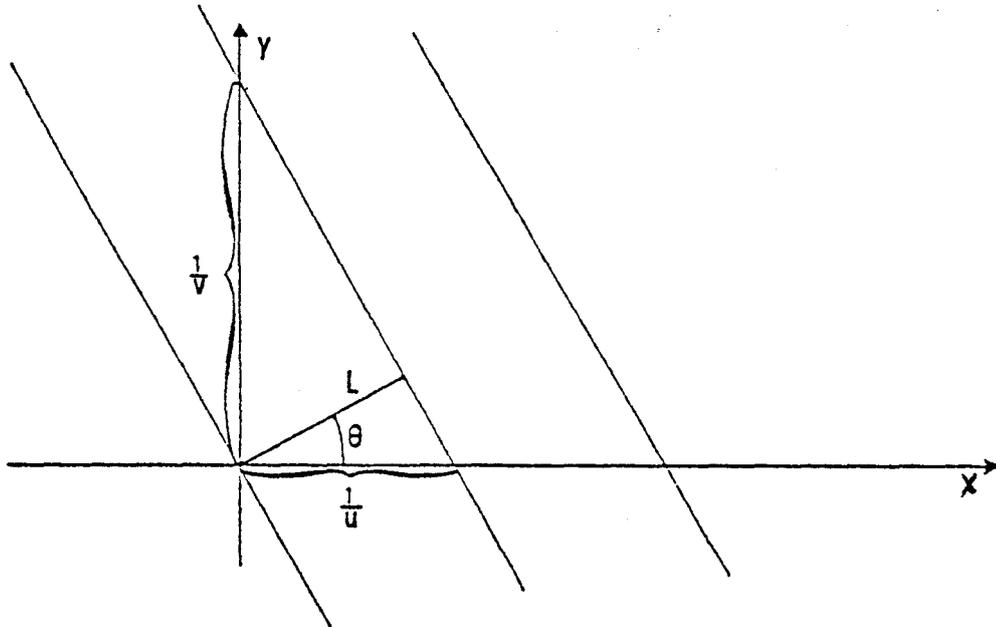


Figure II-3

L'utilisation de la transformée de Fourier permet, comme dans le cas de signaux temporels, de pratiquer des filtrages fréquentiels sur les images.

L'interprétation physique de telles opérations est la suivante : un filtrage passe-bas représente un lissage de l'image, et a tendance à éliminer le bruit et à "écraser" les contours abrupts.

Un filtrage passe-haut, au contraire, renforce les bords, et agit comme les méthodes de gradient (qui sont des cas particuliers de tels filtrages). L'ensemble des méthodes de filtrage par la transformation de Fourier fournit donc une gamme d'opérations globales, qui servent à améliorer sélectivement la qualité de l'image. Elles sont adaptées aux images où apparaissent des fréquences nettes. En aucun cas ce ne sont des méthodes de reconnaissance mais de prétraitement et de préanalyse.

IV-4- METHODES RECURSIVES<sup>[14]</sup>

On assimile ici l'image binarisée à une suite de variables aléatoires discrètes  $\{g_u\}_{u=(x,y)}$  réelles indexées par les coordonnées  $(x,y)$  du point de l'image dans le plan discret,  $g_u$  prendra un nombre fini de valeurs (niveaux de gris). On suppose en plus que la covariance de la suite  $\{g_u\}$  est de type :

$$\begin{aligned} \Gamma(u_1 ; u_2) &= \Gamma(x_1, y_1 ; x_2, y_2) \\ &= T^2 \exp [-\gamma_1 |x_2 - x_1| - \gamma_2 |y_2 - y_1|] \end{aligned}$$

Cette forme de la covariance permet une factorisation en 2 termes : une fonction de  $|x_2 - x_1|$  et l'autre de  $|y_2 - y_1|$ . En plus,  $\Gamma$  ne dépend que de la distance de  $u_1$  à  $u_2$  (cela rappelle la stationnarité temporelle dans le cas d'un processus aléatoire).

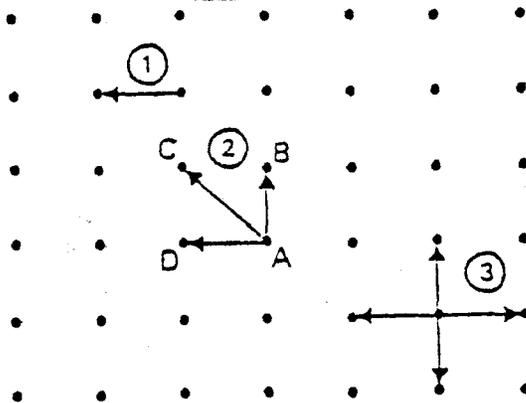


Figure II-4

Quant à la corrélation d'un point de l'image avec son voisinage (fig.II-4), il existe plusieurs modèles. Le plus simple est le type (1) où le point est corrélé avec son antécédent immédiat (dans le sens de balayage de l'image) ; c'est donc un modèle monodimensionnel, le modèle est traité dans<sup>[15]</sup>. Un modèle plus fin de type (2) où l'on tient compte du caractère bidimensionnel de l'image et où le point A est corrélé avec les 3 voisins B, C et D<sup>[16]</sup>. On donne comme exemple un modèle plus élaboré (3) qui suppose que le point est corrélé avec ses 4 voisins.

On pose  $g_u = g(x,y)$        $u = (x,y)$

On prend une corrélation locale de la forme :

$$g(x,y) = \alpha_1 g(x,y+1) + \alpha_2 g(x+1,y) + \alpha_3 g(x,y-1) + \alpha_4 g(x-1,y) + \beta K(x,y)$$

où  $\alpha_1, \alpha_2, \alpha_3, \alpha_4$  et  $\beta$  sont des constantes à déterminer et  $\{K_u\}_{u=(x,y)}$  une suite de variables aléatoires réelles centrées indépendantes indexées par  $u=(x,y)$  et de même variance que celle de  $\{g_u\}_{u=(x,y)}$  et indépendantes des variables aléatoires dans le 4-voisinage.

Ainsi énoncé et ayant les conditions aux limites sur  $g$ , ce problème est de type de Kalman<sup>[17]</sup>.

L'inconvénient des méthodes récursives c'est qu'on impose trop d'hypothèses à priori sur l'image pour avoir un modèle soluble. De plus, les données du second ordre (les moments d'ordre 1 et 2) ne suffisent pas pour déterminer complètement et statistiquement le processus, et la correspondance entre la covariance du processus et sa loi spatiale (tous les moments) n'est pas bijective.

Le filtrage par les méthodes récursives devient lourd et compliqué lorsque l'on veut des bons résultats.

#### IV-5- SQUELETTISATION<sup>[18]</sup>

On désigne sous ce terme un ensemble de techniques qui visent à extraire des images numériques des sous-images plus faciles à traiter, en réduisant par exemple une dimension. Le cas où l'on pratique ces traitements le plus couramment est celui des caractères manuscrits. Un exemple montrera plus clairement le but recherché (fig.II-5)

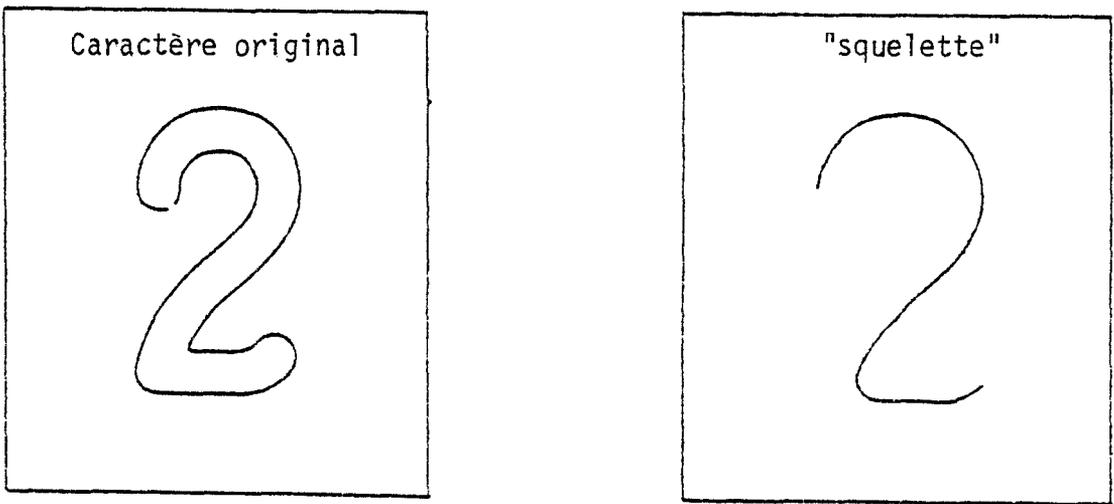


Figure II-5

On ne possède pas de définition qui ait à la fois assez de rigueur topologique et une traduction algorithmique<sup>[19]</sup>. Il faudra donc se contenter d'imposer aux algorithmes de squelettisation des contraintes telles que :

- la largeur du squelette doit être de 1 (le concept de "largeur" étant à définir),
- les propriétés topologiques de l'image originale doivent se retrouver dans celles du squelette (nombre de parties connexes, nombre de trous etc...).

Nous présenterons ici deux méthodes de squelettisation, l'une à critère topologique, l'autre connue sous le nom de "grass-fire".

#### 1ère méthode : "Shrinking algorithms"

Ce sont des algorithmes de réduction de l'image qui préservent la connexité dans l'image. Cette catégorie d'algorithmes est particulièrement intéressante dans le cas du caractère manuscrit (chiffre, lettre, ...). En effet, il apparaît évident que dans ce cas là, la première réduction que l'on doit effectuer sur le caractère aura pour but de tirer un squelette de ce caractère (ligne moyenne d'épaisseur 1), donc de préserver une caractéristique pertinente qui est la connexité.

La squelettisation consistera donc à détecter les points inessentiels et de les remettre à zéro. Ceci peut se faire de deux manières différentes : S(shape) désigne la forme à squelettiser dans l'image E ; SCE.

#### 1ère manière : type série

On lit l'image E séquentiellement et quand on trouve un point inessentiel pour S, on le remet à zéro. La lecture de l'image est itérée jusqu'à la suppression de tous les points inessentiels de S. Cette méthode est bonne et très adaptée à une implantation série (logicielle).

#### 2ème manière : type parallèle

Elle est plus compliquée que la précédente mais supprime en parallèle plusieurs points à la fois.

Ci-contre les quatre points du carré sont :  $\begin{matrix} 1 & 1 \\ 1 & 1 \end{matrix}$  chacun à part, inessentiels, mais si on les supprime tous ensemble on change le nombre des composantes connexes de E. Mais si on commence par supprimer les 2 points supérieurs, par exemple  $\begin{matrix} 0 & 0 \\ 1 & 1 \end{matrix}$ ; ce qui reste est essentiel.

Donc dans la méthode parallèle on décompose le traitement en sous-cycles. Une méthode possible consiste à construire 4 ensembles :

$$E_1 = \{p \in S ; P_3 = 0\}$$

$$E_2 = \{p \in S ; P_3 = 1 \text{ et } P_7 = 0\}$$

$$E_3 = \{p \in S ; P_3 = P_7 = 1 \text{ et } P_1 = 0\}$$

$$E_4 = \{p \in S ; P_3 = P_7 = P_1 = 1 \text{ et } P_5 = 0\}$$

On cherche les points inessentiels de  $E_1$ , puis on les remet à zéro simultanément. On continue successivement pour  $E_2$ ,  $E_3$  et  $E_4$ . On reconstruit  $E_1, E_2, E_3$

Exemple :

	(RAZ $E_1$ )	(RAZ $E_2$ )	(RAZ $E_3$ )	(RAZ $E_4$ )
1 1 1 1 1	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0
1 1 1 1 1	1 1 1 1 1	1 1 1 1 1	1 1 1 1 0	0 1 1 1 0
1 1 0 1 1	1 1 0 1 1	1 1 0 1 1	1 0 0 1 0	1 0 0 1 0
1 1 1 1 1	1 1 0 1 1	1 1 0 1 1	1 1 0 1 0	0 1 0 1 0
1 1 1 1 1	1 1 1 1 1	0 0 1 0 0	0 0 1 0 0	0 0 1 0 0

(RAZ  $E_1$ )  
 0 0 0 0 0  
 0 1 1 0 0  
 1 0 0 1 0  
 0 1 0 1 0  
 0 0 1 0 0

Plus de points inessentiels.

2ème méthode : "grass-fire"

L'approche intuitive est la suivante : on suppose que le fond de l'image est composé d'herbe humide et la figure connexe à squelettiser d'herbe sèche. On allume un feu simultanément sur tous les points frontières de cette figure, et le feu se propage à vitesse uniforme vers l'intérieur. L'ensemble des points où se rencontrent deux lignes de feu formera le squelette de la figure.

Traitons d'abord le cas en supposant l'image non numérisée, avant de voir l'application au cas des images discrètes. Supposons choisie une métrique  $d$  sur l'image : c'est-à-dire qu'à deux points quelconques  $x$  et  $y$  de l'image, on peut associer une valeur  $d(x,y)$  satisfaisant les propriétés d'une distance.

On définira la distance d'un point  $x$  à un ensemble de points  $A$  par :

$$d(x,A) = \text{Min}_{y \in A} \{d(x,y)\}$$

Soit  $B$  la frontière de la figure. Le squelette  $S$  sera alors défini par :

$$S = \{x \mid \exists y, Z \in B : d(x,B) = d(x,y) = d(x,Z)\}$$

Il est intéressant de définir aussi le couple  $(S,q)$  par :

$$(S,q) = \{(x,d(x,B)) \mid x \in S\}$$

On peut remarquer alors que la figure originale peut être reconstituée à partir de  $(S,q)$ , ce qui est une propriété intéressante, liée à la définition de  $d$  sur l'image.

Cette définition du squelette serait très intéressante si on pouvait l'appliquer directement aux images discrétisées. Il suffirait alors pour chaque point de la figure de décider s'il appartient ou non au squelette, selon la formule présentée. Malheureusement, la cohérence de la topologie du contenu ne peut pas se transmettre dans le cas discret, et on arrive à certaines contradictions.

Pour rendre cependant opératoire la méthode "grass-fire", on peut procéder pas à pas, en partant de la frontière de l'image, et donc simuler le comportement de l'incendie en question.

## V - TRAITEMENTS LOCAUX

On désigne par le terme de "traitements locaux" dans les images numérisées une série de techniques permettant d'en extraire des paramètres, à l'examen de certaines parties de l'image. Ces points seront en quelque sorte individualisés à l'examen de leur contexte, et non pas un traitement identique à tous.

Il est parfois difficile d'attribuer l'étiquette "local" ou "global" à tel ou tel algorithme de traitement (la squelettisation, par exemple). Nous présenterons ici quelques méthodes classiques ; elles concernent d'abord la paramétrisation des frontières des figures (ou des courbes, de façon plus générale) : c'est en effet un problème important que de pouvoir décrire le pourtour d'une figure en termes simples de segments de droite. On réalise alors une grande compression d'information, et donc une bonne paramétrisation.

Pour cela, il faudra d'abord posséder une représentation de la frontière permettant de la suivre dans un certain sens (et non pas la connaître simplement par l'ensemble de ces points non ordonné, comme par exemple pour le "grass-fire"). Des algorithmes de "suivi de contour" sont donc à développer avant toute paramétrisation de ce contour.

Une attention particulière sera aussi portée sur la détection de points supposés importants pour représenter l'image : angles, arêtes, points d'intersection de frontières etc...

### V-1- SUIVI DE CONTOUR

Les méthodes présentées ici sont plus classiques de forme. Parmi les méthodes globales les techniques de gradient permettent de trouver un ensemble de points frontière, et que la définition donnée dans la méthode de squelettisation "grass-fire" fournit elle aussi cet ensemble. Ce sont des méthodes globales, en ce sens que l'ensemble des points trouvés n'est pas ordonné, et qu'aucune relation de connexité n'est fournie entre les points.

V-1-1- Suivi de bord

Soit une surface  $S$  composée d'éléments à valeur "1" sur un fond  $\bar{S}$  d'éléments à valeur "0", l'ensemble constituant une image binaire ; on supposera cette figure 4-connexe, et on cherchera à définir sa frontière comme une suite de points, chacun étant 4-connexe au précédent.

On appellera "bord" de la figure  $S$  un ensemble de 2 éléments 4-connexes, dont l'un est dans  $S$  et l'autre dans  $\bar{S}$ .

Soit  $(x_k, y_k)$  un bord de  $S$  avec  $x_k = 1$  et  $y_k = 0$ . On peut considérer sans perte de généralité avoir une configuration du type "frontière à main gauche".

a	b
$x_k$	$y_k$

On cherche donc quels éléments composeront le bord suivant  $(x_{k+1}, y_{k+1})$  en fonction des valeurs de  $a$  et  $b$ . Le tableau suivant fournit un algorithme cohérent :

a	b	$x_{k+1}$	$y_{k+1}$
0	0	$x_k$	a
1	0	a	b
0	1	b	$y_k$

On démontre alors le théorème suivant dû à Rosenfeld<sup>[20]</sup>.

Soit  $(x_0, y_0)$  un bord quelconque de la figure connue 4-connexe  $S$ . Il existe un entier  $m$  tel que  $(x_m, y_m) = (x_0, y_0)$ , et chaque bord de  $S$  apparaît une fois et une seule dans la suite  $\{(x_i, y_i) \mid i = 1, m\}$ .

Cet algorithme assure donc un suivi de contour, par une suite minimale de "bords" de la figure.

La suite des points  $x_k$  est la frontière de la figure que l'on réduit immédiatement de la suite de bords trouvés par cet algorithme.

### V-1-2- Suivi direct de frontière

On peut démontrer que l'algorithme suivant est équivalent à la détection de bord présentée au paragraphe précédent, et donc qu'il permet un suivi intégral de la frontière de la figure.

Soit  $x_0$  un élément quelconque de la frontière cherchée. Soit  $y_1, y_3, y_5$  et  $y_7$  les éléments de son 4-voisinage (Si aucun des 4 ne vaut 1, la figure se compose uniquement de  $x_0$ ).

Soit  $y_{2i+1}$  le premier d'entre eux de valeur 1 (dans un certain ordre arbitraire). Soit  $x_1$  le point frontière suivant cherché. On a alors :

$$\begin{array}{lll} x_1 = y_{2i+1} & \text{si} & y_{2i} = 0 \\ x_1 = y_{2i} & \text{si} & y_{2i} = 1 \end{array}$$

L'ordre choisi sur les  $y$  imposera le sens de rotation autour de la figure.

Dans le cas de 8-connexité, l'algorithme sera simplement :  
 $x_1$  est le premier 8-voisin de  $x_0$  à valeur 1 (dans un ordre arbitraire imposé aux  $y$ ).

### V-2- APPROXIMATION D'UNE COURBE PAR UN POLYGONE<sup>[21][22]</sup>

On cherche ici particulièrement à réduire l'information contenue dans une courbe discrète, en codant celle-ci sous la forme d'un ou plusieurs segments de droite, caractérisés par leurs extrémités (figure II-6). Une telle courbe pourra être en particulier la frontière d'une figure ; dans ce cas, la paramétrisation de celle-ci conduira à la représenter sous la forme du polygone qui approche le mieux sa frontière, ce qui représente une intéressante réduction d'information sur un critère opératoire.

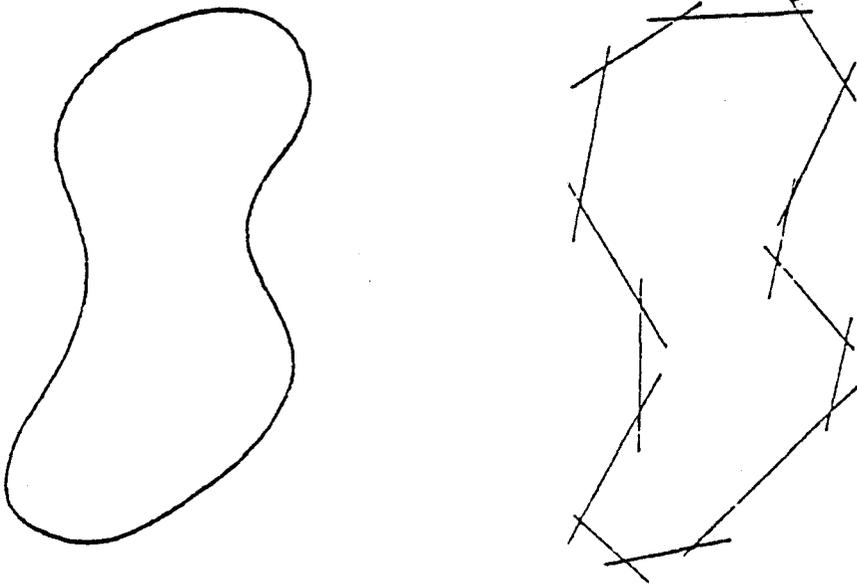


Figure II-6

V-2-1- Approximation d'un ensemble par une droite

Dans un premier temps, on cherche à trouver la meilleure droite représentant l'ensemble des points de la courbe. Ceci est tout à fait identique à la recherche de l'axe principal d'inertie de cet ensemble de points. On peut en effet naturellement chercher à minimiser le terme  $d^2 = \sum_{i=1}^n d_i^2$

où  $d_i$  est la distance du  $i^{\text{ème}}$  point de la courbe à la droite cherchée.

On démontre facilement dans un premier temps que cette droite doit passer par le centre de gravité des points ; on y prendra donc l'origine du repère, il reste à déterminer le vecteur directeur de la courbe (ou plus spécifiquement ici, son vecteur normal).

Soit  $N$  le vecteur normal de la droite cherchée,

Soit  $V_i = (x_i, y_i)$  le vecteur représentant chaque point. Il faut minimiser pour les  $n$  points :

$$\begin{aligned} d^2 &= \sum_{i=1}^n d_i^2 = \sum_{i=1}^n (N^t V_i)^2 \\ &= \sum_{i=1}^n (N^t V_i)(N^t V_i)^t = \sum_{i=1}^n N^t V_i V_i^t N \\ &= N^t \sum_{i=1}^n (V_i V_i^t) N \end{aligned}$$

Si on note  $S$  la matrice  $(\sum_{i=1}^n (V_i V_i^t))$ , matrice (symétrique) de covariance de l'ensemble des points, il faut chercher  $N$  tel que :

$$d^2 = N^t S N \text{ soit minimal}$$

Sous la contrainte  $\|N\| = 1$ , on démontre que le minimum de  $d^2$  est atteint pour  $N$  vecteur propre de  $S$  associé à la plus petite valeur propre.

Ce qui implique que la direction de la droite qui approxime le mieux l'ensemble de points qui est perpendiculaire à  $N$  est parallèle au vecteur propre associé à la plus grande valeur propre de  $S$  (axe principal d'inertie).

### V-2-2- Approximation polygonale d'une courbe

On a ici à résoudre en parallèle deux problèmes ; le premier est de segmenter les points en sous-ensembles, le second est de trouver le meilleur segment pour chacun des sous-ensembles. On présente ici quelques exemples d'algorithmes classiques.

#### V-2-2-1- Approximation itérative

On procède ici à une extension présentée au paragraphe précédent, de la façon suivante : on va calculer itérativement la valeur  $d^2$  en prenant les points un par un, dans l'ordre de la courbe numérique, à partir d'une certaine origine.

Si  $d^2$  dépasse un seuil fixé à l'avance, on conserve alors la droite trouvée et on réinitialise l'opération : le point courant devient une nouvelle origine pour le calcul.

Cet algorithme fournira des réponses différentes selon l'origine choisie au départ. Il n'est donc pas indifférent de la sélectionner soigneusement. Par exemple :

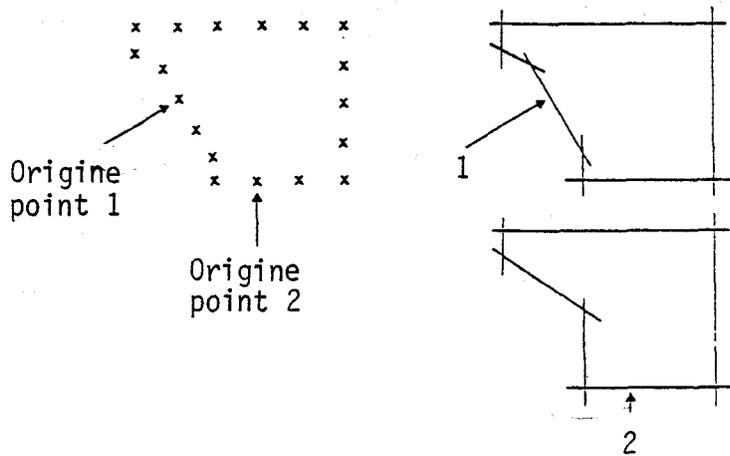


Figure II-7

Il est intéressant de noter que cet algorithme peut être couplé en parallèle à un suivi de contour. Il faut pour cela savoir calculer à partir de l'axe principal d'inertie d'un ensemble  $S_u$  de points, celui de l'ensemble  $S_{u+1} = S_u \cup \{x_{u+1}\}$  où  $x_{u+1}$  est le point suivant détecté sur la frontière. Ce calcul est classique en analyse de données ; mais l'implantation n'en est pas forcément un gain de temps. D'autre part, le réglage du seuil est crucial pour cet algorithme, car il influe directement sur le nombre de segments trouvés en fin de calcul.

#### V-2-2-2- Approximation directe par un polygone

##### Algorithme des "points frontière"

Cette méthode traite des courbes fermées, mais peut leur être étendue. Elle se décrit comme suit. Dans l'ensemble des points, on cherche les deux points extrêmes A et B et on trace la droite en les joignant. Si la somme des distances des points à la droite est inférieure à un seuil donné, le processus s'arrête. Si non, on cherche le point C le plus distant de AB, et on trace les droites AC et CB. On répète le processus indépendamment sur les deux droites AC et BC avec éventuellement une nouvelle détermination des seuils et on poursuit jusqu'à la décision d'arrêt sur tous les segments (figure II-8). Cet algorithme fournit de bons résultats sur des courbes peu bruitées.

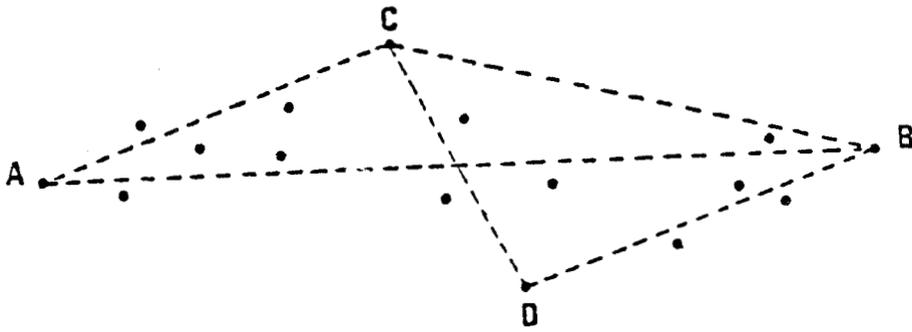


Figure II-8

$$(AB) \rightarrow ((AC),(CB)) \rightarrow ((AC),((CD),(DB)))$$

V-2-2-3- Un autre type d'approximation globale peut être défini en se fixant un critère global sur l'ensemble des points.

C'est ainsi qu'il existe des algorithmes cherchant à approximer une courbe fermée par un polygone ayant un nombre fixé de sommets ou ayant une longueur minimale. Ces procédés très intéressants ne sont pas détaillés ici, mais on trouvera leur description dans les références [21] et [22].

V-3- CODAGE DE FREEMAN

Le codage de Freeman<sup>[23]</sup> consiste à remplacer une courbe (suite de points dans le plan discret) par une suite de valeurs exprimant les positions relatives de 2 points consécutifs par la direction prise pour passer de l'un à l'autre de ces 2 points dite direction de transitions.

Les directions de transitions entre 2 éléments consécutifs de la courbe sont codées de la façon suivante (figure II-9) :

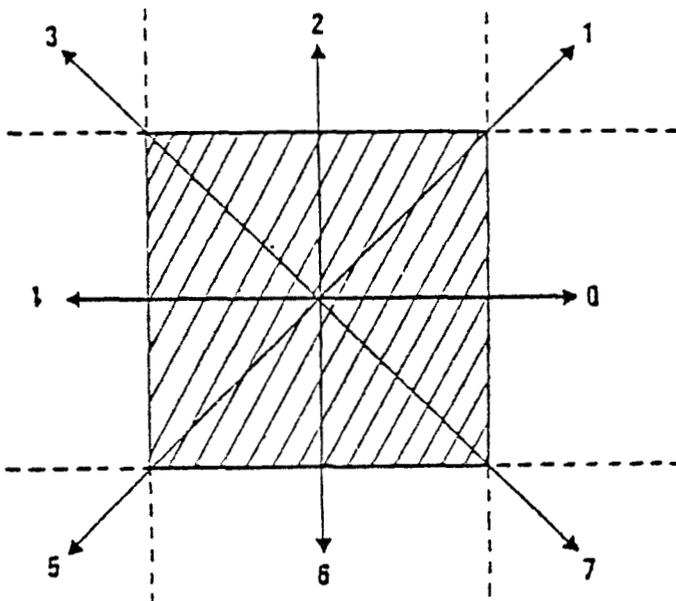
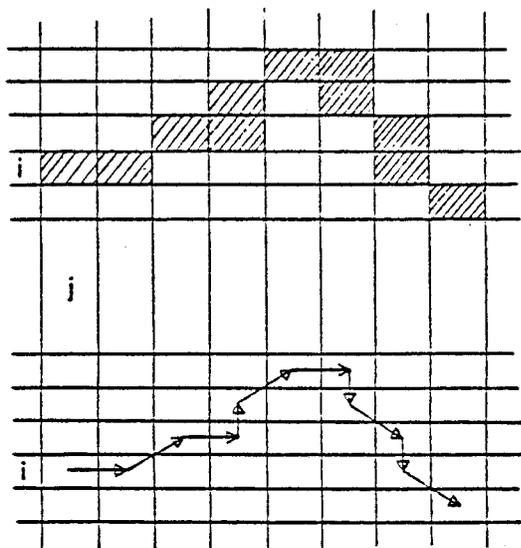


Figure II-9

La courbe numérisée suivante sera donc codée ainsi (figure II-10) :

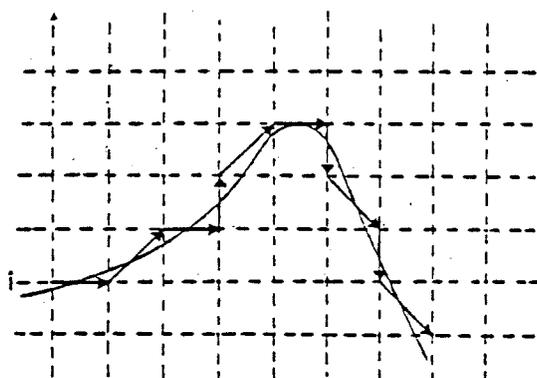


Point de départ : (i;j)  
Codage de la courbe 01021066767

Figure II-10

On peut aussi présenter ce codage en supposant avoir à coder une courbe analogique d'épaisseur nulle, sur laquelle on superpose une grille (de maille arbitraire). A chaque intersection avec une ligne de cette grille, on affecte à la courbe le noeud de la grille la plus proche,

On code ensuite les successions de tels noeuds (figure II-11).



Point de départ : (i,j)

Courbe

Codage de cette courbe  
0102106767

Figure II-11

Le codage de Freeman se prête à deux utilisations, l'une pour la transmission, l'autre pour le prétraitement et l'extraction de paramètres.

Dans le premier cas, il s'agit de représenter une image par un codage qui réduise son information, afin de diminuer son coût de transmission. La technique de Freeman peut être utilisée pour coder les lignes d'égalité d'intensité ("courbe de niveau") ; on peut d'autre part remplacer une succession de valeurs égales par un surcodage encore plus économique, et d'une façon générale employer les "astuces" habituelles de la transmission numérique.

Dans le second cas, les propriétés intrinsèques du codage de Freeman peuvent être utilisées aux fins de prétraitement. Il permet en effet de fournir un encadrement assez précis de la longueur exacte de la courbe codée, de trouver des lignes perpendiculaires, ou faisant un angle donné avec cette courbe ; on peut aussi calculer l'approximation de la surface d'une courbe fermée. Cette dernière est donnée par :

$$S = \sum_{i=1}^n a_{ix} (y_{i-1} + \frac{1}{2} a_{iy})$$

où  $a_{ix}$  (resp.  $a_{iy}$ ) est la composante selon l'axe horizontal (resp. vertical) du 1er élément  $a_i$  de la chaîne code et  $y_{i-1}$  l'ordonnée du point de départ de l'élément  $a_i$  (mesurée à partir d'un axe arbitraire).

Le codage de Freeman permet de même, pour la reconstitution d'éléments manquants (lignes cachées, par exemple) de synthétiser la chaîne correspondant à la droite joignant deux points donnés ou celle de longueur minimum joignant deux points.

Les lignes droites ont des codages possédant des propriétés intéressantes. Une bonne approximation peut en être faite par des grammaires régulières ; les contraintes sur la suite des éléments de leur codes sont en effet très fortes, et la mise sous forme de "grammaire" de ces contraintes permet de les modéliser.

Si le codage de Freeman est insensible à la translation, il ne l'est pas à la rotation, et l'on ne peut pas déduire du codage d'une courbe celui de sa transformée par rotation.

Malgré ce problème de bruit engendré par la rotation, ce codage peut néanmoins servir à comparer la forme de deux segments de courbe, et donc à extraire des formes primitives d'une image. La corrélation de deux chaînes, dont l'une sert de référence  $A = a_1, \dots, a_n$  et  $B = b_1, \dots, b_m$  ( $n < m$ ) est donnée par :

$$\phi_{ab}(j) = \frac{1}{n} \sum_{i=1}^n \cos(a_i - b_{i+j}) \frac{\pi}{4}$$

On peut se servir de cette formule pour décomposer une frontière d'image en formes primitives dont on connaît un codage standard ; ceci suppose que l'orientation globale de la partie de frontière à analyser soit connue, afin d'éviter l'erreur de direction.

VI - LES K FORMES. DEFINITION ET APPLICATIONS NOUVELLES EN TRAITEMENT  
D'IMAGES [ 24 ]

VI-1- IMAGE : TABLEAU ET GRAPHE

Une image digitalisée peut être considérée comme un tableau de pixels ou comme un graphe planaire. Dans ce dernier cas on peut étiqueter soit les sommets par la quantification de l'intensité de gris (figure II-12-b), soit les arêtes par la quantification de la différence de niveau de gris (ou gradient) (figure II-12-b).

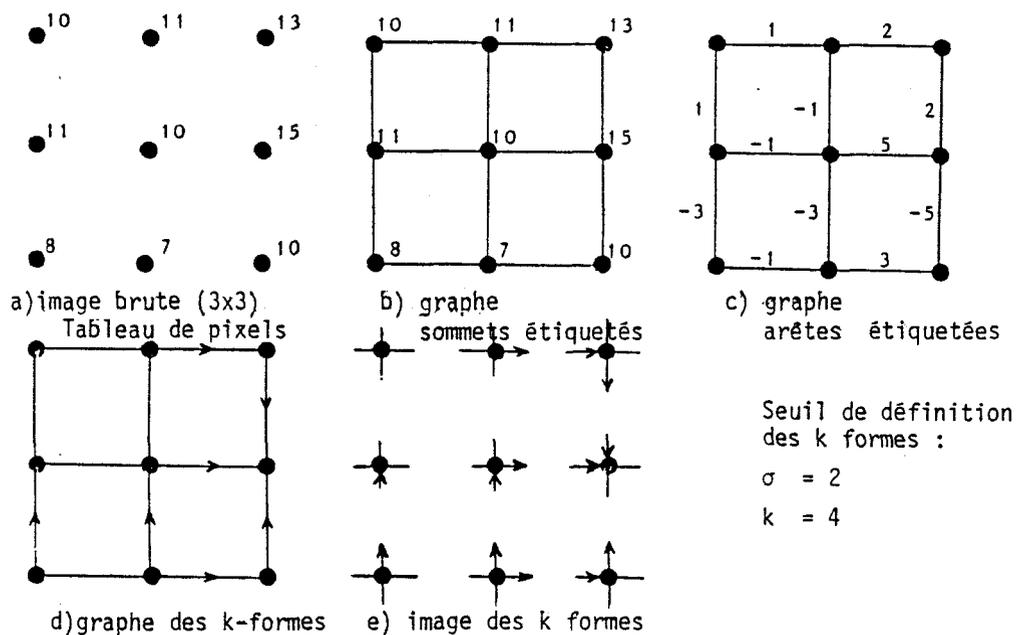


Figure II-12 : Images et graphes associés

La représentation sous forme de tableau n'est universellement adoptée qu'à cause de l'organisation mémoire des ordinateurs et donc la possibilité très pratique de l'accès direct à un pixel.

VI-2- CODIFICATION DES K-FORMES (MALLAGE RECTANGULAIRE DROIT)

Soit la fonction de quantification à trois niveaux (0,1,2) définie comme suit :

$$\theta_{\sigma} : \mathbb{R} \rightarrow \{0,1,2\} \quad : \text{seuil donné } \sigma \in \mathbb{R}$$

$$x \rightarrow \theta_{\sigma}(x) = 1 + \text{sign} [\min (\max(x, -\sigma), x + \sigma)]$$

ou plus simplement sous forme d'algorithme et de graphe (figure II-13)

$$\theta_{\sigma}(x) = \begin{cases} 0 & \text{si } x < -\sigma \\ 1 & \text{si } -\sigma \leq x \leq \sigma \\ 2 & \text{si } x > \sigma \end{cases}$$

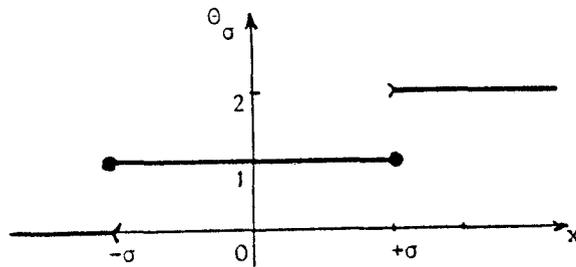
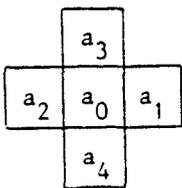


Figure II-13

Pour définir le plus simplement possible la codification décimale  $f_{\sigma}^k(v)$  d'une 4-forme associée à un pixel de référence de 4-voisinage  $v = (a_0, a_1, a_2, a_3, a_4)$  on peut procéder comme suit :



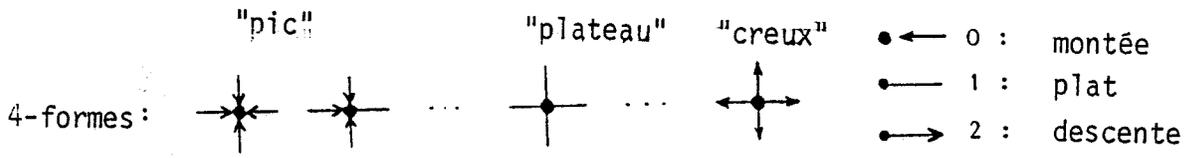
$$f_{\sigma}^k(v) = \sum_{i=1}^k 3^{i-1} \theta_{\sigma} (\Delta_i(a_0)) \quad k = 2 \text{ ou } 4$$

La pente dans la direction  $i \in [1, K]$  sera notée  $\Delta_i(a_0)$  avec ici  $K = 4$ ,

$\Delta_i(a_0) = a_i - a_0$  (dans le cas le plus simple).

Pour une image digitalisée à maillage rectangulaire droit on utilise surtout les 2-formes et les 4-formes (figures II-14 et II-15).

Il y a donc  $3^4$  soit 81 4-formes possibles décrivant l'information structurelle locale au niveau du pixel. La 4-forme codée 0 correspond à un "pic" en intensité de gris, et celle codée 80 à un creux (figure II-14).



décimal : 0    1    ...    40    ...    80  
 ternaire : 0000 0001 ... 1111 ... 2222

Figure II-14 : codage des 4-formes

Les K-formes traduisent le "relief" en intensité de gris d'une image. Décrire ce relief, c'est s'intéresser à la texture ou à la segmentation en régions ; extraire ou suivre les lignes de crête et/ou les lignes de plus grande pente, c'est faire de la détection de contours ou du suivi de contours ; modifier adéquatement ce relief c'est faire du lissage, du filtrage, de l'amélioration ou de la restauration d'image.

La figure II-15 donne la correspondance directe entre les 4-formes (K=2) et la forme topologique correspondante.

code décimal	code ternaire	représentation à une dimension	représentation topographique
0	00		pic
1	01		épaulement ouest
2	02		versant ouest
3	10		épaulement est
4	11		plateau
5	12		pied de versant ouest
6	20		versant est
7	21		pied de versant est
8	22		vallée

Figure II-15 : analogie topologique pour les 2-formes

On peut entrevoir l'universalité de l'emploi des K-formes en prétraitement d'images. Les K-formes peuvent se déduire également par une quantification de la dérivée première de la fonction intensité de gris (gradient G) et de la dérivée seconde (Laplacien) (figure II-16).

L \ G	$G > \sigma$	$ G  \leq \sigma$	$G < -\sigma$
$L < -\sigma$			
$ L  \leq \sigma$			
$L > \sigma$			

Figure II-16 : correspondance entre 2-formes et gradient (G)-Laplacien (L)

### VI-3- APPLICATIONS DES K-FORMES POUR LA DETECTION DE CONTOURS

Toute bi-partition judicieuse de l'ensemble des  $3^k$  K-formes permet de définir un type de détection de contours. A partir de l'image brute, convenablement lissée, on tire l'image des K-formes. On attribue ensuite à la partition contenant les K-formes correspondant à un contour la couleur blanche et à l'autre la couleur noire par exemple. Par simple manipulation de la table de transfert sur l'image des K-formes on obtient donc l'image des contours. Cette méthode doit surtout être utilisée pour choisir au mieux la bi-partition des K-formes.

## VII - EROSION ET DILATATION D'UNE IMAGE BINAIRE [25][26]

### VII-1- INTRODUCTION

L'image simplifiée peut être résumée par un ensemble de points qui l'échantillonne, chaque point ayant deux valeurs conventionnelles possibles.

Cette image peut être l'image de départ de l'étude ou l'image intermédiaire succédant à plusieurs étapes ayant transformé une image couleur ou une image à niveaux de gris en une image binaire.

L'image simplifiée suppose que l'on a su isoler deux "objets" ou deux "thèmes" ou deux "phases" dans l'image et que l'on se propose d'étudier l'un et l'autre.

Les études d'images binaires sont très nombreuses. Ce paragraphe décrit deux opérations élémentaires, l'érosion et la dilatation et montre leurs propriétés quant aux "effets" et aux "mesures" possibles sur l'image pour la caractériser et la quantifier.

L'érosion et la dilatation sont deux opérations élémentaires fondamentales de la morphologie mathématique. Elles sont à la base de nombreuses méthodes d'analyses.

## VII-2- L'EROSION

Cette opération consiste à enlever une couche d'épaisseur constante au contour de l'ensemble étudié (figure II-17),

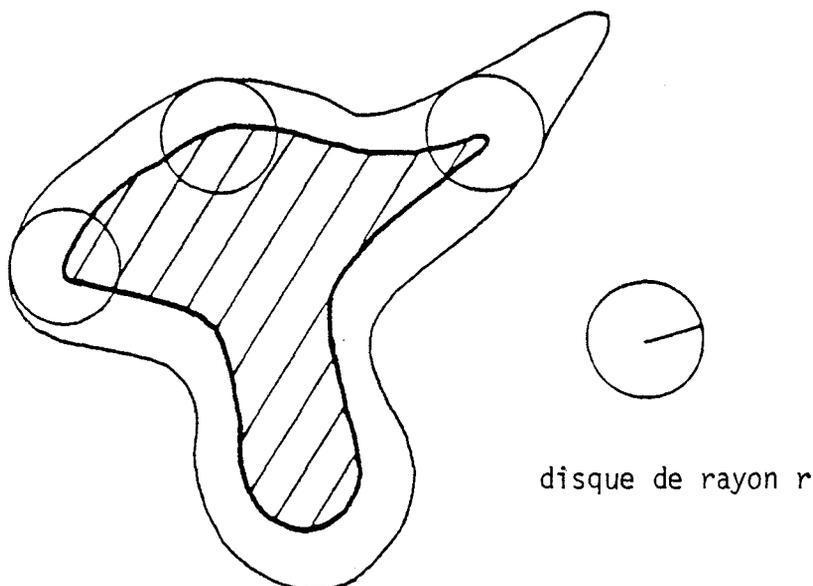
Définition : Soit A un ensemble. L'érodé de A par un disque de rayon r est l'ensemble de tous points de A, centres d'un disque de rayon r totalement inclus dans A.

Notation : l'érodé de A par un disque de rayon r est noté :  $A \ominus B_r$ .

Exemple :

### Méthode de construction

De l'ensemble A on a conservé les points centres d'un disque de rayon r inclus dans A.



Ensemble A

Erodé de l'ensemble A

Figure II-17

Propriété : L'érodé de A par un disque de rayon r, lui-même érodé par un disque de rayon r', est aussi l'érodé de A par un disque de rayon r + r'

$$(A \ominus B_r) \ominus B_{r'} = A \ominus B_{(r+r')}$$

### VII-3- LA DILATATION

Cette opération est l'inverse de la précédente. Elle consiste à rajouter une épaisseur constante au contour de l'ensemble à étudier (figure II-18).

Définition : Soit A un ensemble, le dilaté de A par un disque de rayon r est l'ensemble de tous les points centres d'un disque de rayon r rencontrant A.

Notation : Le dilaté de A par un disque de rayon r est noté :  $A \oplus B_r$

#### Exemples

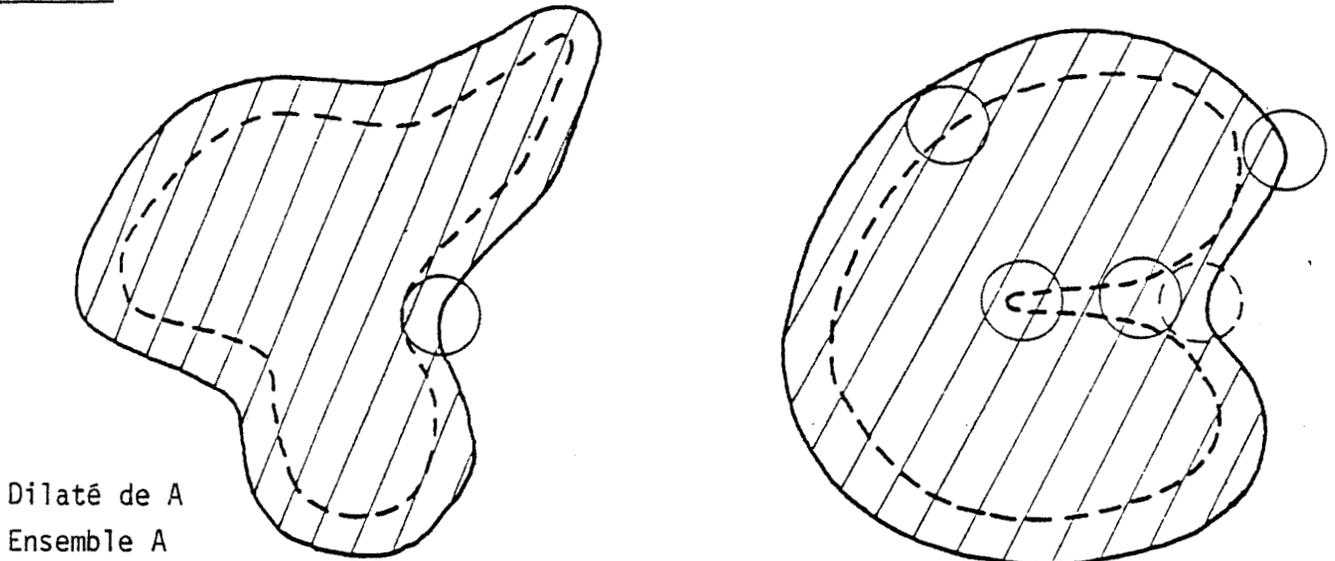


Figure II-18

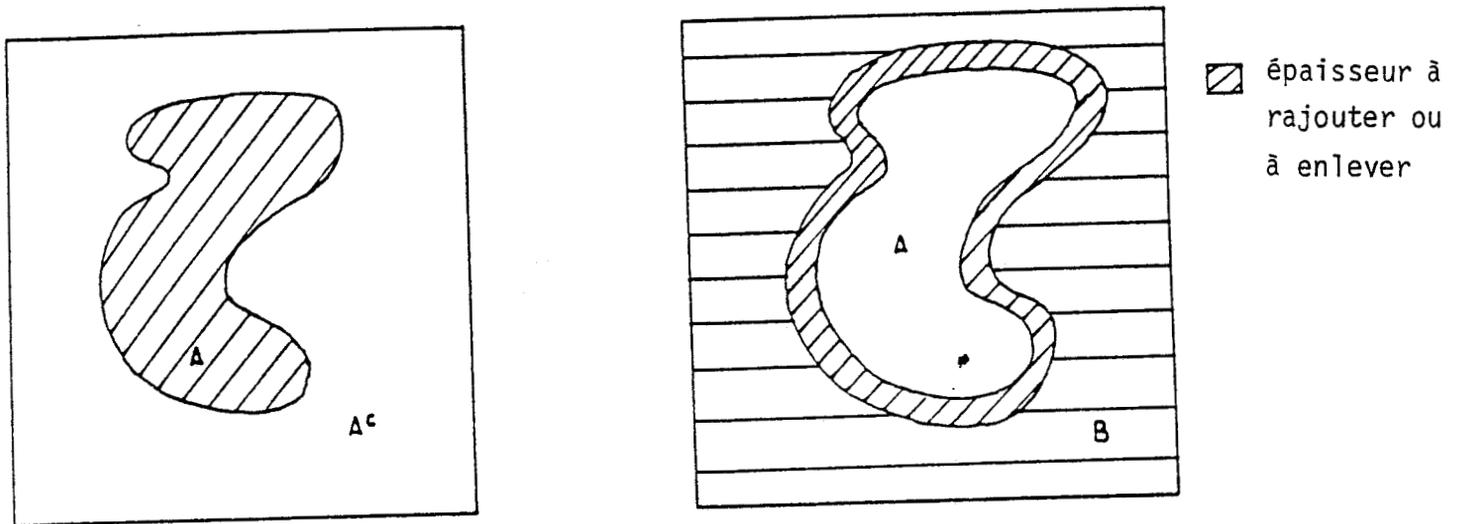
Sur l'exemple de la figure II-18, on peut constater que l'érosion de taille r fait disparaître les pointes d'épaisseur inférieure à 2 r, alors que la dilatation de taille r comble les trous d'épaisseur inférieure à 2 r.

Propriété : Le dilaté de A par un disque de rayon r, lui-même dilaté par un disque r', est le dilaté de A par un disque de rayon r+r'

$$(A \oplus B_r) \oplus B_{r'} = A \oplus B_{(r+r')}$$

VII-4- LIEN ENTRE EROSION ET DILATATION

Une érosion consiste à enlever une couche d'épaisseur constante alors qu'une dilatation consiste à en rajouter une. On conçoit que l'on peut passer facilement d'une opération à l'autre. Ainsi, pour rajouter une couche d'épaisseur constante, il suffit d'enlever la même épaisseur à son complémentaire et de prendre le complémentaire du nouvel ensemble ainsi déterminé. Réciproquement pour rajouter une couche d'épaisseur constante à un ensemble, il suffit d'enlever une même épaisseur à la partie complémentaire (figure II-19).



A union la partie hachurée - représente le dilaté de A  
 B union la partie hachurée - représente le complémentaire de A :  $A^c$   
 B seul, représente l'érodé du complémentaire de A.

Figure II-19

Propriété : Soit A un ensemble,  $A^c$  son complémentaire, on a les propriétés (figure II-20) :

$$(A \oplus B_r)^c = A^c \ominus B_r$$

$$(A \ominus B_r)^c = A^c \oplus B_r$$

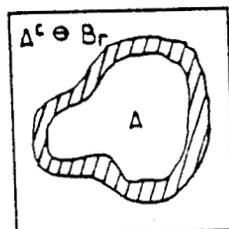


Figure II-20

VII-5- L'EROSION ET LA DILATATION FACE A L'UNION ET L'INTERSECTION

VII-5-1- L'érosion

L'érodé d'une intersection est bien l'intersection des deux érodés (figure II-21).

$$(X_1 \cap X_2) \ominus B_r = (X_1 \ominus B_r) \cap (X_2 \ominus B_r)$$

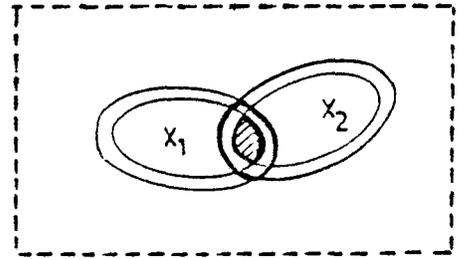


Figure II-21

Mais l'érodé d'union peut être strictement plus grand que l'union des érodés.

$$(X_1 \cup X_2) \ominus B_r \supset (X_1 \ominus B_r) \cup (X_2 \ominus B_r)$$

VII-5-2- La dilatation

Par passage au complémentaire, on montre de même que la dilatation se comporte bien face à l'union, mais ne donne pas de bons résultats face à l'intersection :

$$(X_1 \cup X_2) \oplus B_r = (X_1 \oplus B_r) \cup (X_2 \oplus B_r)$$

$$(X_1 \cap X_2) \oplus B_r \subset (X_1 \oplus B_r) \cap (X_2 \oplus B_r)$$

VII-6- OUVERTURE ET FERMETURE

Nous avons vu que :  $(A \oplus B_r) \oplus B_{r'} = A \oplus B_{(r+r')}$

$$(A \ominus B_r) \ominus B_{r'} = A \ominus B_{(r+r')}$$

Par contre :  $(A \oplus B_r) \ominus B_r \neq A$

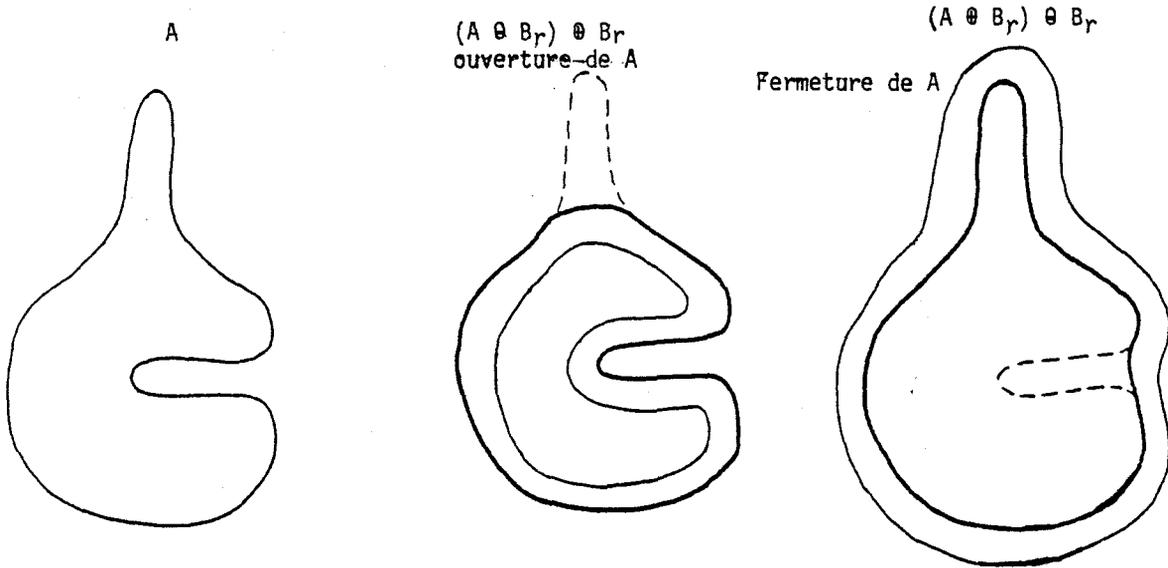


Figure II-22

On voit donc que les résultats obtenus sont différents de A et que d'une façon générale les contours de A sont régularisés.

Définition : La première opération qui consiste à faire une érosion suivie d'une dilatation s'appelle une ouverture.

La seconde qui consiste à faire une dilatation suivie d'une érosion s'appelle fermeture.

Propriété : Les opérations d'ouverture et de fermeture d'un ensemble sont idempotentes. L'ouvert de l'ouvert (respectivement le fermé du fermé) d'un ensemble est l'ouvert (respectivement le fermé) de cet ensemble.

Cette propriété est indispensable pour la reconnaissance des formes puisqu'elle assure une certaine stabilité.

Ces deux résultats sont en pratique largement utilisés, car leur résultat est représentatif de l'ensemble, et que d'autre part il régularise ses contours et permet donc un meilleur traitement ultérieur de l'image.

Elles constituent une première application des deux opérations fondamentales : dilatation et érosion.

## VIII - CONCLUSION

Dans ce paragraphe, nous avons décrit quelques méthodes de traitement d'images (lissage, détection de contour, filtrage binaire ...) pour pouvoir fournir une approche théorique au problème de détection du relief Braille.

Dans la première partie du chapitre III, nous allons présenter une méthode de reconnaissance issue des algorithmes cités ci-dessus.

Cette méthode n'étant pas bien adaptée à la détection du relief, nous en mettons une autre en oeuvre, ce qui va permettre sa reconnaissance sans ambiguïté.

CHAPITRE III : DÉTECTION DU RELIEF BRAILLE

## I - INTRODUCTION

Dans ce chapitre, nous nous intéressons à la reconnaissance du relief des points Braille. La première analyse de l'image effectuée au chapitre I a montré que dans une première phase il faut séparer les objets du fond. Un objet est défini comme une tâche connexe claire sur fond sombre.

On rappelle que les images comportent :

- des objets de petite taille et faiblement contrastés,
- des objets fins et très faiblement contrastés,
- des objets proches, ce qui peut produire des interférences de luminosité,
- des objets de formes complexes.

Dans ce domaine très proche de détection de contour, il existe un grand nombre de méthodes dont une partie a été citée au chapitre II. Cependant, la plupart des objets se distinguent essentiellement par la différence du niveau de gris des points qui les composent. Pour assurer leur extraction, nous avons été amenés à privilégier les opérateurs de détection de contour sensible à la variation du niveau de gris moyen (Gradient, Laplacien, Sobel, etc..., travaillant sur un voisinage d'une dizaine à quelques dizaines de points autour du point étudié) qui font appel aux techniques de filtrage linéaire bidimensionnel<sup>[27][28]</sup>.

Les opérateurs que nous présentons sont définis sur des voisinages (1x5), (3x3) ou (5x5) d'un point d'une image. Pour la plupart des calculs, nous avons utilisé la notion de vecteur image. Par exemple, pour un opérateur (3x3), le vecteur image  $f_m$  a pour composantes les valeurs des intensités des neuf points de la fenêtre centrée en  $p$  (figure III-1).

Tout opérateur linéaire est défini par un vecteur constant  $\theta$ . Le résultat de l'application d'un tel opérateur correspond à la valeur du produit scalaire  $\langle f_m, \theta \rangle$  qui est attribuée au point central de la fenêtre image.

y \ x	-1	0	1
-1	$f(-1,-1)$	$f(0,-1)$	$f(1,-1)$
0	$f(-1,0)$	$f(0,0)$	$f(1,0)$
1	$f(-1,1)$	$f(0,1)$	$f(1,1)$

$$\begin{pmatrix} f(-1,-1) & f(0,-1) & f(1,-1) \\ f(-1,0) & f(0,0) & f(1,0) \\ f(-1,1) & f(0,1) & f(1,1) \end{pmatrix}$$

Fenêtre image

Vecteur image correspondant

$$f_m = (f(-1,-1), f(-1,0), \dots, f(1,1))^T$$

$$\begin{pmatrix} \theta(-1,-1) & \theta(0,-1) & \theta(1,-1) \\ \theta(-1,0) & \theta(0,0) & \theta(1,0) \\ \theta(-1,1) & \theta(0,1) & \theta(1,1) \end{pmatrix}$$

Masque opérateur

Vecteur opérateur

$$\theta = (\theta(-1,-1), \theta(-1,0), \dots, \theta(1,1))^T$$

Figure III-1

La valeur fournie par l'opérateur est :

$$h = \sum_x \sum_y f(x,y) \theta(x,y)$$

L'image filtrée est obtenue en répétant la même opération pour chaque point de l'image, sauf pour ceux du bord.

Dans ce chapitre, nous mettons également en oeuvre quelques filtres agissant sur l'image seuillée (image binaire), ce qui nous permet de reconnaître le relief Braille sans ambiguïté.

PREMIÈRE PARTIE DU CHAPITRE III : PREMIÈRE MÉTHODE DE DETECTION  
DU RELIEF.

## II - LE PRETRAITEMENT DE L'IMAGE EN RELIEF BRAILLE

Les filtres passe-bas utilisés pour lisser l'image afin de réduire l'effet du bruit sont en général les suivants<sup>[33]</sup>:

$$\Theta_1 = \frac{1}{9} \begin{vmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{vmatrix} \quad \Theta_2 = \frac{1}{10} \begin{vmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 1 \end{vmatrix} \quad \Theta_3 = \frac{1}{16} \begin{vmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{vmatrix}$$

Ces filtres "moyenneurs" ne sont pas adaptés à la forme du signal représentant le relief Braille. En effet, les figures III-2, III-3 et III-4 montrent que les filtres  $\Theta_1$ ,  $\Theta_2$ ,  $\Theta_3$  atténuent efficacement le bruit de l'image, mais qu'ils lissent fortement les contours des objets, ce qui va rendre plus difficile l'extraction du relief.

En effet, si l'on veut éliminer le bruit, il faut limiter les hautes fréquences et donc dégrader un peu plus l'information haute fréquence provenant de l'objet. Inversement, lorsque l'on veut compenser la dégradation introduite par le système d'imagerie, il faut amplifier essentiellement les hautes fréquences. Le bruit est alors amplifié simultanément. Malheureusement, ces deux phénomènes sont opposés et on voit à quel point les notions de contraste-bruit-résolution spatiale sont imbriquées.

Si on remplace chaque point de l'image par la combinaison linéaire suivante :

$$g(x,y) = \frac{1}{3} \left[ f(x,y-1) + f(x,y) + f(x,y+1) \right]$$

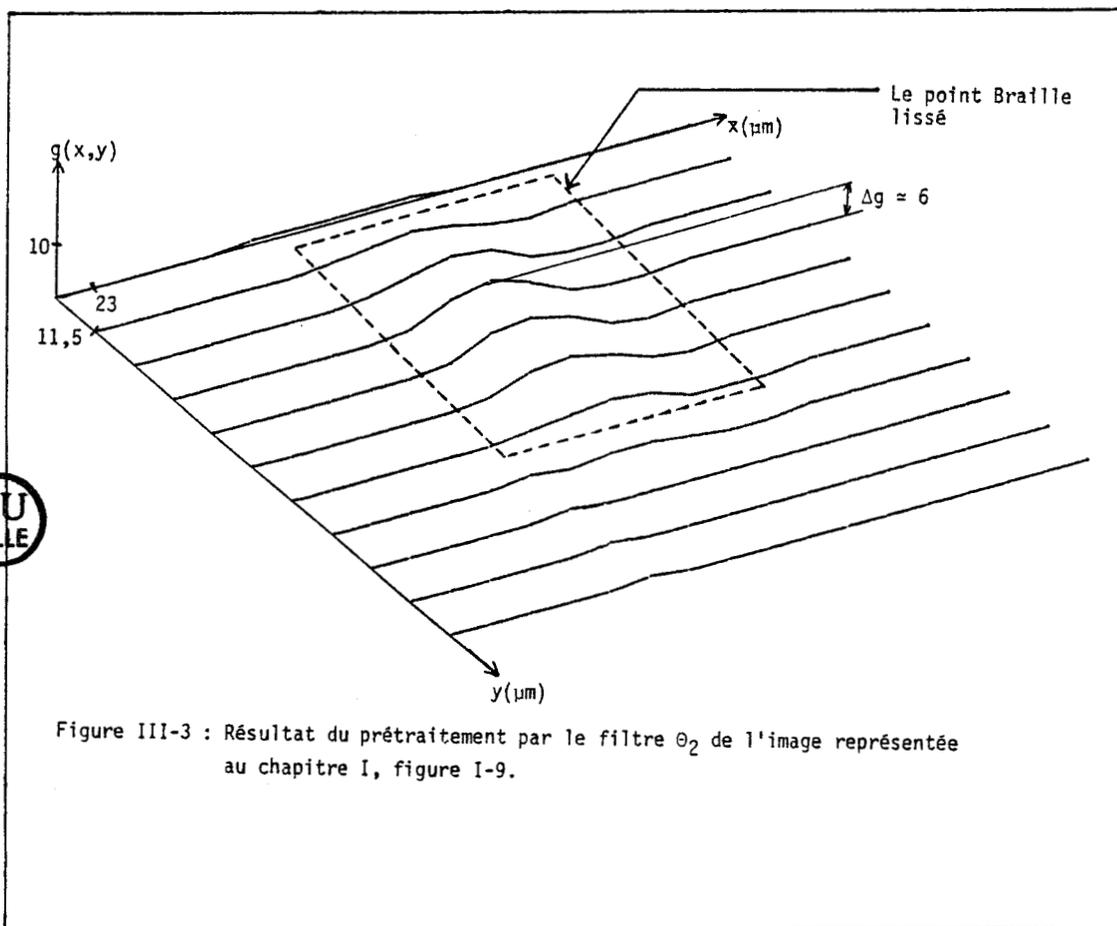
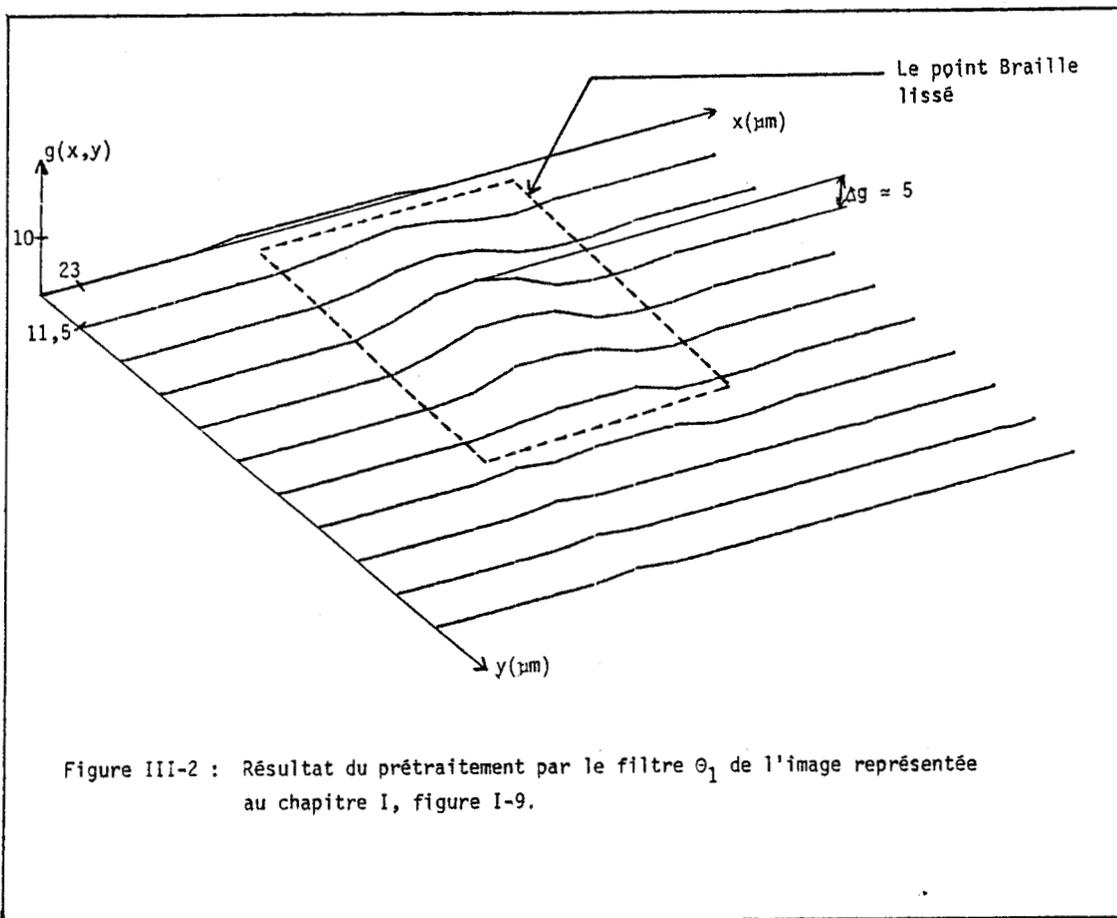
$$y = -\frac{N}{2} + 1, -\frac{N}{2} + 2, \dots, 0, \dots, \frac{N}{2} - 1$$

et

$$\begin{cases} g(x, -\frac{N}{2}) = f(x, -\frac{N}{2}) \\ g(x, \frac{N}{2}) = f(x, \frac{N}{2}) \end{cases}$$

On obtient un filtre moyeneur "vertical" défini par le masque opérateur  $\Theta_4$  :

$$\Theta_4 = \frac{1}{3} \begin{vmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{vmatrix}$$



Ce filtre adoucit les contours des objets, ce qui met en valeur l'information à extraire (il améliore la forme de l'objet), mais il atténue faiblement le bruit (figure III-5). Dans un premier temps, nous retenons le filtre  $\Theta_4$  pour le prétraitement de l'image en relief Braille.

### III - SEGMENTATION D'IMAGES

Lors de l'analyse de l'image, il est nécessaire d'avoir une bonne segmentation du support afin de ne pas fausser le calcul des paramètres qui parfois dépendent beaucoup de l'état de l'image segmentée tels les paramètres de formes basés sur la frontière de l'objet.

La segmentation peut être vue sous différents aspects (statistique ou déterministe) mais elle demeure basée sur deux notions essentielles :

- Définition de classes (critère d'uniformité de niveau de gris, critère de détection de points frontières...).
- Regroupement des points d'une même classe (notion de contour, notion de région).

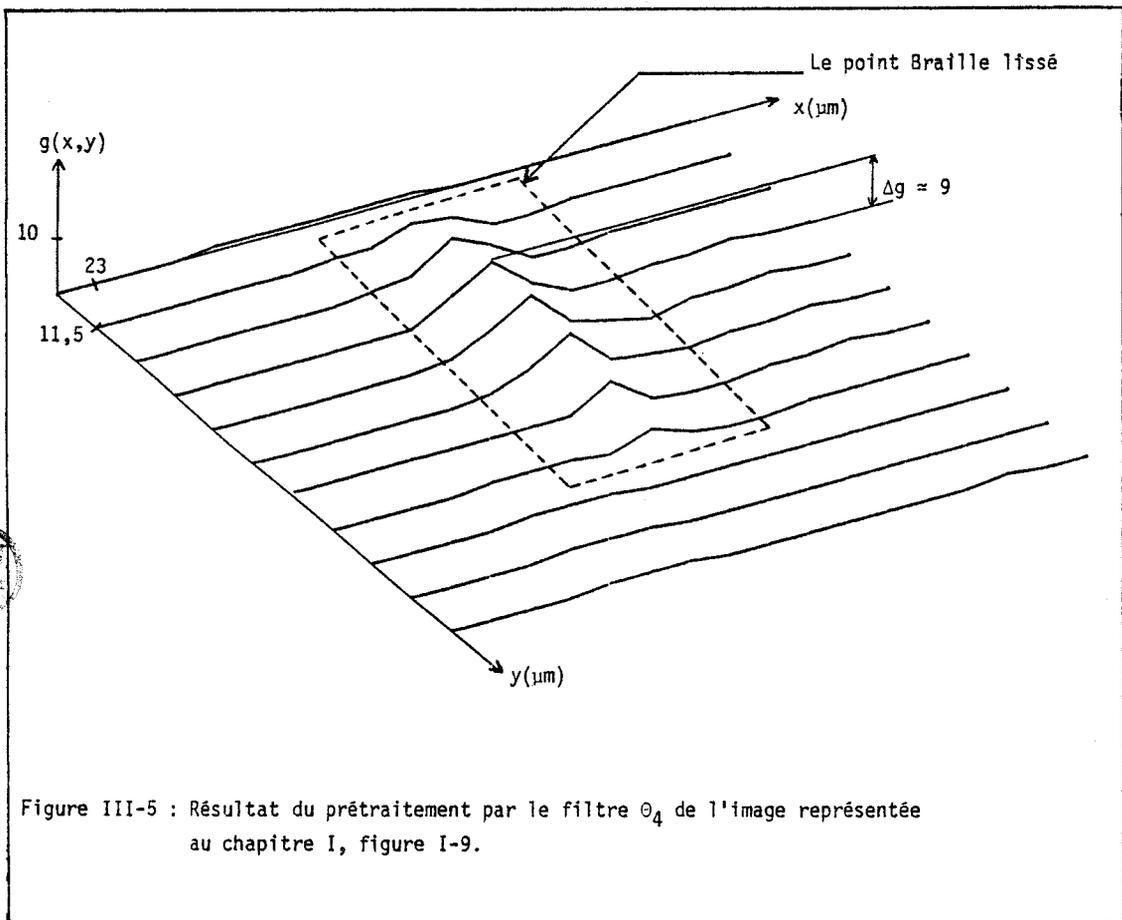
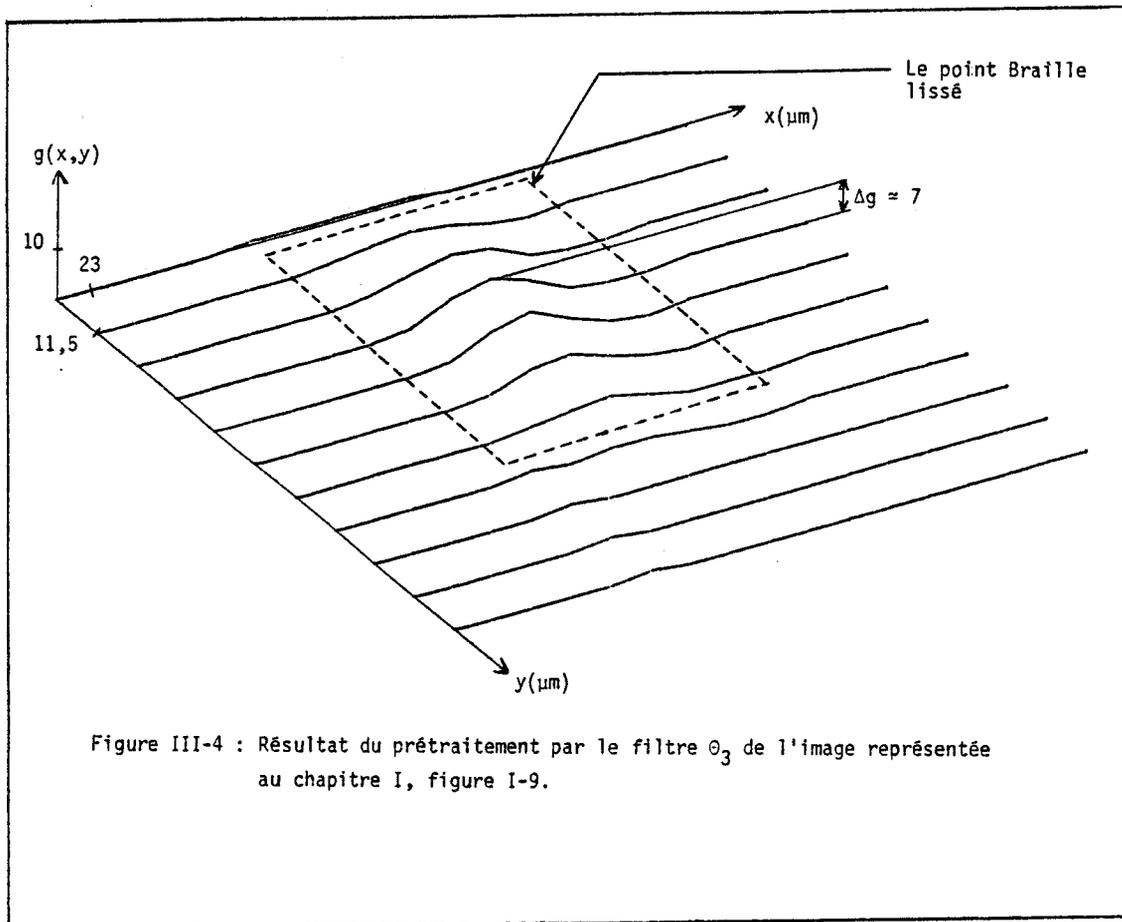
Plusieurs méthodes ont été définies au cours de recherches en analyse d'images. Il existe un lien étroit entre le problème abordé et la méthode employée. Ce lien peut être décrit par la nécessité d'utiliser des critères de couleur, de forme, de texture, ce choix étant étroitement lié au contexte présent dans l'image.

On peut distinguer deux façons d'aborder la segmentation : l'élément de contour et la construction de régions.

L'algorithme de régions prend naissance en un point de propriétés déterminées, par une recherche séquentielle de points similaires et de proche en proche, la région s'agrandit. Elle est entièrement délimitée lorsque des discontinuités de la caractéristique considérée apparaissent.

La segmentation par contours recherche des régions de discontinuités de niveaux de gris (puisque c'est la seule caractéristique que nous traitons) et on extrait par seuillage les éléments de contour significatifs.

C'est cette démarche qui va être suivie dans un premier temps.



IV - EXTRACTION D'ATTRIBUTS DU RELIEF BRAILLE

IV-1- RENFORCEMENT LINEAIRE DES CONTOURS DU RELIEF BRAILLE

Comme il a été dit au chapitre I, la localisation de points Braille se fera par renforcement après prétraitement de l'image. C'est une approche non optimale de détection de contour de type "gradient" qui a donnée les résultats les plus satisfaisant. Les détecteurs sont heuristiques, il n'y a pas de modèles associés et il est difficile de discuter de leur optimalité.

A titre d'illustration, la figure III-6 représente l'effet d'un opérateur gradient linéaire sur un élément de contour idéal.

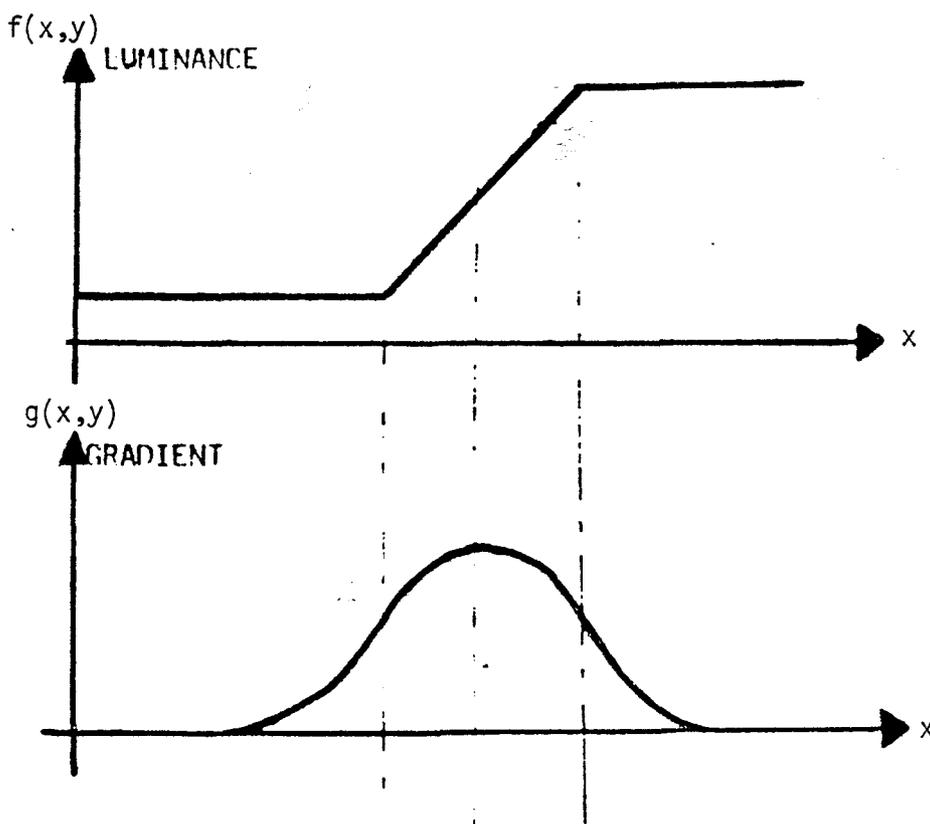


Figure III-6



IV-1-1- Renforcement linéaire des contours dans les directions orthogonales

L'objet représentant le point Braille se manifeste par une augmentation du contraste d'une surface se composant d'un certain nombre de pixels qui est fonction de la qualité du relief.

Comme le montre la figure III-7, différentes formes peuvent se présenter :

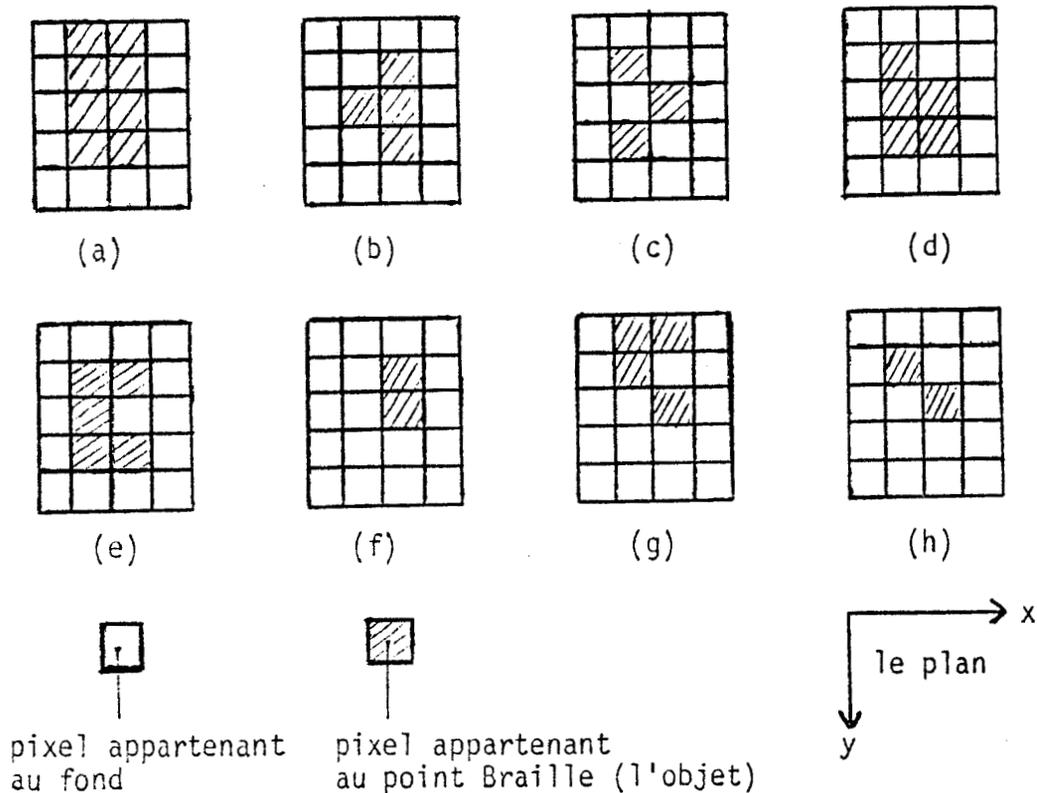


Figure III-7 : Représentation de quelques points Braille dans le plan discret.

Les lignes de contour sont mises en évidence par la convolution de l'image avec les masques opérateurs "vertical et horizontal" suivants<sup>[34]</sup>:

$$\frac{1}{6} \begin{vmatrix} -1 & 2 & -1 \\ -1 & 2 & -1 \\ -1 & 2 & -1 \end{vmatrix}$$

$$\frac{1}{6} \begin{vmatrix} -1 & -1 & -1 \\ 2 & 2 & 2 \\ -1 & -1 & -1 \end{vmatrix}$$

Nous avons convolué l'image prétraitée avec le filtre (3x3) vertical pour que les caractéristiques particulières de présence du relief avantagent l'extraction des primitives sous formes de segments.

Le bruit à dominante verticale s'amplifie et les résultats obtenus ne sont satisfaisants que pour quelques formes du relief. Ces reliefs sont ceux des figures III-7-a, b, d et e.

Un exemple de ce traitement est représenté par la figure III-8.

IV-1-2- Renforcement linéaire des contours par les compas directionnels

Dans le paragraphe précédent, nous avons supposé que les contours de l'image sont susceptibles d'exister dans deux directions orthogonales. Une autre approche de Freeman (chapitre II) consiste à diviser le plan en 8 directions auxquelles sont rattachés des masques de convolution (sensibles à la rotation, mais invariants à la translation : figure III-9)<sup>[35]</sup>

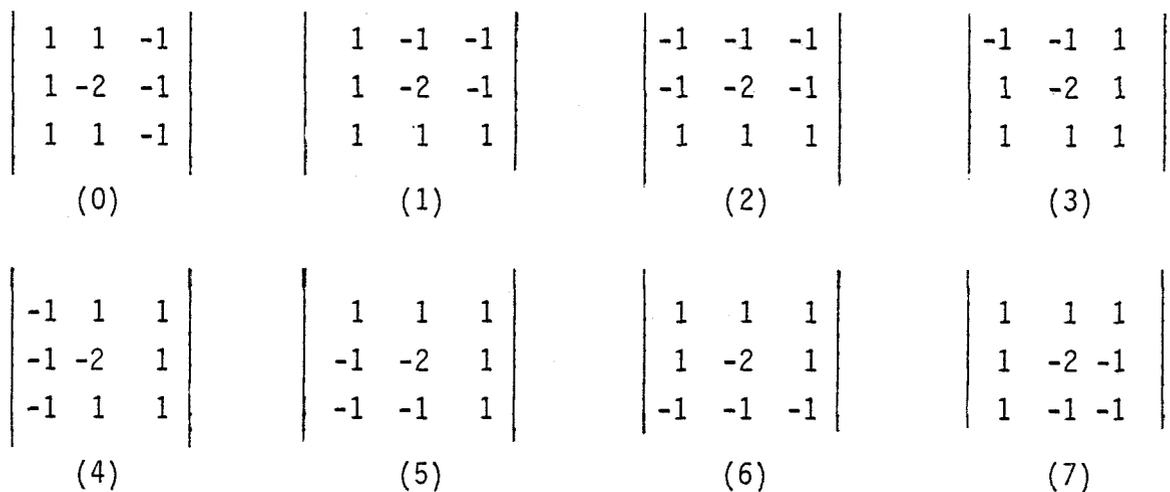
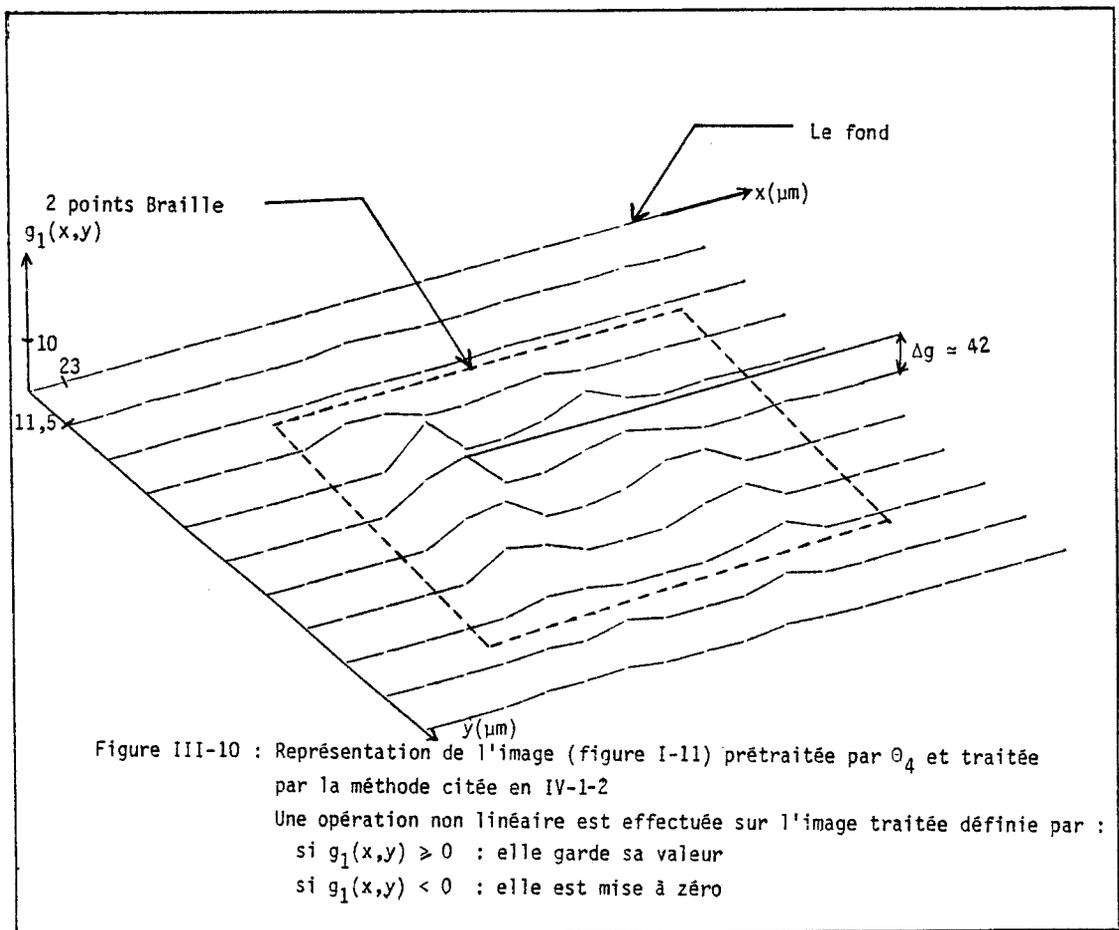
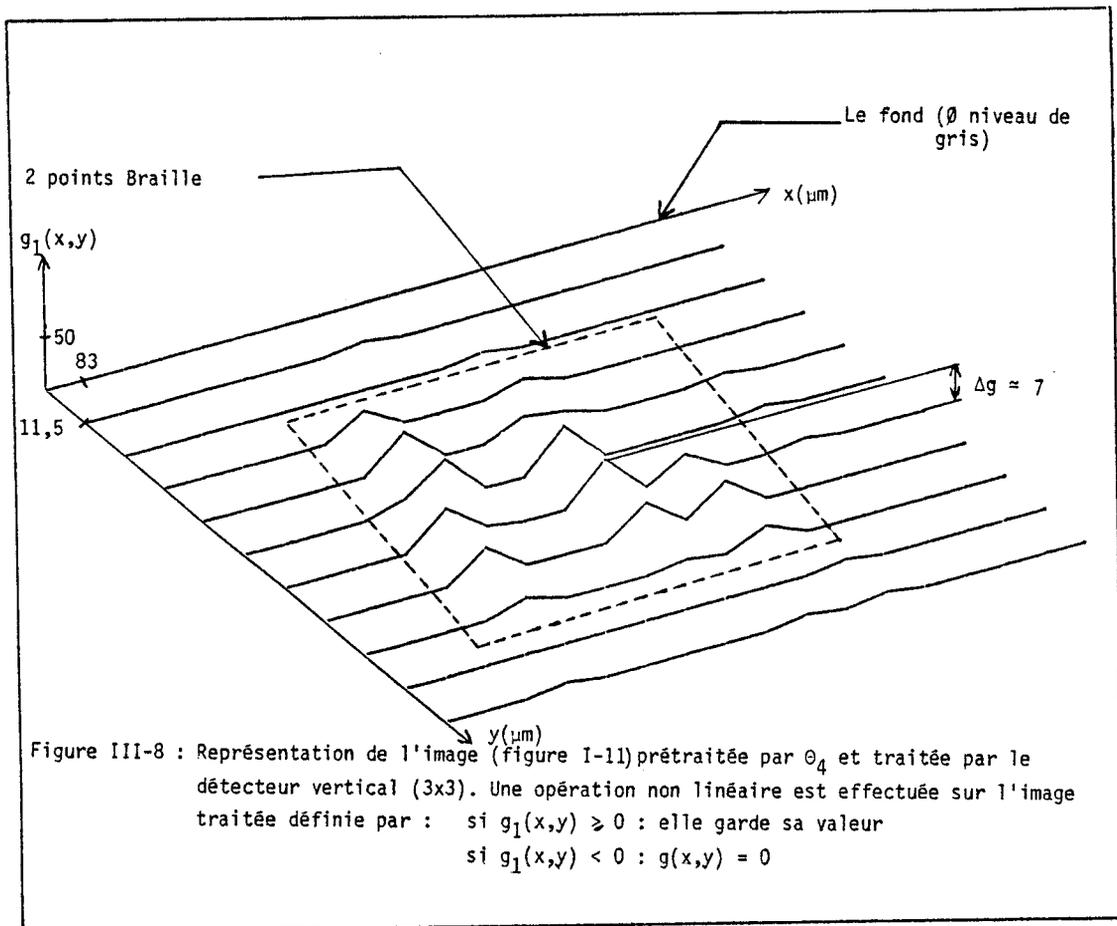


Figure III-9 : Compas ou gradients directionnels de Prewitt

La réponse de ces filtres est nulle pour les régions à luminance constante, on a :

$$\sum_x \sum_y g(x,y) \Theta(x,y) = g(x,y) \sum_x \sum_y \Theta(x,y) = 0$$

La figure III-10 représente un exemple du résultat obtenu par la convolution de l'image de la figure I-11 du chapitre I prétraitée par  $\Theta_4$  avec les deux masques de Prewitt suivant les directions 0 et 7 de Freeman.



Ce traitement a consisté à attribuer à chaque point de l'image l'amplitude maximale résultante de la convolution avec chacun des deux opérateurs de Prewitt 0 et 7.

Après avoir effectué par cette démarche un nombre important de mesures, nous nous sommes aperçus que les conditions de mesures (éclairage, nature de la fabrication du relief ...) privilégient deux directions.

Les masques qui ont donné le maximum de réponse sont ceux de Prewitt suivant les directions 0 et 7 de Freeman.

## V - LA BINARISATION

### V-1- LE SEUILLAGE

Le seuillage est une opération non linéaire. Elle permet de ne retenir parmi les zones de transitions mises en évidence par le filtrage que celles qui correspondent à un contour recherché. Le choix du seuil de luminance  $S$  dans l'image filtré peut être obtenu automatiquement en utilisant les statistiques du bruit présent dans l'image d'origine ( $S = N \sigma_b$ , le facteur constant  $N$  est défini par avance expérimentalement sur un grand nombre d'images et  $\sigma_b$  est estimé lors du traitement en fonction des connaissances à priori de la caméra et son réglage de gain), soit déterminé par la méthode d'histogramme pour les images présentant des objets nettement discernables (calcul du seuil par maximum de vraisemblance), soit défini d'une façon déterministe.

Les points Braille peuvent être représentés par des contours monodimensionnels suivant les lignes horizontales (Exemple de contours : figure III-11).

L'amplitude de la luminance représentant les extrémités des deux premiers segments appartenant aux contours du relief Braille est définie par :

$$h_1 = g_1(x,y)$$

$$h_2 = g_1(x+1,y)$$

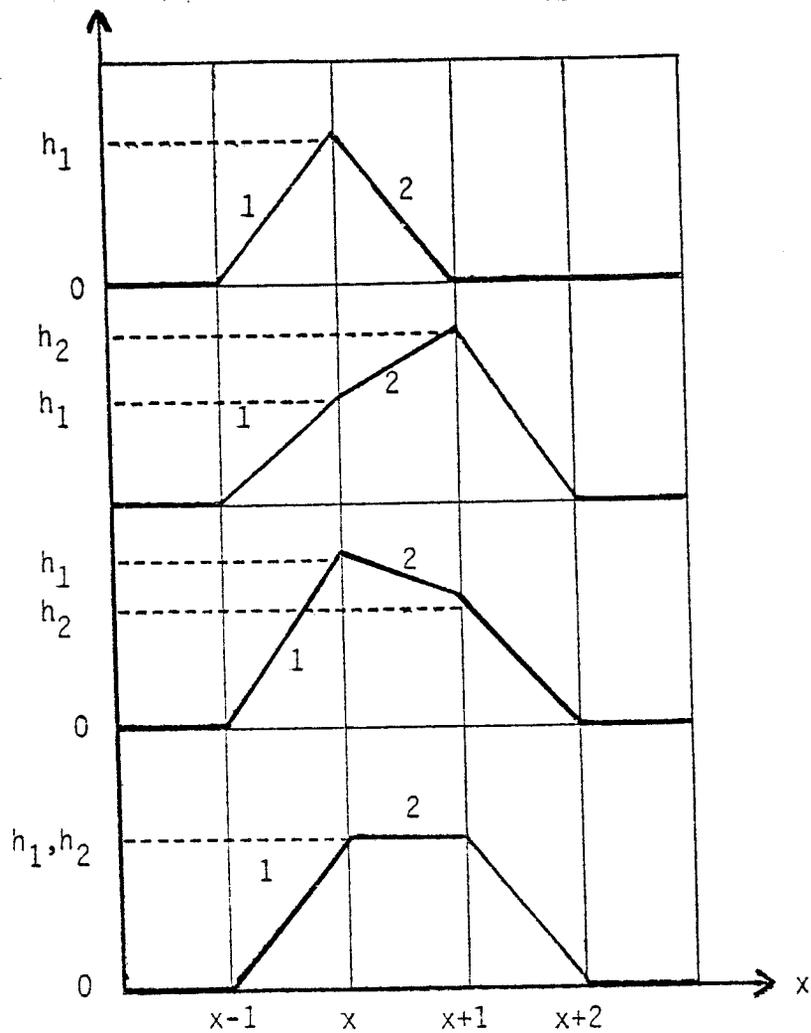


Figure III-11

Le seuillage de l'image est obtenu par :

$$c(x,y) = T [g_1(x,y)]$$

$$\text{avec } T = \begin{cases} c(x,y) = 0 & , \text{ si } h_1 + h_2 < S \\ c(x,y) = 1 & , \text{ si } h_1 + h_2 \geq S \end{cases}$$

Par exemple, la figure III-12 représente  $c(x,y)$  ou  $g_1(x,y)$  seuillée de l'image représentée par la figure III-10 avec un seuil choisi de façon déterministe égal à 35. Cette méthode de seuillage utilisant le principe de cumul d'observations successives<sup>[36]</sup> (méthode séquentielle de détection de contour) permet la détection des objets très faiblement contrastés et l'élimination des fausses alarmes principales.

V-2- LE FILTRAGE BINAIRE

L'image binaire, c'est-à-dire l'image précédemment seuillée, peut contenir outre les formes utiles, trois types de défauts essentiellement dus aux perturbations du bruit

- des coupures dans les formes (exemple : figure III-13-(a))
- des fausses alarmes, apparaissant souvent sous formes de bruit impulsionnel (exemple : figure III-13-(b))
- des interférences de luminosité entre deux points Braille appartenant au même graphème suivant les lignes horizontales (exemple : fig. III-13-(c))

Pour remédier à ces défauts, nous avons appliqué à l'image binaire deux filtres de lissage, ce qui permet la reconnaissance du relief sans ambiguïté.

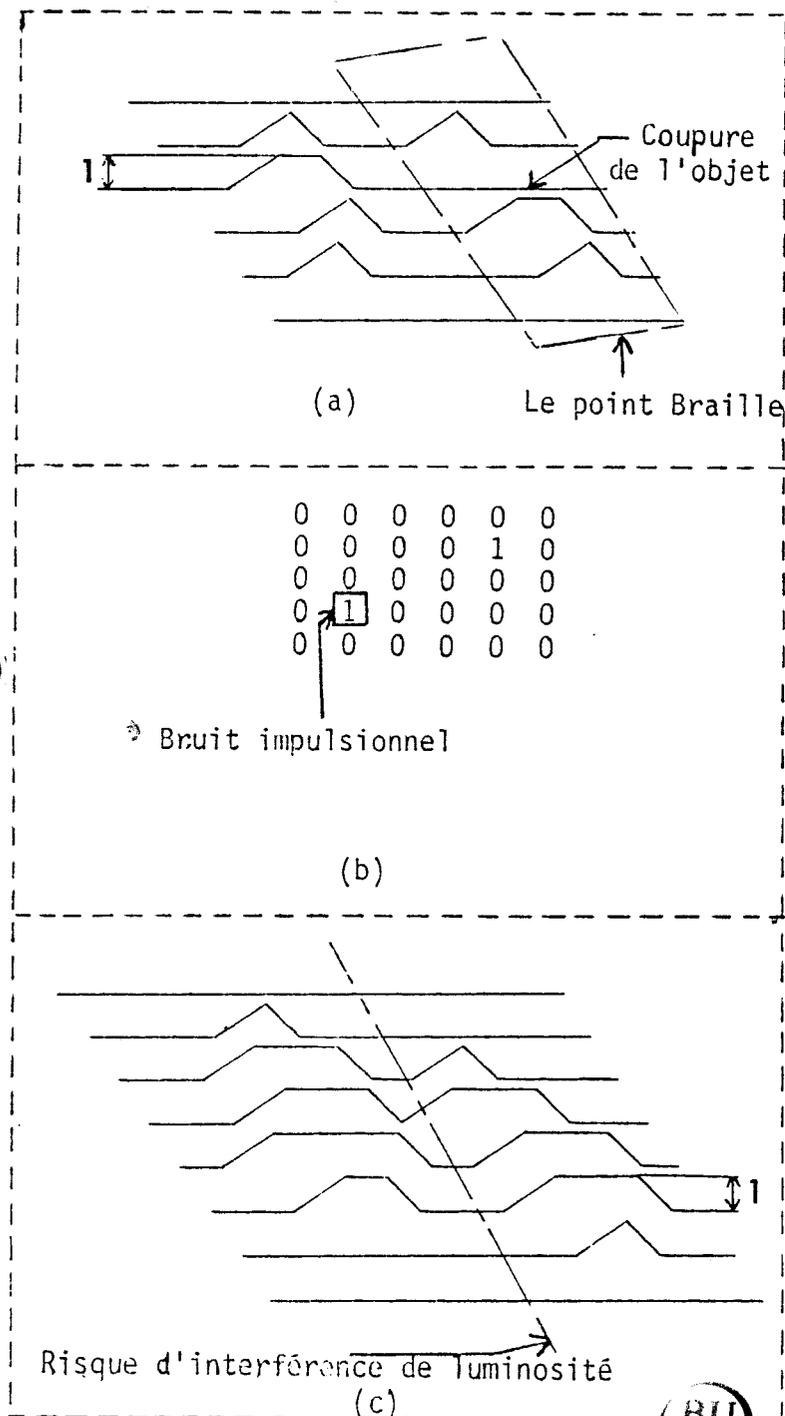


Figure III-13



V-3- LE LISSAGE DE L'IMAGE BINAIRE

Pour le lissage binaire, deux filtres ont été retenus :

- un filtre médian monodimensionnel agissant verticalement sur un voisinage de trois points. Il élimine le bruit impulsionnel et lisse les contours "spéciaux" du relief Braille.

La figure III-14 représente l'effet du filtre médian monodimensionnel 3 points sur l'image de la figure III-13-(a) et la figure III-15 représente cet effet sur l'image III-12. On constate que ce filtre effectue un alignement le long d'une direction parallèle à oy. Par contre, on voit sur la figure III-13-(a) qu'il peut provoquer un désalignement suivant une direction parallèle à ox.

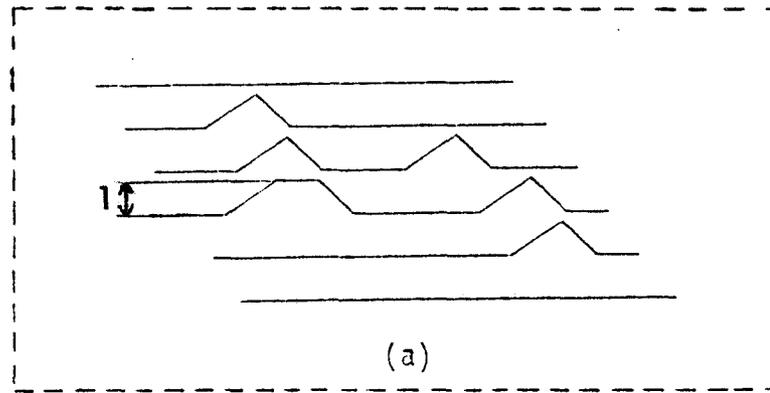


Figure III-14 : Image de la figure III-13-(a) lissée par le filtre médian

Ce désalignement n'est pas gênant car il existe une zone de recouvrement des reliefs de deux points adjacents d'un même graphème.

Une dilatation qui constitue le deuxième filtre retenu est susceptible suivant le module choisi de fermer les trous, faciliter la reconnaissance et éviter les interférences de luminosité des points Braille.

La dilatation revient, pour chaque point de l'image binarisée de valeur  $P \in \{0,1\}$  à satisfaire la relation booléenne :

$$P : = P + \text{voisinage} \quad \text{avec voisinage} = \text{somme booléenne des valeurs des points appartenant au module de définition du voisinage.}$$

Dans le cas d'une image représentant des points Braille, la dilatation ferme les trous et évite les interférences de luminosité entre les points lorsque le module est le premier des trois modules représentés figure III-16-(a).

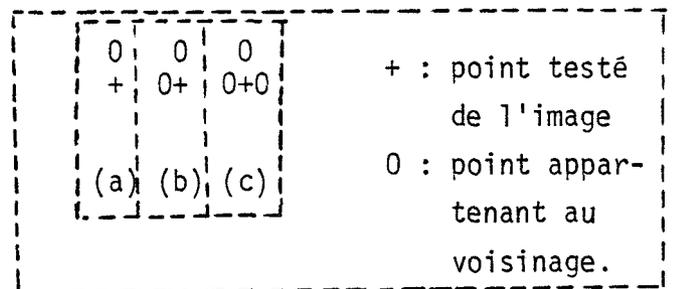


Figure III-16

La figure III-17-b,c,d illustre l'effet des 3 modules de dilatation appliqués au graphème figure III-17-a composé de 4 points résultant des prétraitements précédents.

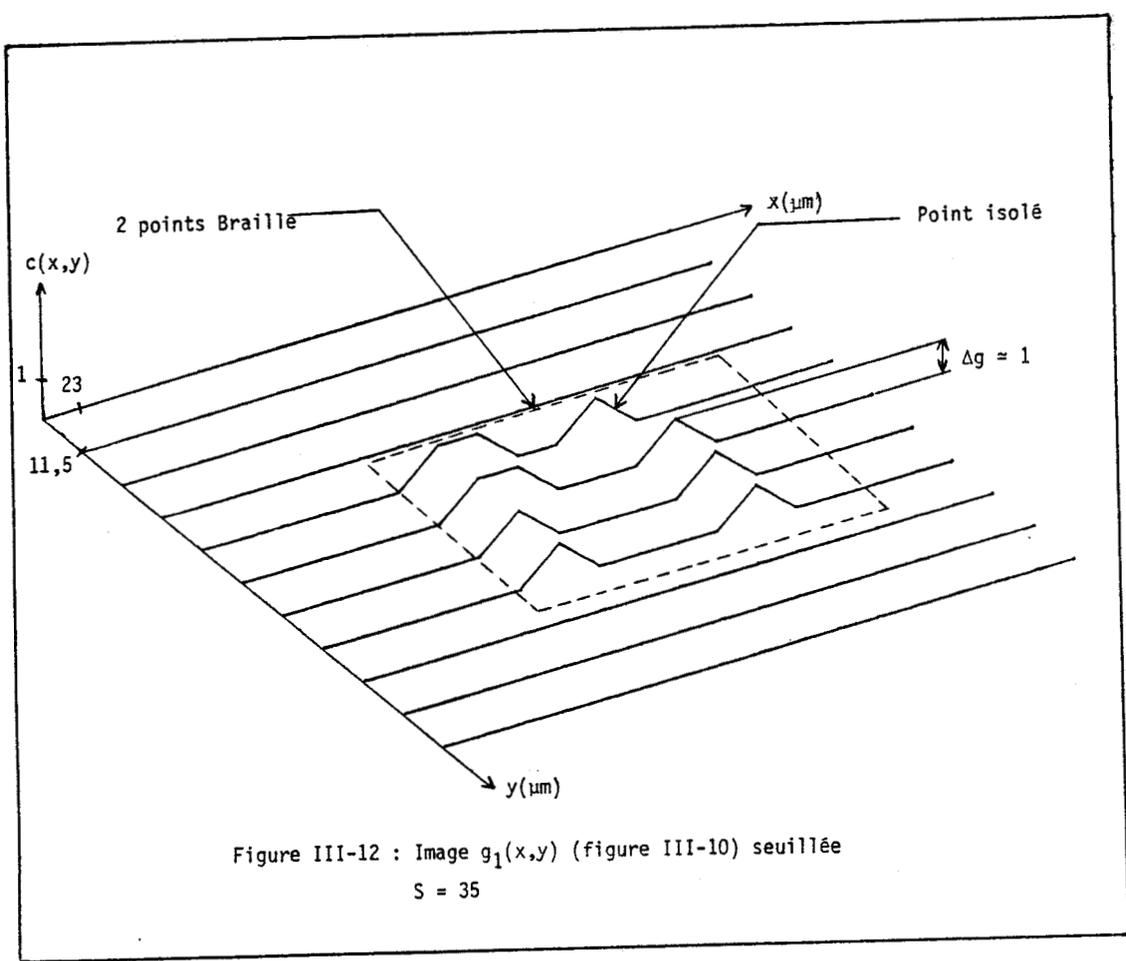


Figure III-12 : Image  $g_1(x,y)$  (figure III-10) seuillée  
S = 35

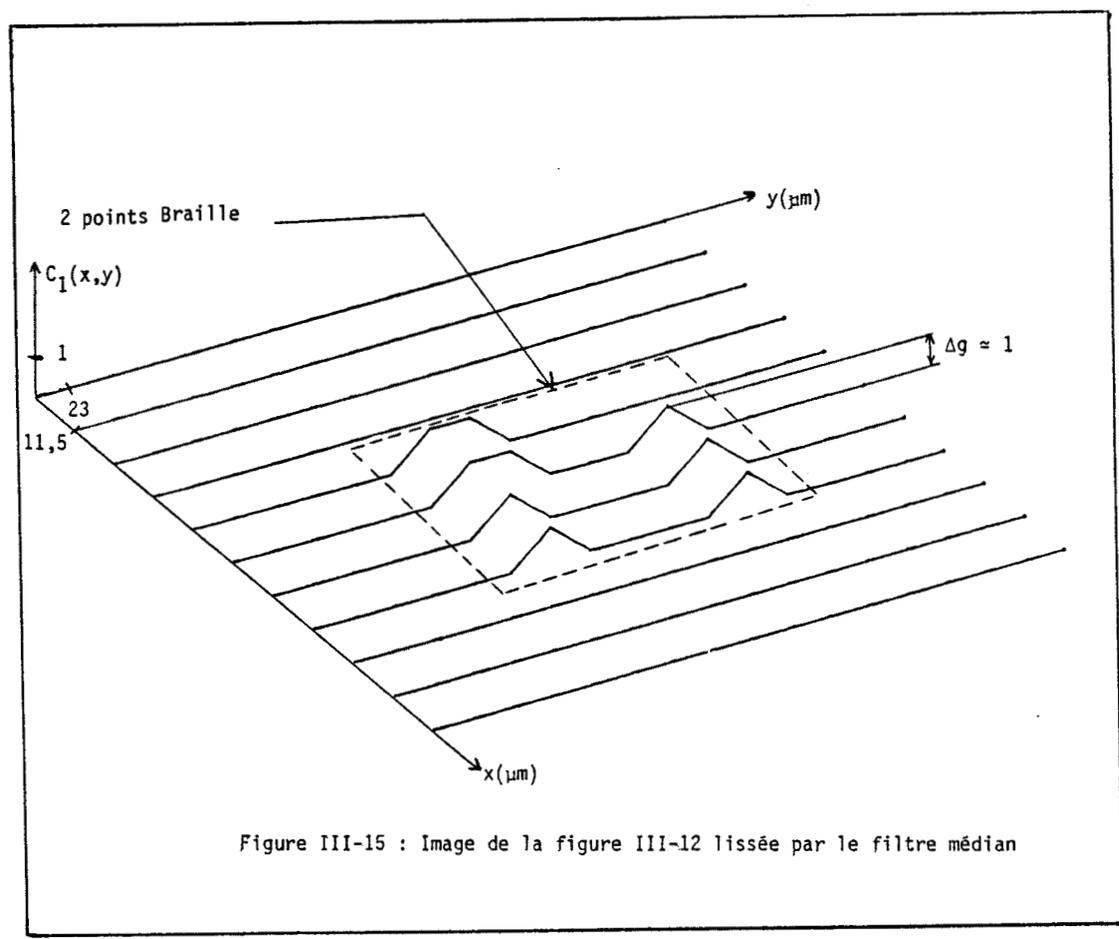
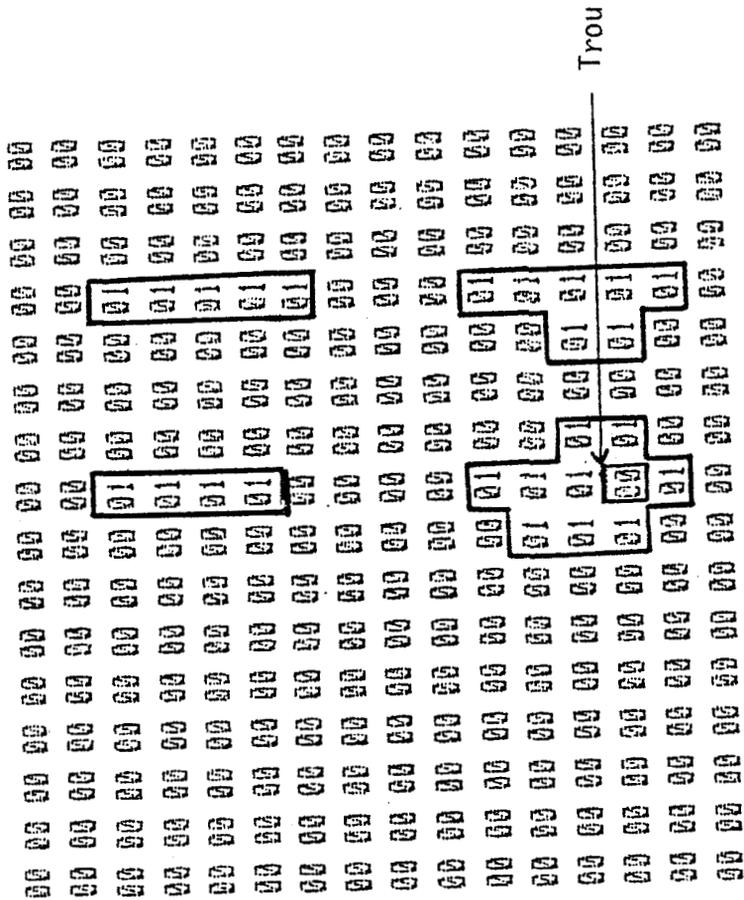


Figure III-15 : Image de la figure III-12 lissée par le filtre médian





Le caractère Braille à dilater

(a)

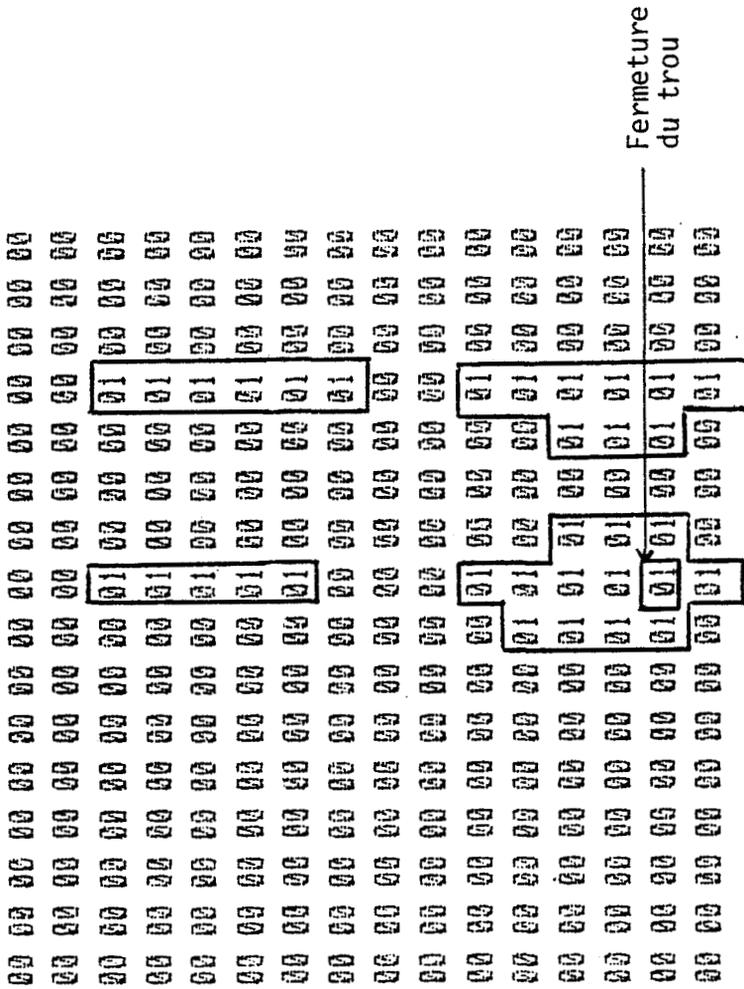


Image (a) dilatée  
Module de dilatation : x

(b)

Figure III-17

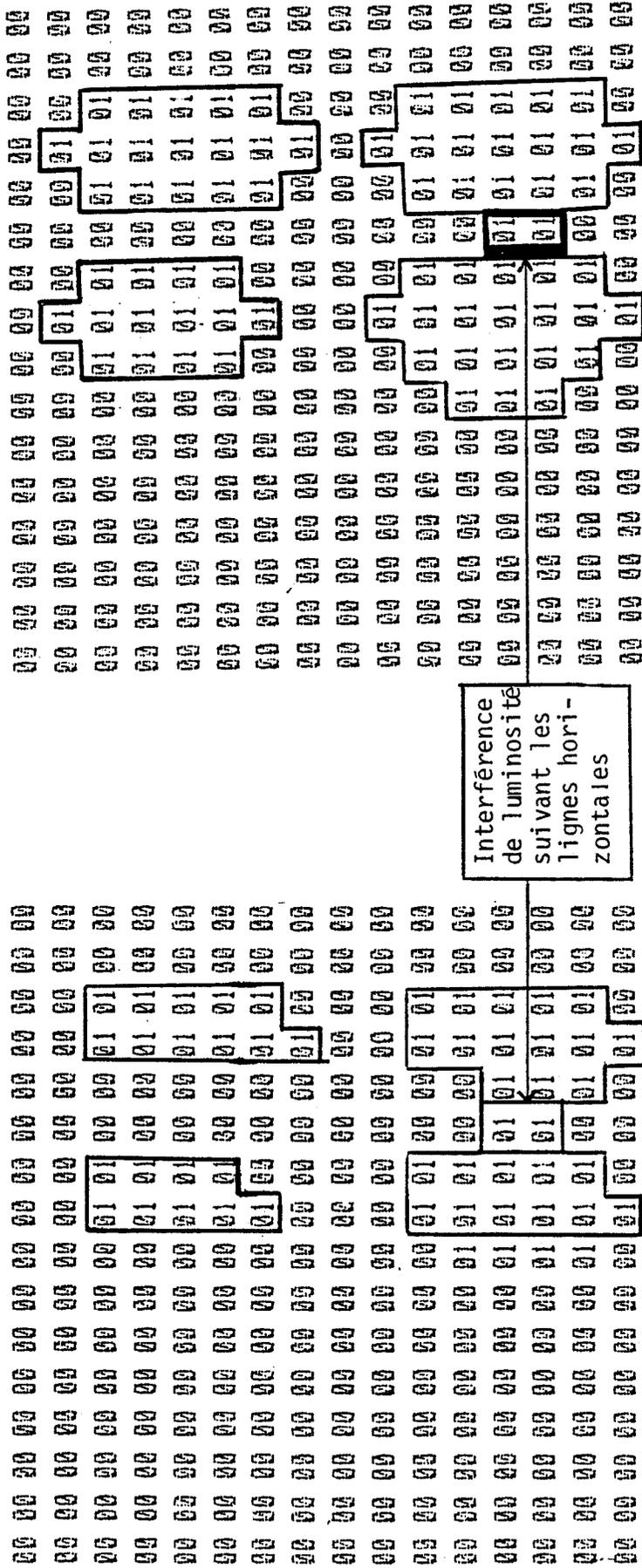


Image (a) dilatée  
Module de dilatation : 0 x o

(d)

Figure III-17

(c)



VI - DISCUSSION

Cette première méthode de reconnaissance du relief citée ci-dessous peut être résumée par le schéma suivant :

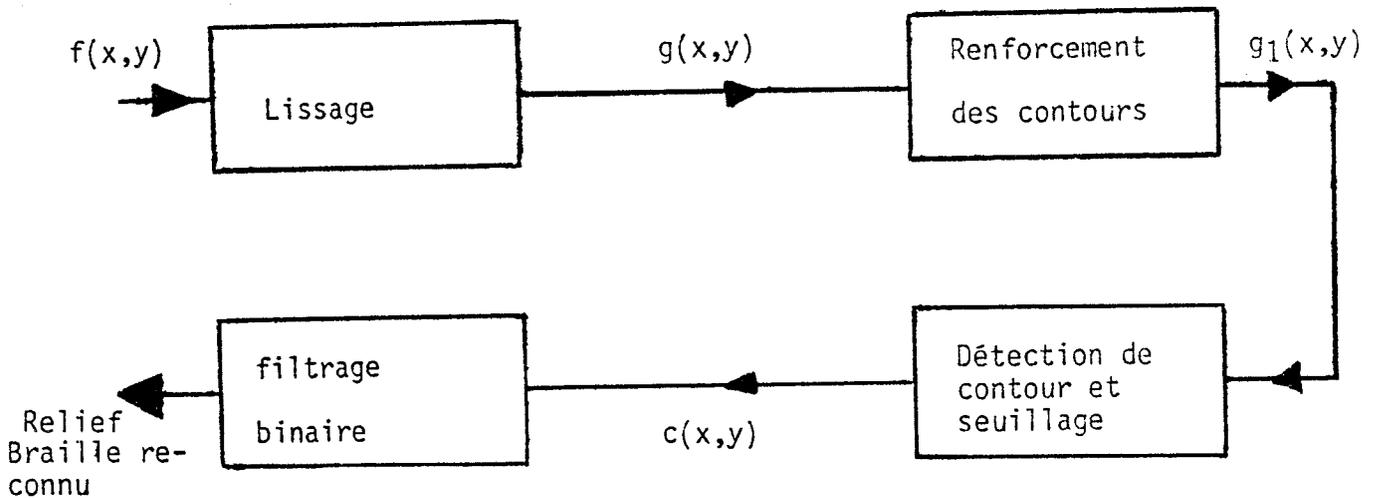


Figure III-18

Cette méthode heuristique extrait les caractères de classe A et B. Mais, l'inconvénient majeur est celui de ne pouvoir estimer sa robustesse en face du bruit et définir un seuil optimal en fonction des propriétés photométriques des documents.

Si l'on fait apparaître le paramètre coefficient de réflexion dû à la couleur du papier (dans nos conditions de mesure), on peut se demander si l'étude précédente reste valable.

L'examen de la figure III-19 montre que la seule différence existante entre documents de couleurs différentes est la fluctuation des niveaux de gris du fond qui est fonction des propriétés photométriques du document considéré.

Pour nous permettre de nous affranchir de la fluctuation des niveaux de gris due aux propriétés photométriques du document, nous allons développer dans la deuxième partie de ce chapitre des méthodes de filtrage qui sont adaptées aux caractéristiques particulières du relief Braille dont l'efficacité d'extraction de caractères peut être estimée approximativement.

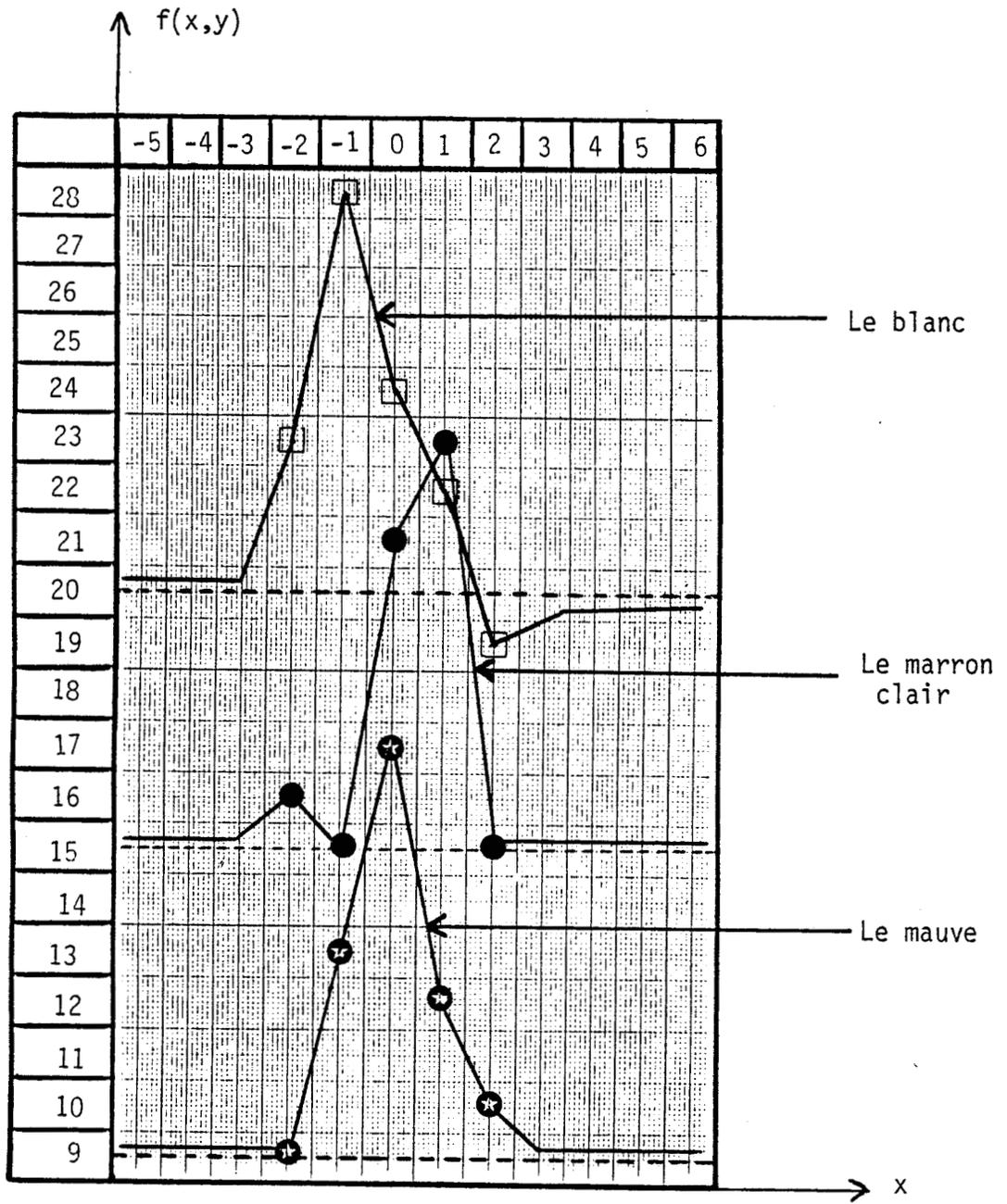


Figure III-19 : Représentation de la luminance suivant une ligne d'un relief Braille appartenant à 3 documents de couleur différente.



DEUXIÈME PARTIE DU CHAPITRE III : APPROXIMATION DE LA LUMINANCE  
DU RELIEF BRAILLE PAR UN MODÈLE POLYNOMIAL À L'AIDE DE RÉGRESSION  
LINÉAIRE.

VII - DETECTION ET MISE EN VALEUR DU RELIEF BRAILLE EN APPROXIMANT LA FONCTION INTENSITE LUMINEUSE PAR UN MODELE POLYNOMIAL

VII-1- LISSAGE DU RELIEF BRAILLE PAR APPROXIMATION POLYNOMIALE

Un point Braille embossé correctement engendre une fonction "intensité lumineuse"  $f(x,y)$  qui diminue graduellement à l'approche des limites du point Braille. En se basant sur cette observation, nous allons d'abord approximer la fonction "intensité lumineuse" par un polynôme monodimensionnel du second degré défini sur le voisinage (1x5) par :

$$z = f(x,y) = b_1 + b_2 x + b_3 x^2 \quad \forall y \quad (1)$$

$x$  et  $y$  sont les coordonnées de pixel connues,  $b_1, b_2, b_3$  des paramètres estimés à partir d'observations discrètes autour d'un pixel central et  $z$  la luminance entachée de bruit.

On considère chaque pixel  $(x,y)$  en relation avec son voisinage (1x5). On suppose que les intensités lumineuses des 5 pixels forment un polynôme du second degré altéré par du bruit. On utilise la méthode (de régression) des moindres carrés pour adapter le modèle aux données de l'image<sup>[37][38][39]</sup>. Le modèle de régression sert à calculer la valeur égalisée  $\hat{z} = \hat{f}(x,y)$ , c'est-à-dire l'intensité lumineuse épurée du bruit définie par :

$$\hat{z} = \hat{f}(x,y) = \hat{b}_1 + \hat{b}_2 x + \hat{b}_3 x^2 \quad (2)$$

$\hat{b}_1, \hat{b}_2, \hat{b}_3$  sont des paramètres inconnus à estimer en utilisant le critère des moindres-carrés.

Il supprime le bruit, tout en laissant intact les segments du second degré. Cette méthode convient pour dériver des masques symétriques et asymétriques de tailles diverses.

VII-1-1- Détermination des masques de lissage du relief Braille par la méthode des moindres carrés

Considérons un processus excité par un certain nombre d'échantillons  $u_i$  à l'entrée et délivrant des échantillons  $z_i$  à sa sortie ( $1 \leq i \leq n$ ). On notera  $u$  le vecteur d'entrée et  $z$  le vecteur de sortie. Nous poursuivrons une expérience et nous enregistrerons les entrées et les sorties :

$$(u_1, z_1), (u_2, z_2), \dots, (u_n, z_n)$$

Supposons que le processus puisse être décrit par le modèle :

$$z = b_1 h_1(u) + \dots + b_k h_k(u) = h^T(u)b \quad (3)$$

$h(u) = [h_1(u), \dots, h_k(u)]^T$  est le vecteur de fonctions d'entrées connues.

La matrice (n x k) des entrées des expériences, le vecteur des sorties d'expériences et le vecteur des sorties du modèle sont définis par :

$$U = \begin{vmatrix} h_1(u_1), \dots, h_k(u_1) \\ \vdots \\ h_1(u_n), \dots, h_k(u_n) \end{vmatrix} \quad (4)$$

$$z = \begin{vmatrix} z_1 \\ \vdots \\ z_n \end{vmatrix} = Ub \quad (5)$$

$$\hat{z} = \begin{vmatrix} \hat{z}_1 \\ \vdots \\ \hat{z}_n \end{vmatrix} = U\hat{b} \quad (6)$$

On détermine les paramètres  $\hat{b}_1, \hat{b}_2, \hat{b}_3$  en minimisant l'écart entre l'estimation  $\hat{z}$  et l'observation  $z$ . En utilisant un critère au sens des moindres carrés, alors :

$$S = \sum_{n=1}^{n=5} (\hat{z}_n - z_n)^2$$

$$S = (z - U\hat{b})^T (z - U\hat{b})$$

S est minimale lorsque  $\frac{\partial S}{\partial \hat{b}_k} = 0 \quad \forall k$

Il en est ainsi pour :

$$\hat{b} = [U^T U]^{-1} U^T z \quad (7)$$

Cette estimation  $\hat{b}$  est utilisée pour calculer la sortie correspondante à une entrée donnée :

$$\hat{z} = \hat{b}_1 h_1(u) + \dots + \hat{b}_k h_k(u) = h^T(u) \hat{b} \quad (8)$$

Si on considère un voisinage (1x5) pour l'identification comme le montre la figure ci-dessous :

x	-2	-1	0	1	2
f(x,y)	z <sub>1</sub>	z <sub>2</sub>	z <sub>3</sub>	z <sub>4</sub>	z <sub>5</sub>

(9)

L'application de l'expression (1) pour chacune des 5 abscisses x conduit à une équation matricielle qui, identifiée à l'expression (5), permet d'écrire U pour un modèle à une seule dimension :

$$U = \begin{vmatrix} 1 & -2 & 4 \\ 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{vmatrix} \quad (10)$$

A partir de l'expression (8) qui décrit la réponse du modèle à la sollicitation  $h_i(u)$ , on peut calculer l'intensité  $\hat{z}$  du pixel central. Soit :

$$\hat{z} = h^T(x_3, y) \hat{b} = h_3^T \hat{b} \quad (11)$$

avec le vecteur d'entrée  $h_3$  défini par :

$$h_3 = [1 \ 0 \ 0]^T \quad (12)$$

Pour  $x = 0$  et à partir de (6), (7), (10), nous obtenons :

$$\hat{z}_3 = \hat{b}_1$$

et

$$\hat{z} = \frac{1}{12} [-z_1 + 4z_2 + 6z_3 + 4z_4 - z_5] \quad (13)$$

La formule (13) est équivalente au masque fixe suivant :

$$MS1 = \frac{1}{12} [-1 \quad 4 \quad 6 \quad 4 \quad -1] \quad (14)$$

L'application du masque (14) correspond à l'opération suivante : On suppose que le voisinage (1x5) présente un point Braille décrit par le polynôme (1) suivant les lignes horizontales et est altéré par du bruit. On fait correspondre les coefficients optimums aux sens des moindres carrés aux données et on calcule le pixel central. Par comparaison avec l'intensité lumineuse d'origine, cette valeur contient un bruit fortement adouci.

Le masque (14) n'est pas adapté aux pixels du bord de l'épure de l'image.

Le voisinage du pixel du bord est limité des deux côtés suivant les lignes horizontales.

Dans le cas du côté gauche du bord, l'abscisse  $x$  prend les valeurs 0, 1, 2, 3, 4 comme le montre la figure ci-dessous.

x	0	1	2	3	4
z	$z_1$	$z_2$	$z_3$	$z_4$	$z_5$

(15)

Dans ces conditions, la matrice  $U$  et le masque fixe permettant de lisser les pixels du bord sont définis par :

$$U = \begin{vmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \\ 1 & 4 & 16 \end{vmatrix} \quad (16)$$

et 
$$MS2 = \frac{1}{12} [11 \quad 3 \quad -1 \quad -2 \quad 1] \quad (17)$$

Cependant, l'examen de la distribution des niveaux de gris de la luminance suivant les lignes horizontales et verticales montre que les points Braille se manifestent de la même manière suivant ces deux directions. C'est pour cette raison, que nous allons approximer également la fonction luminance par un polynôme du second degré sur les voisinages (3x3) et (5x5) (figure ci-dessous) défini par :

$$z = f(x,y) = b_1 + b_2 x + b_3 x^2 + b_4 y + b_5 y^2 \quad (18)$$

$y \quad x$	-1	0	1	(19)
-1	$z_1$	$z_2$	$z_3$	
0	$z_4$	$z_5$	$z_6$	
1	$z_7$	$z_8$	$z_9$	

Définition du voisinage (3x3)

Comme pour le modèle (1), la méthode des moindres carrés s'applique également au modèle (18) et nous déduisons U pour un modèle à deux dimensions :

$$U = \begin{vmatrix} 1 & -1 & 1 & -1 & 1 \\ 1 & 0 & 0 & -1 & 1 \\ 1 & 1 & 1 & -1 & 1 \\ 1 & -1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & -1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{vmatrix} \quad (20)$$

Le vecteur d'entrée  $h_5$  :

$$h_5 = [1 \ 0 \ 0 \ 0 \ 0]^T \quad (21)$$

Le masque fixe MS3 qui supprime le bruit en laissant intactes les surfaces du second degré est de la forme :

$$MS3 = \frac{1}{9} \begin{vmatrix} -1 & 2 & -1 \\ 2 & 5 & 2 \\ -1 & 2 & -1 \end{vmatrix} \quad (22)$$

Si nous considérons par exemple le pixel d'angle pour les points appartenant aux bords de l'épure de l'image, le voisinage du pixel d'angle est limité des deux côtés et les coordonnées x et y prennent les valeurs 0, 1, 2 et -2, -1, 0.

Dans ces conditions, la matrice U, le vecteur d'entrée  $h_5$  et le masque fixe MS4 sont définis par :

$$U = \begin{vmatrix} 1 & 0 & 0 & -2 & 4 \\ 1 & 1 & 1 & -2 & 4 \\ 1 & 2 & 4 & -2 & 4 \\ 1 & 0 & 0 & -1 & 1 \\ 1 & 1 & 1 & -1 & 1 \\ 1 & 2 & 4 & -1 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 2 & 4 & 0 & 0 \end{vmatrix} \quad (23)$$

$$h_5 = [1 \ 1 \ 1 \ -1 \ 1]^T \quad (24)$$

$$MS4 = \frac{1}{9} \begin{vmatrix} 5 & 2 & 2 \\ 5 & -1 & -1 \\ 2 & -1 & -1 \end{vmatrix} \quad (25)$$

Supposons que les surfaces des objets à détecter ne soient pas contenues dans le voisinage (3x3), mais soient plus étendues. Dans ce cas, nous pouvons lisser de deux manières distinctes :

- soit en définissant l'approximation polynomiâle sur un voisinage (5x5), (7x7) ... etc
- soit en effectuant plusieurs passages sur l'image avec le même masque de convolution linéaire, par exemple (3x3).

L'examen du relief Braille effectué au chapitre I montre que les objets ont une faible étendue et leur détection ne peut être faite sur un voisinage supérieur à (5x5) points élémentaires.

Nous avons choisi d'effectuer deux passages successifs du filtre MS3 centré sur l'image en relief Braille.

Dans ce cas, on montre que deux passages successifs par un masque (3x3) de convolution linéaire d'images ayant la forme ci-dessous sont équivalents à un seul passage par un filtre défini sur un voisinage (5x5) points élémentaires.

$$\begin{vmatrix} c & b & c \\ b & a & b \\ c & b & c \end{vmatrix}$$

Forme du masque 3x3

Nous déduisons alors le filtre MS5 défini sur le voisinage (5x5) par :

$$\begin{aligned} &\Rightarrow \\ &\begin{vmatrix} c & b & c \\ b & a & b \\ c & b & c \end{vmatrix} * \begin{vmatrix} c & b & c \\ b & a & b \\ c & b & c \end{vmatrix} = \begin{vmatrix} a_5 & a_4 & a_2 & a_4 & a_5 \\ a_4 & a_3 & a_1 & a_3 & a_4 \\ a_2 & a_1 & a_0 & a_1 & a_2 \\ a_4 & a_3 & a_1 & a_3 & a_4 \\ a_5 & a_4 & a_2 & a_4 & a_5 \end{vmatrix} \\ &\frac{1}{9} \begin{vmatrix} -1 & 2 & -1 \\ 2 & 5 & 2 \\ -1 & 2 & -1 \end{vmatrix} * \frac{1}{9} \begin{vmatrix} -1 & 2 & -1 \\ 2 & 5 & 2 \\ -1 & 2 & -1 \end{vmatrix} = \frac{1}{81} \begin{vmatrix} 1 & -4 & 6 & -4 & 1 \\ -4 & -2 & 12 & -2 & -4 \\ 6 & 12 & 45 & 12 & 6 \\ -4 & -2 & 12 & -2 & -4 \\ 1 & -4 & 6 & -4 & 1 \end{vmatrix} \quad (26) \end{aligned}$$

$$\begin{aligned} \text{avec} \quad a_0 &= 4c^2 + 4b^2 + a^2, & a_1 &= 2bc + ab + 2bc + ab \\ a_2 &= b^2 + 2c^2, & a_3 &= ac + 2b^2 + ac \\ a_4 &= 2bc, & a_5 &= c^2 \end{aligned}$$

Les masques MS1, MS2 et MS5 sont des filtres qui atténuent le bruit, lissent les contours du relief et laissent intacts les segments et les surfaces du second degré.

Dans le paragraphe suivant, nous allons calculer les masques de détection du relief Braille à l'aide des masques de lissage précédents.

### VII-2- DETECTION DU RELIEF BRAILLE

Pour détecter les points Braille, nous allons utiliser un détecteur fixe qui a les propriétés suivantes :

- Il met en valeur les segments du second degré.
- La sortie du filtre est positive si la zone d'image qu'il recouvre présente un objet convexe et négative si l'objet est concave.
- La sortie du filtre est nulle pour les surfaces à luminance constante.

Nous avons dérivé le détecteur du relief Braille directement du masque MS1 (14). La réponse de ce filtre doit être nulle pour les régions à luminance constante. Il faut donc contraindre la somme des coefficients du détecteur à être nulle.

Pour réaliser cette condition, il faut ajouter une constante à chaque coefficient du filtre. Pour le masque (14), la somme des coefficients est égale à 1 et le nombre de poids est égal à 5. Par conséquent, on doit soustraire de chaque coefficient la constante  $1/5$  pour contraindre la somme des poids à être nulle et on obtient le détecteur du relief Braille suivant :

$$MS6 = \frac{1}{120} \begin{bmatrix} -34 & 16 & 36 & 16 & -34 \end{bmatrix} \quad (27)$$

De la même manière, à partir de MS3 (22) et MS5 (26), nous obtenons les détecteurs définis sur les voisinages (3x3) et (5x5) :

$$MS7 = 1/9 \begin{vmatrix} -2 & 1 & -2 \\ 1 & 4 & 1 \\ -2 & 1 & -2 \end{vmatrix} \quad (28)$$

$$MS8 = \frac{1}{8I} \begin{vmatrix} 4 & -4 & 9 & -4 & 4 \\ -4 & -14 & 0 & -14 & -4 \\ 9 & 0 & 36 & 0 & 9 \\ -4 & -14 & 0 & -14 & -4 \\ 4 & -4 & 9 & -4 & 4 \end{vmatrix} \quad (29)$$

Comme nous allons voir au paragraphe suivant où nous décrivons la partie expérimentale de ce travail, MS7 (28) et MS8 (29) ne sont pas adaptés à la détection du relief Braille.

### VII-3- PARTIE EXPERIMENTALE

Dans ce paragraphe, nous allons présenter et commenter les résultats expérimentaux obtenus. Nous définissons également une méthode simple permettant la reconnaissance du relief Braille sans ambiguïté.

Les figures III-20, III-21, III-22, III-23, III-24 représentent l'effet du filtre MS3 (22) sur les figures I-9, I-11, I-12, I-13, I-14 du chapitre I.

Ce filtre défini sur le voisinage (3x3) lisse et met en valeur les contours des graphèmes Braille appartenant aux classes A, B et une partie de D comme le montrent les figures III-20, III-21, III-22 et III-24.

Mais, par contre, ce masque opérateur ne permet pas de lisser les points Braille dont la variation de luminosité est importante suivant les lignes, ce qui va se traduire par des trous dans la forme des points Braille lors du passage du détecteur du relief.

Il laisse aussi intactes les surfaces décrites par :

$$f(x,y) = b_1 + b_2 x + b_3 y \quad (30)$$

Les propriétés (18) et (30) ont comme inconvénient de ne pas permettre au masque de lisser les points Braille dont la variation de luminosité est importante suivant les lignes (nous supposons que le voisinage (3x3) est adapté à la forme du signal) et de ne pas atténuer le bruit dont la forme du signal peut être décrite par (18) ou par (30).

Deux exemples illustrant ce problème sont représentés figures III-23 et III-24.

La figure III-23 présente le résultat du passage de MS3 (22) sur l'image I-13. Un seul passage de ce filtre sur l'image I-13 n'est pas suffisant pour atténuer la fluctuation de la luminance suivant les lignes. Nous aurons alors tendance à effectuer un deuxième passage du même filtre sur l'image pour remédier à ce problème ou effectuer un seul passage sur I-13 du filtre (26) défini sur le voisinage (5x5).

Un deuxième passage du masque MS3 sur l'image III-23 laissera inchangée la forme du signal. En plus, il peut relier entre eux les points Braille voisins suivant les lignes horizontales et faire disparaître les objets très faiblement contrastés dont la forme n'est pas contenue dans le voisinage (5x5).

La figure III-24 présente un exemple où le passage du filtre MS3 sur l'image figure III-14 n'a pas atténué fortement le bruit. Pour l'atténuer efficacement, la solution consiste à augmenter la définition du voisinage de façon que la surface définissant le bruit ne soit plus contenue dans le voisinage, ce qui est contradictoire avec ce que nous avons cité ci-dessus.

Dans la deuxième partie expérimentale de ce paragraphe, nous avons appliqué à l'image Braille les masques (14) et (27) définis sur le voisinage (1x5) pour lisser et reconnaître le relief ; ce qui va nous permettre de mettre en oeuvre une méthode simple et efficace pour l'extraire. Les masques (3x3) et (5x5) n'étant pas parfaitement adaptés à la forme du signal.

Les figures III-25-a, III-26-a, III-27-a, III-28-a, III-29-a présentent le résultat du passage du masque (14) sur les images du chapitre I, figures I-9, I-11, I-12, I-13 et I-14.

Les réponses de ce filtre obtenues montrent que ce masque lisse met en valeur et adoucit les contours des graphèmes appartenant aux classes A, B et une partie de D.

Ce masque opérateur a tendance à atténuer faiblement l'amplitude du signal, ce qui peut provoquer une perte d'information surtout pour les graphèmes

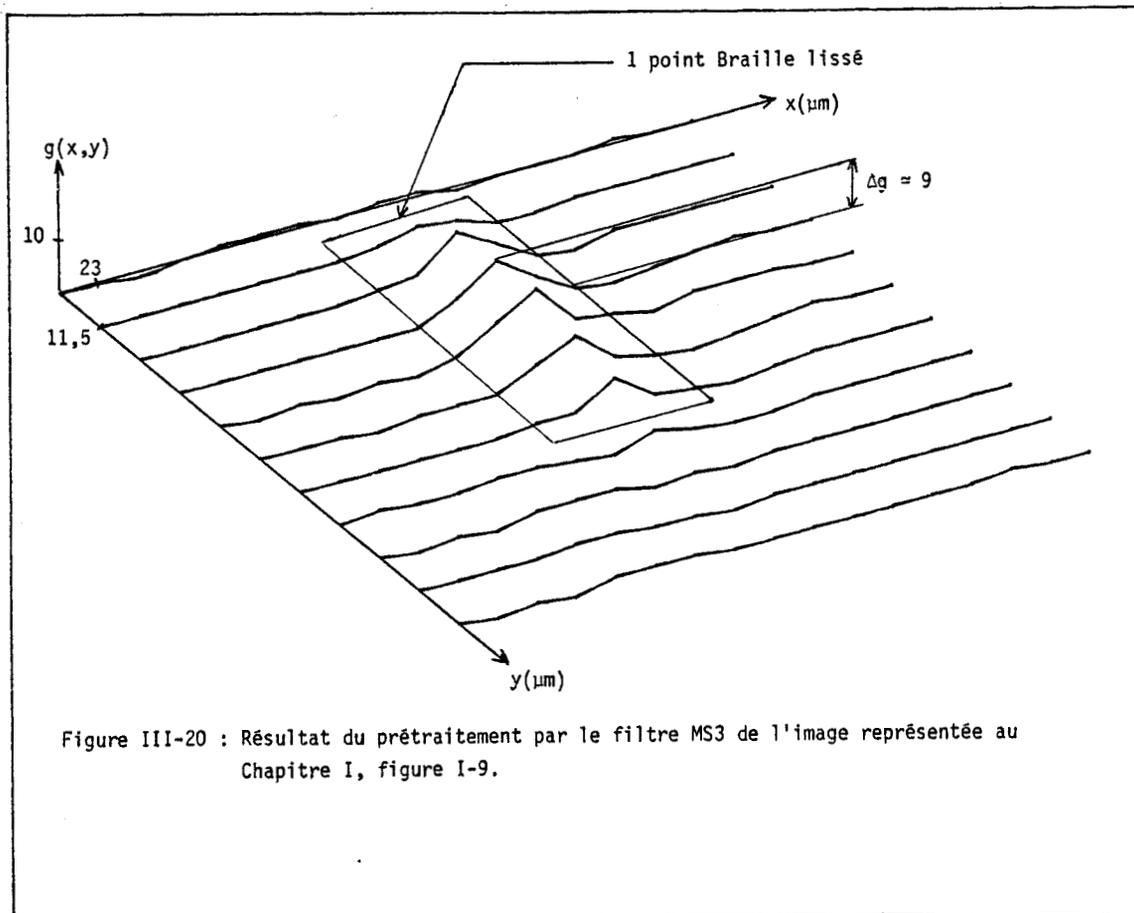


Figure III-20 : Résultat du prétraitement par le filtre MS3 de l'image représentée au Chapitre I, figure I-9.

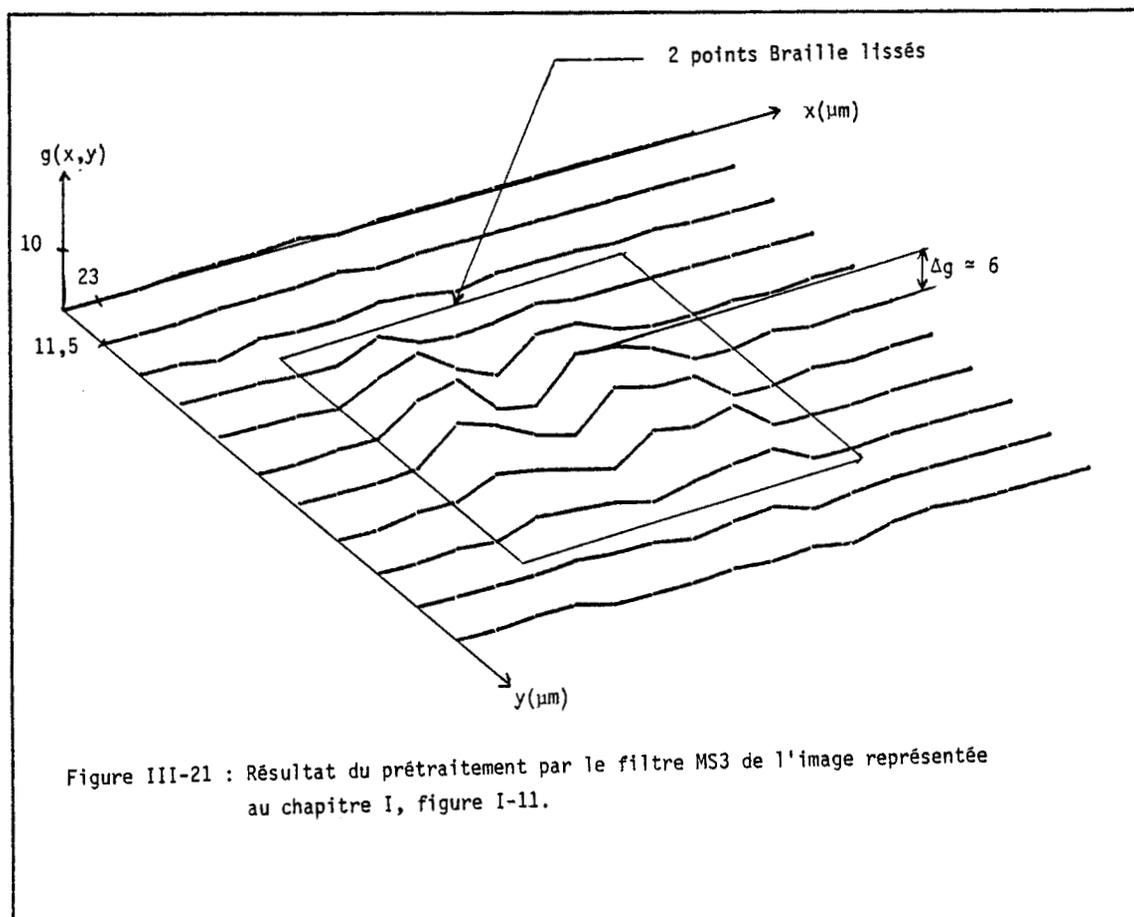


Figure III-21 : Résultat du prétraitement par le filtre MS3 de l'image représentée au chapitre I, figure I-11.



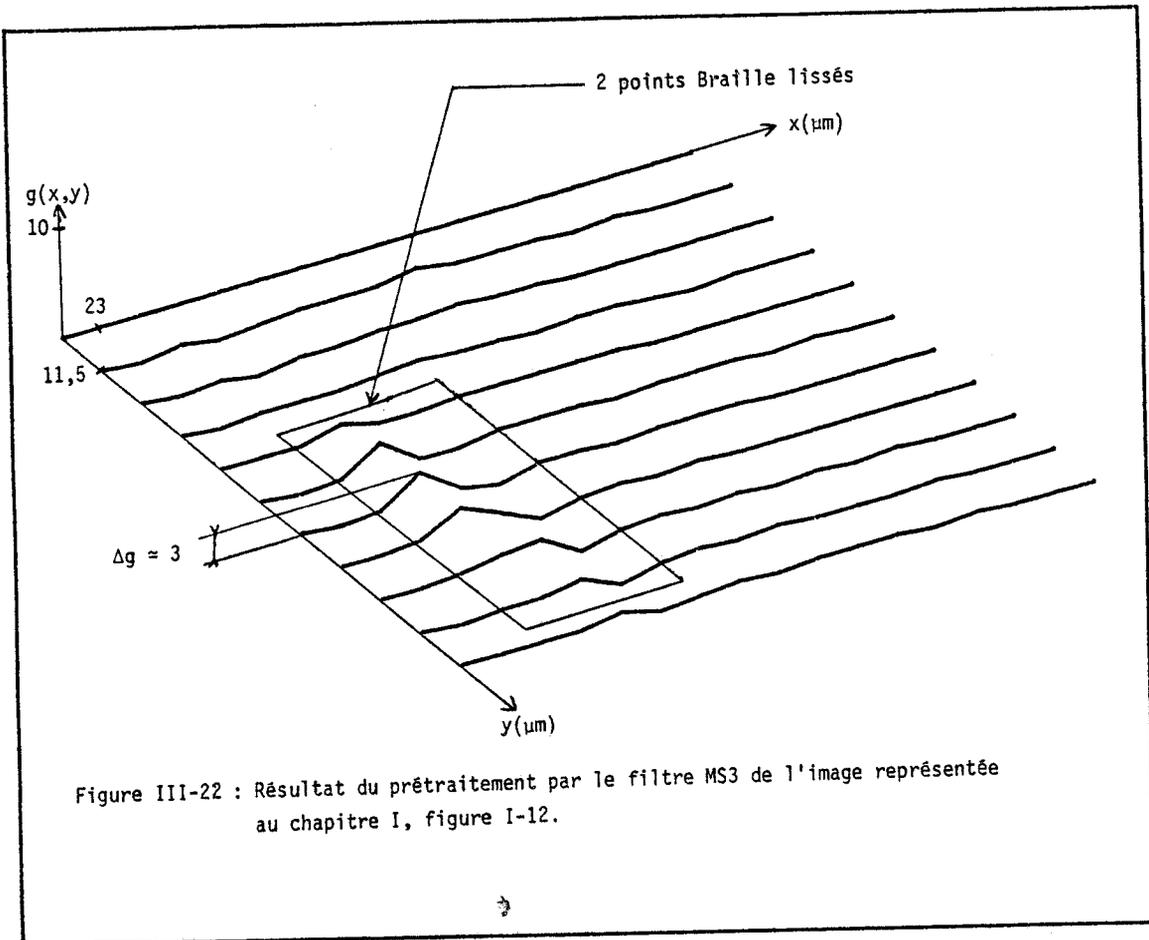


Figure III-22 : Résultat du prétraitement par le filtre MS3 de l'image représentée au chapitre I, figure I-12.

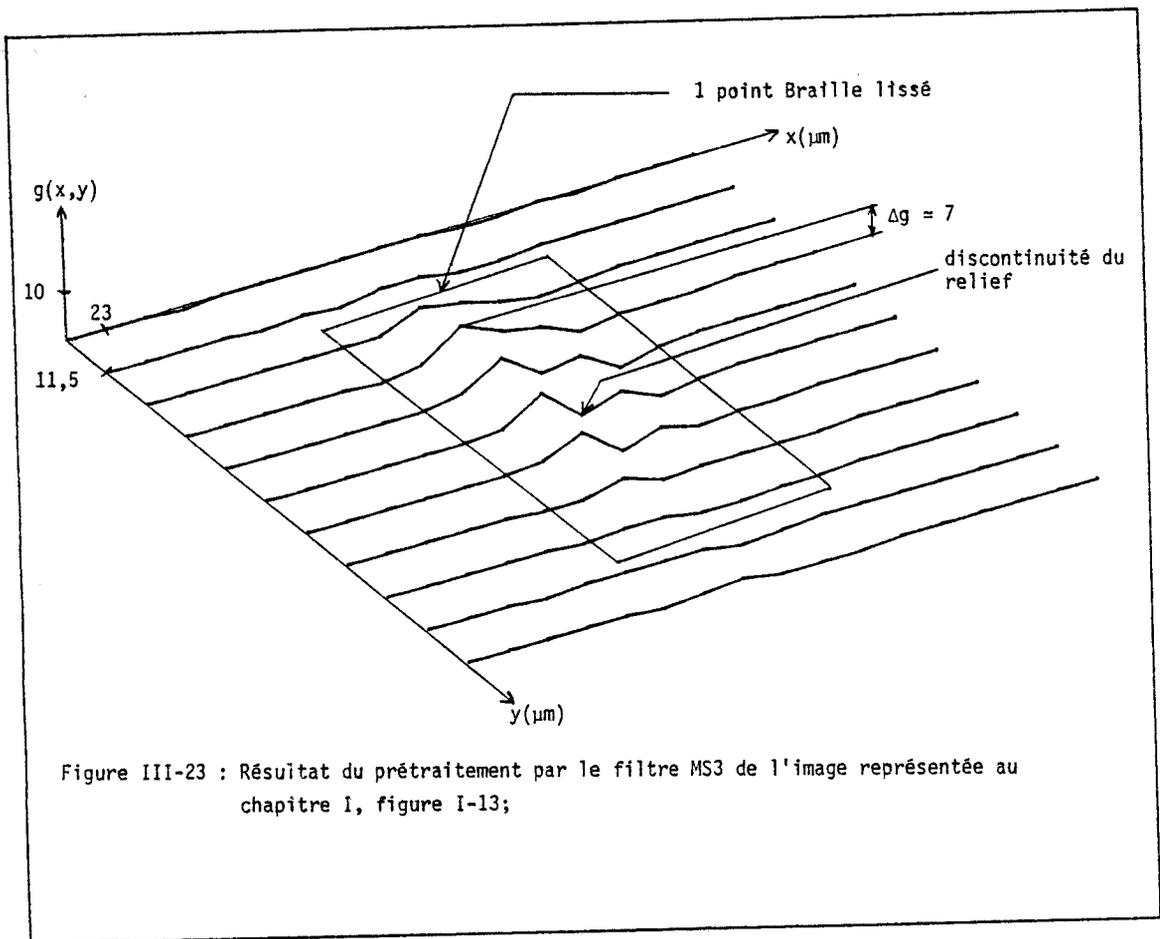


Figure III-23 : Résultat du prétraitement par le filtre MS3 de l'image représentée au chapitre I, figure I-13;

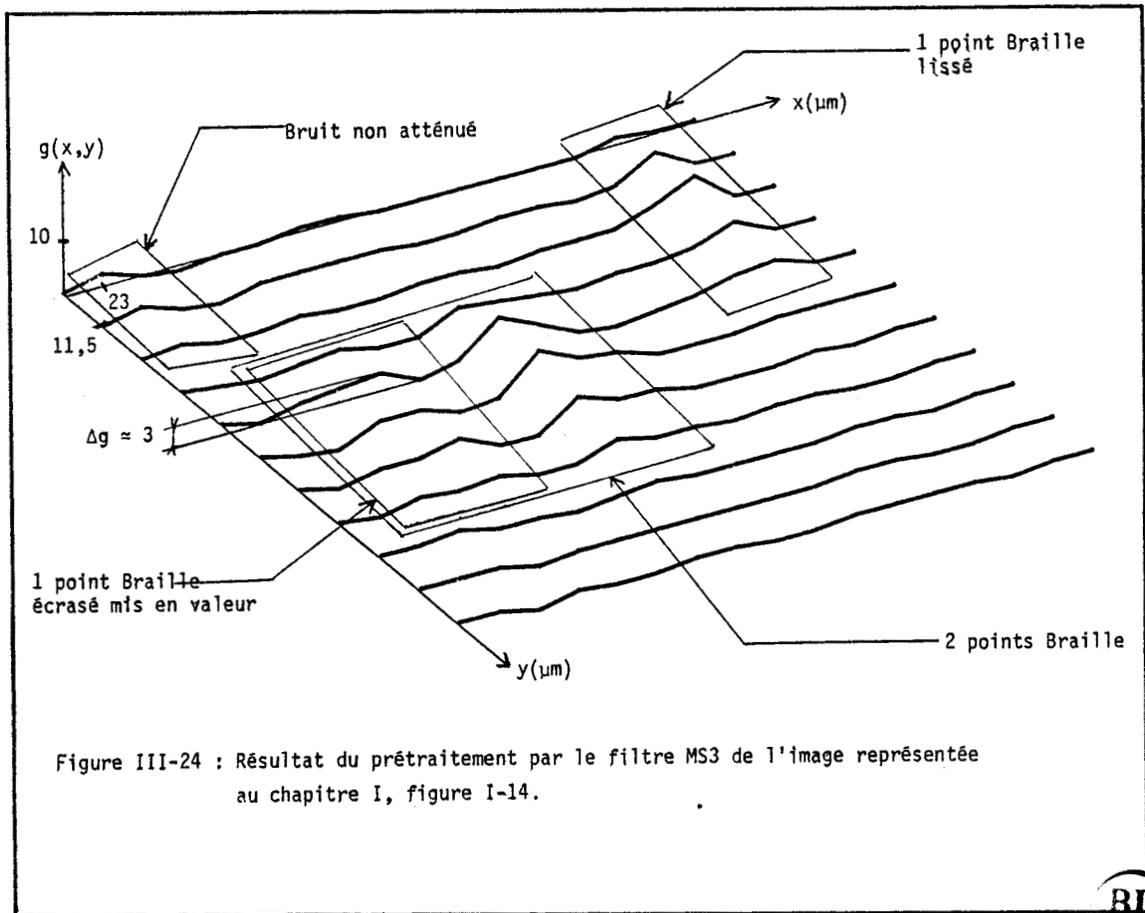


Figure III-24 : Résultat du prétraitement par le filtre MS3 de l'image représentée au chapitre I, figure I-14.

mal embossés. Dans le cas des objets bien décrits par (1), même pour ceux qui sont très faiblement contrastés, cet effet ne se produit pas, ce qui va augmenter le taux de réussite d'extraction du relief. Par exemple, la figure III-29-a représente le résultat du prétraitement d'un point Braille dont la luminosité suivant les lignes horizontales varie de 2 à 3 niveaux de gris qui est bien prétraité et montre aussi la limite de lissage du masque (14).

Ce filtre lisse parfaitement les points Braille dont la variation de luminosité suivant les lignes est importante, ce qui évite l'apparition des trous dans la forme du relief lors de la détection. Un exemple est représenté figure III-28-a. Il atténue fortement le bruit, mais aussi la luminance de l'objet dont la forme du signal se confond avec celle du bruit, ce qui peut provoquer une perte d'information.

Un exemple est représenté figure III-29-a.

Le masque (14) laisse intactes les surfaces planes et les segments décrits par :

$$f(x,y) = b_1 + b_2 x \quad (31)$$

Dans le cas de la figure III-29-a, le bruit et le point Braille ne sont pas décrits par (31), ce qui entraîne une atténuation de la luminance de ces deux derniers.

Cette méthode utilisant un masque de régression fixe basée sur un certain nombre d'observations autour d'un point ne peut différencier le relief et le bruit dont la forme peut se confondre comme dans le dernier exemple cité ci-dessus. Mais, par contre, elle permet un prétraitement efficace du relief, ce qui va faciliter sa détection et améliorer le taux de son extraction. C'est le masque (14) qui approche le mieux la forme du relief. Nous le retenons alors pour le prétraitement de l'image Braille.

Dans le paragraphe précédent, nous avons vu que la présence d'un point se manifeste lorsque la sortie du masque (27) est positive et dans le cas des surfaces à luminance constante, la réponse du filtre est nulle. Pour détecter les points Braille prétraités par (14), il suffit de convoluer l'image avec le masque (27) et de ne garder que les valeurs positives de la sortie.

Les figures III-25-b, III-26-b, III-27-b, III-28-b, III-29-b illustrent l'effet de cette opération sur les images figures III-25-a, III-26-a, III-27-a, III-28-a, III-29-a.

Les résultats obtenus montrent qu'un seul passage sur l'image est suffisant et efficace pour extraire le relief. Un deuxième passage du filtre (27) va entraîner la perte des objets très faiblement contrastés.

Cependant, outre le relief détecté, l'image contient aussi un bruit impulsionnel.

Le lissage du relief et l'élimination du bruit sont effectués par un filtre médian monodimensionnel suivant les lignes verticales défini sur un voisinage de 3 points (décrit dans la partie filtrage binaire de ce chapitre).

Le passage du filtre médian 3 points sur l'image peut entraîner une perte d'information lorsque les formes du signal ne sont pas adaptées à la définition de son voisinage.

La figure III-29-c présente un exemple où le passage du filtre médian 3 points sur un point Braille se composant d'un seul pixel a provoqué sa perte. En effet, ce filtre traitant globalement l'image ne peut différencier un objet qui peut par exemple être confondu avec un bruit impulsionnel.

Cependant, la distribution de la luminance des objets détectés étant au minimum deux fois plus importante suivant les lignes verticales qu'horizontales, le risque de perte d'information par le filtre médian 3 points est très faible.

Les figures III-25-c, III-26-c, III-27-c, III-28-c, III-29-c représentent l'effet du filtre médian 3 points vertical qui lisse les contours du relief et élimine le bruit sur les figures III-25-b, III-26-b, III-27-b, III-28-b et III-29-b.

Pour seuiller l'image, nous effectuons l'opération non linéaire définie par :

$$c(x,y) = T [ g_2(x,y) ]$$

$$\text{avec } T = \begin{cases} c(x,y) = 0 & , \text{ si } g_2(x,y) \leq 0 \\ c(x,y) = 1 & , \text{ si } g_2(x,y) > 0 \end{cases}$$

$g_2(x,y)$  étant l'image prétraitée par MS1, traitée par MS6 et lissée par le filtre médian 3 points.

Le seuillage permet d'obtenir une image binaire représentant l'information "relief" par des "1" et son absence par des "φ".

Pour faciliter la reconnaissance des graphèmes Braille, nous effectuons sur chaque point de l'image binarisée deux opérations :

- la première opération revient, pour chaque point de l'image binarisée de valeur 1 à tester son 8-voisinage (définition de 8-voisinage figure ci-dessous) et satisfaire la relation suivante :

$P_1$	$P_2$	$P_3$		+ : point testé de l'image
0	0	0		
$P_8$	0	+	0	$P_4$
	0	0	0	
$P_7$	$P_6$	$P_5$		

Définition de 8-voisinage



$$S = \sum_{i=1}^{i=8} P_i$$

si  $S = 0$ , nous attribuons la valeur 1 aux points  $P_4, P_5, P_6$

si  $S \neq 0$ , chaque point du 8-voisinage garde sa valeur

- la deuxième opération revient, pour chaque point de l'image binarisée de valeur φ à tester son voisinage (définition de voisinage figure ci-dessous) et satisfaire la relation suivante :

0	$P_1$	+ : point testé de l'image
	+	
0	$P_2$	0 : point appartenant au voisinage

Définition du voisinage

$$S = \sum_{i=1}^{i=2} P_i$$

Si  $S = 2$ , nous attribuons la valeur 1 au point testé de l'image.

Si  $S \neq 2$ , le point testé de l'image garde la valeur  $\phi$

Dans le cas d'une image représentant le relief Braille, les deux filtres cités ci-dessus permettent de détecter les points Braille se composant d'un seul pixel, de les dilater et de fermer les trous sans risque d'interférences de luminosité entre les points.

Les figures III-25-d, III-26-d, III-27-d, III-28-d, III-29-d illustrent l'effet des deux filtres appliqués aux images figures III-25-c, III-26-c, III-27-c, III-28-c, III-29-c, résultant des traitements précédents.

La figure III-29-c représente un point Braille défini par un seul point élémentaire qui a été détecté et dilaté par le filtre 1 comme le montre la figure III-29-d. Ce point Braille étant élément d'un caractère Braille de classe D.

Les résultats obtenus figures III-25-d, III-26-d, III-27-d, III-28-d, III-29-d résultant du dernier traitement de détection du relief montrent l'efficacité d'extraction de cette méthode qui peut être résumée par le schéma suivant :

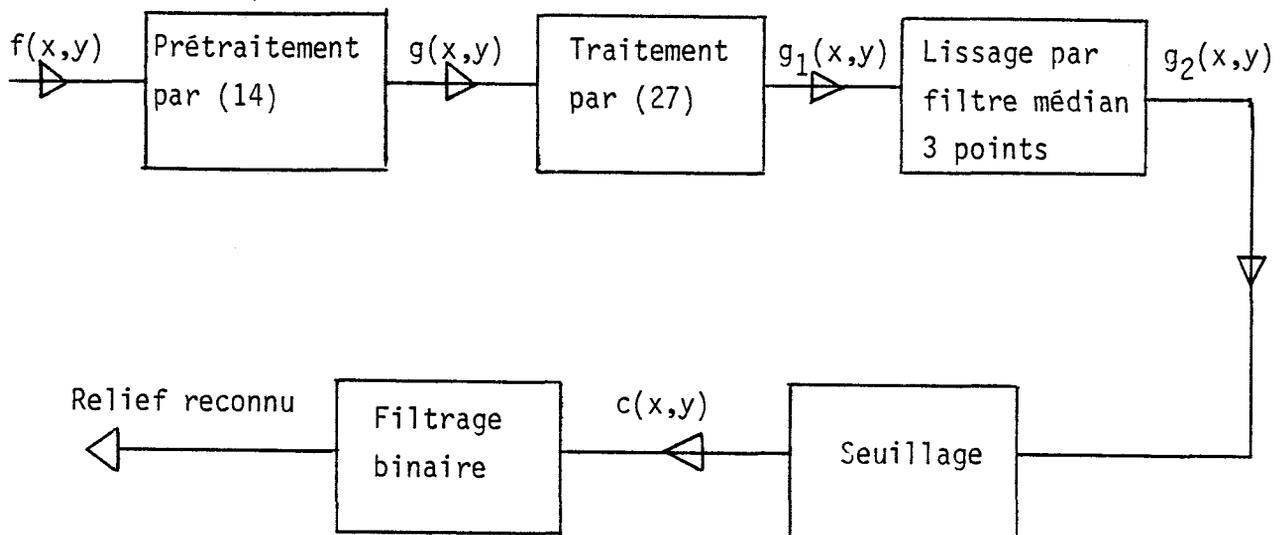
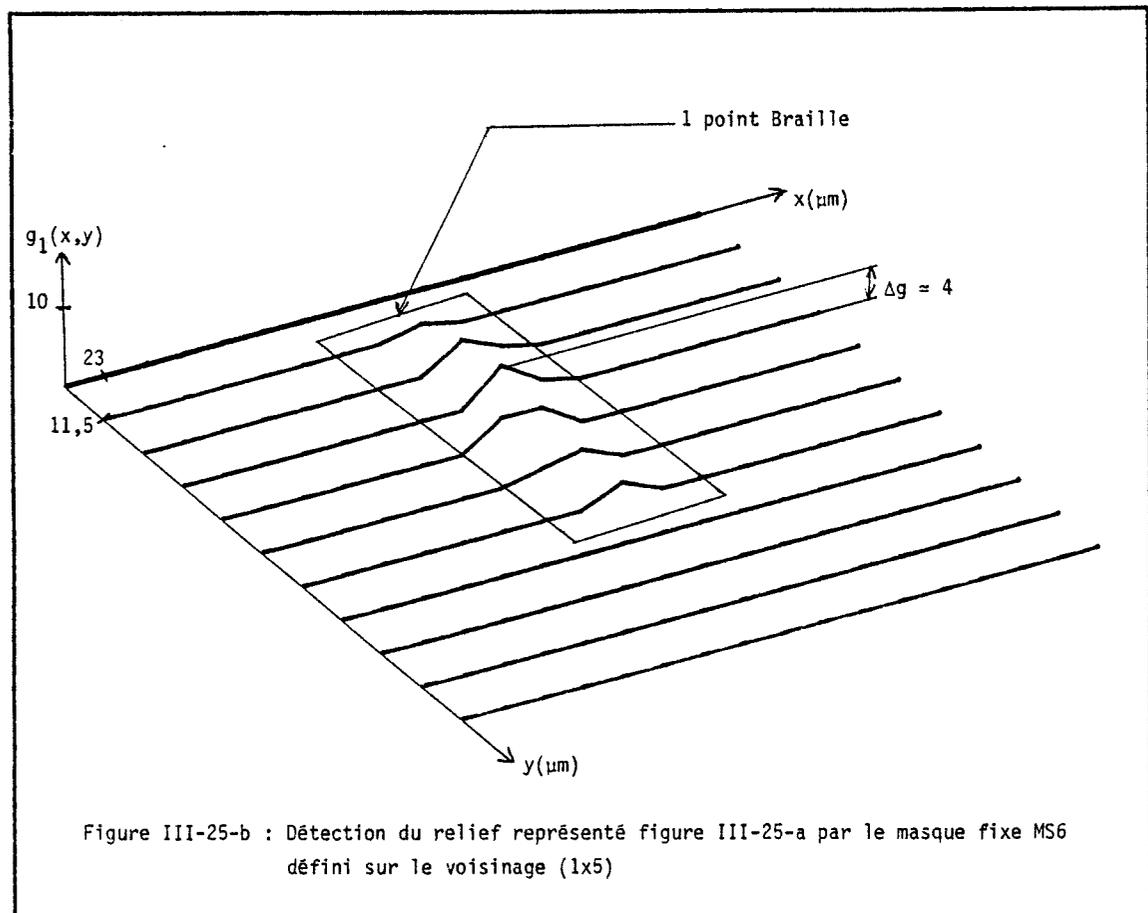
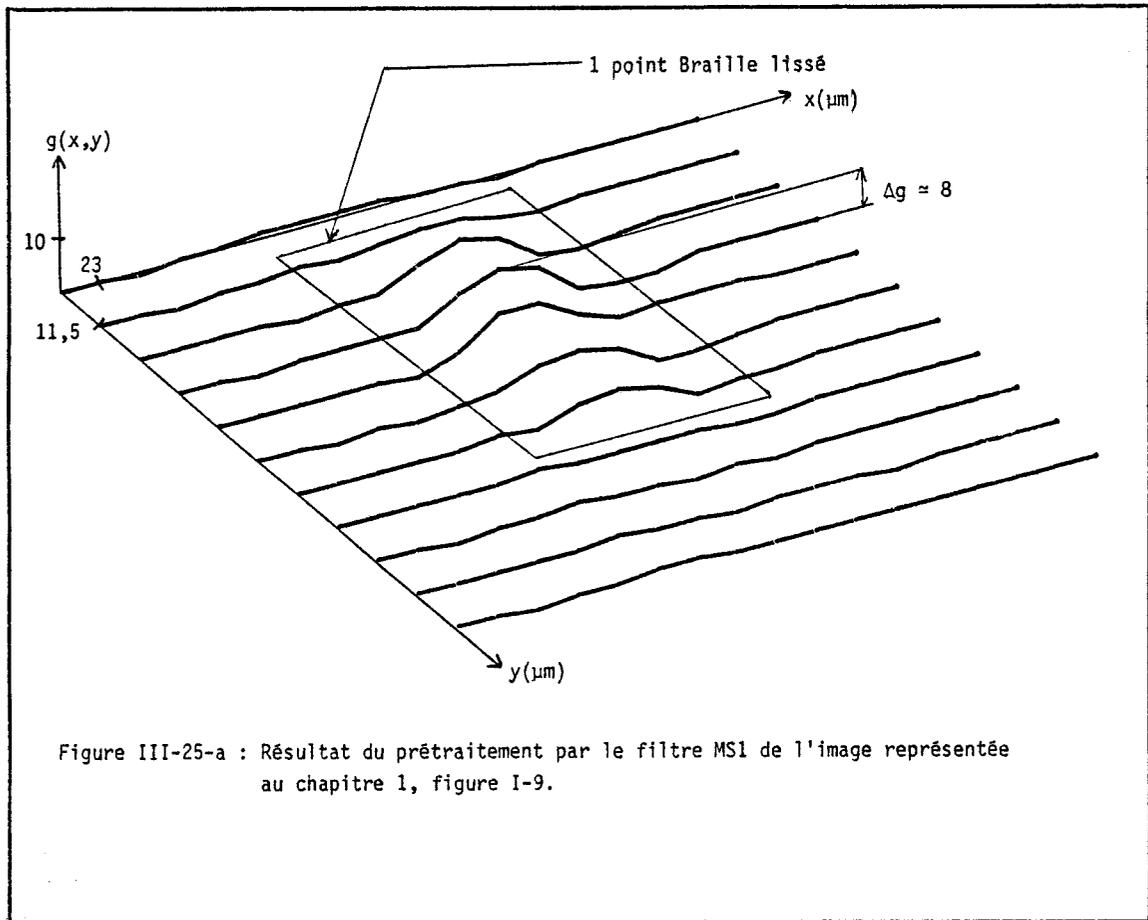
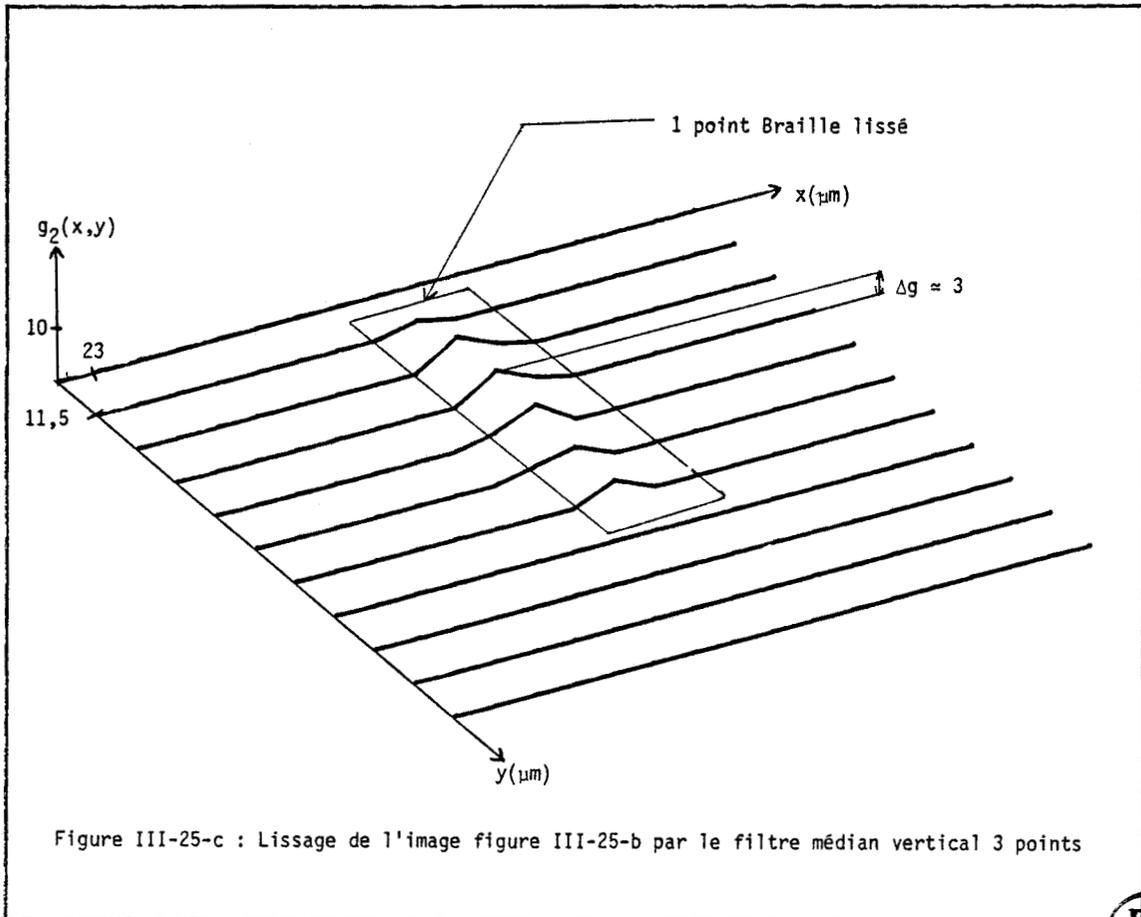


Figure 30 : Schéma de principe de détection du relief.

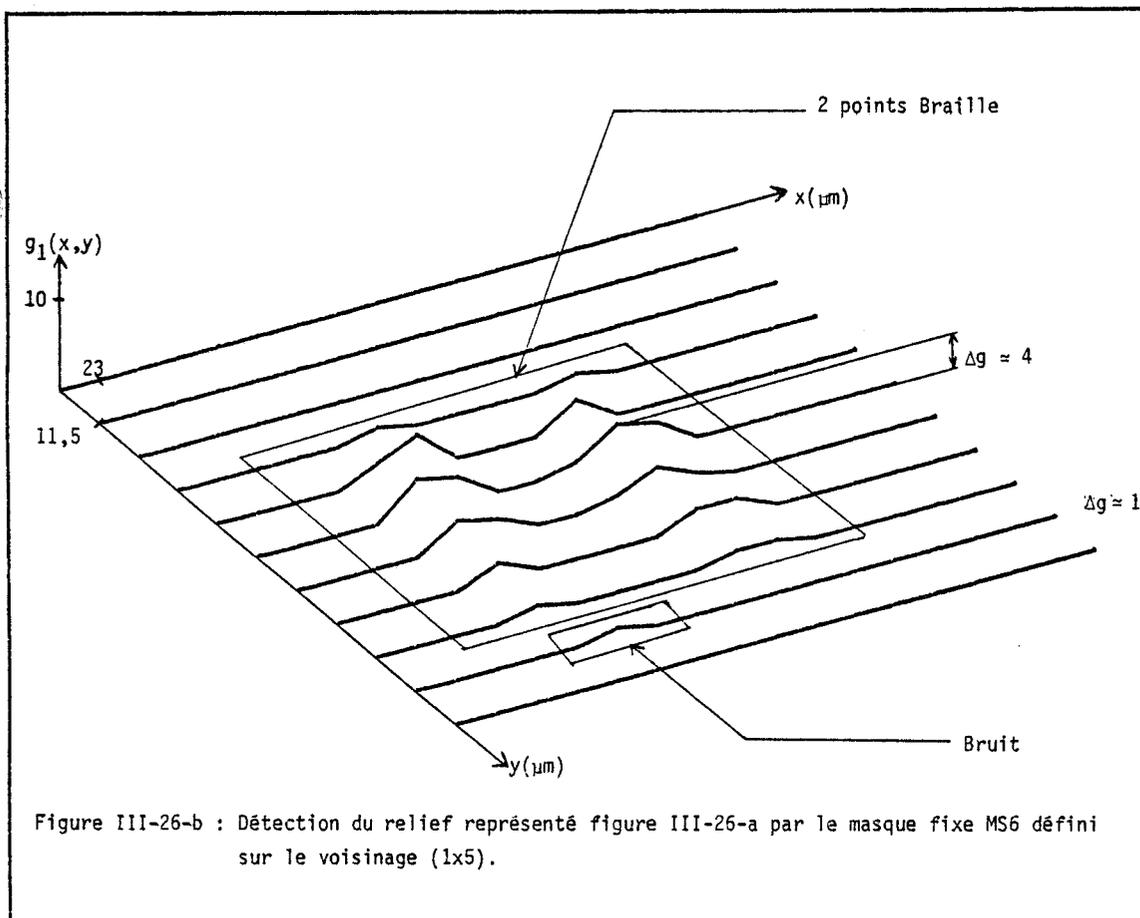
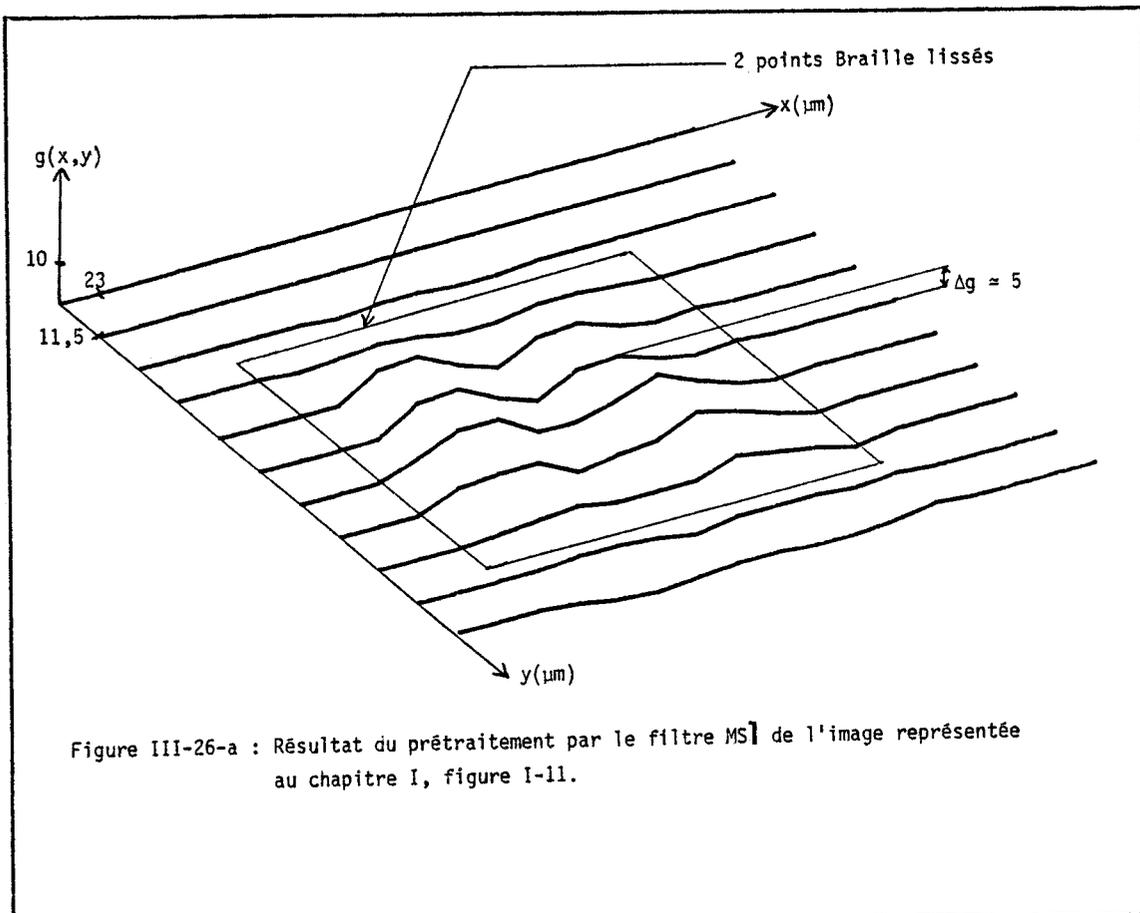


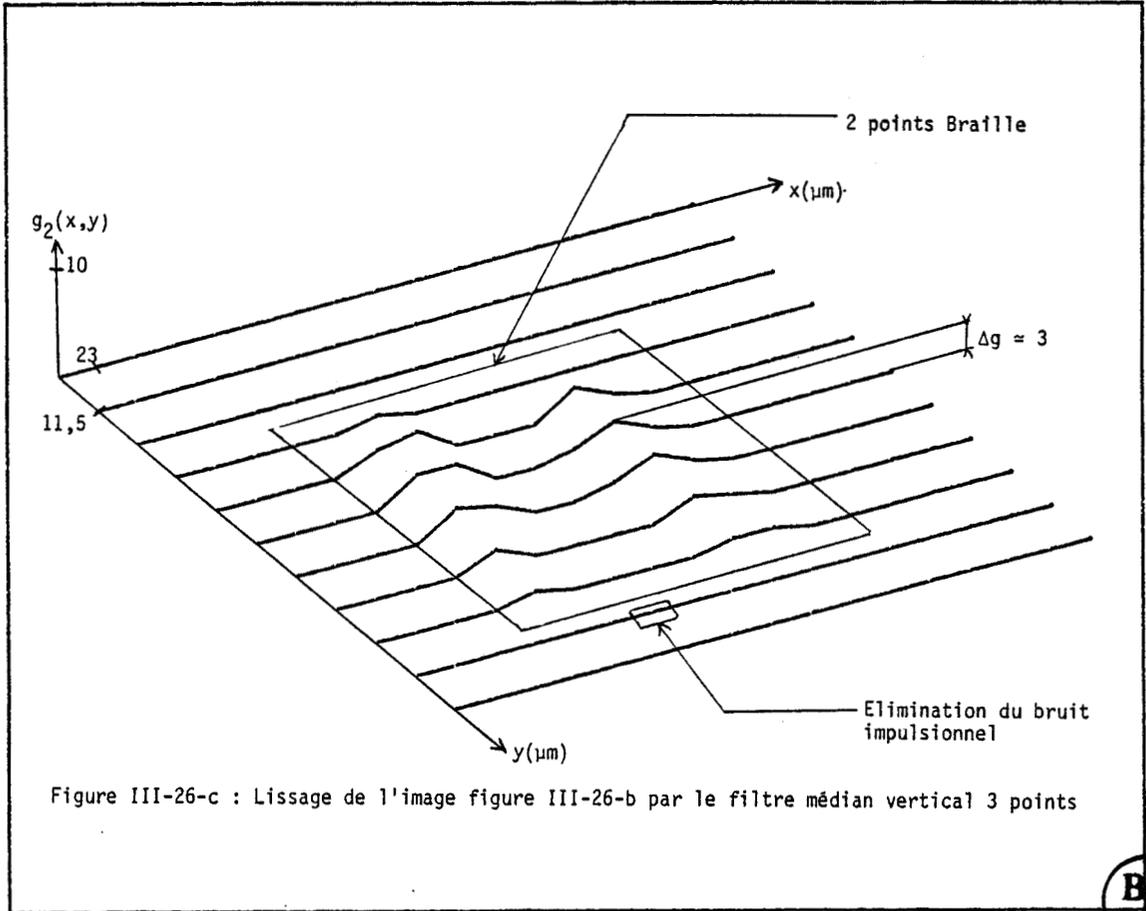


```

00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 01 01 00 00 00 00 00 00
00 00 00 00 00 00 00 00 01 01 00 00 00 00 00 00
00 00 00 00 00 00 00 00 01 01 00 00 00 00 00 00
00 00 00 00 00 00 00 00 01 01 00 00 00 00 00 00
00 00 00 00 00 00 00 00 01 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
    
```

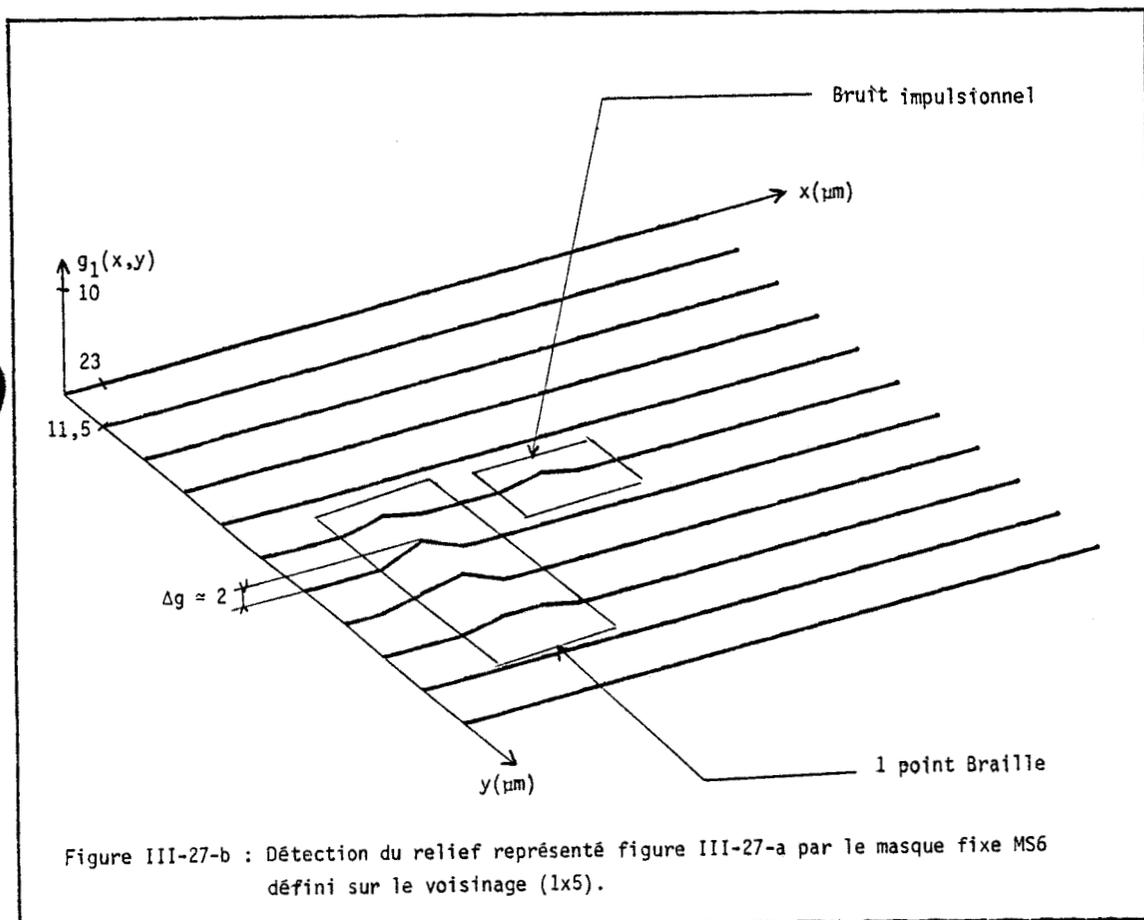
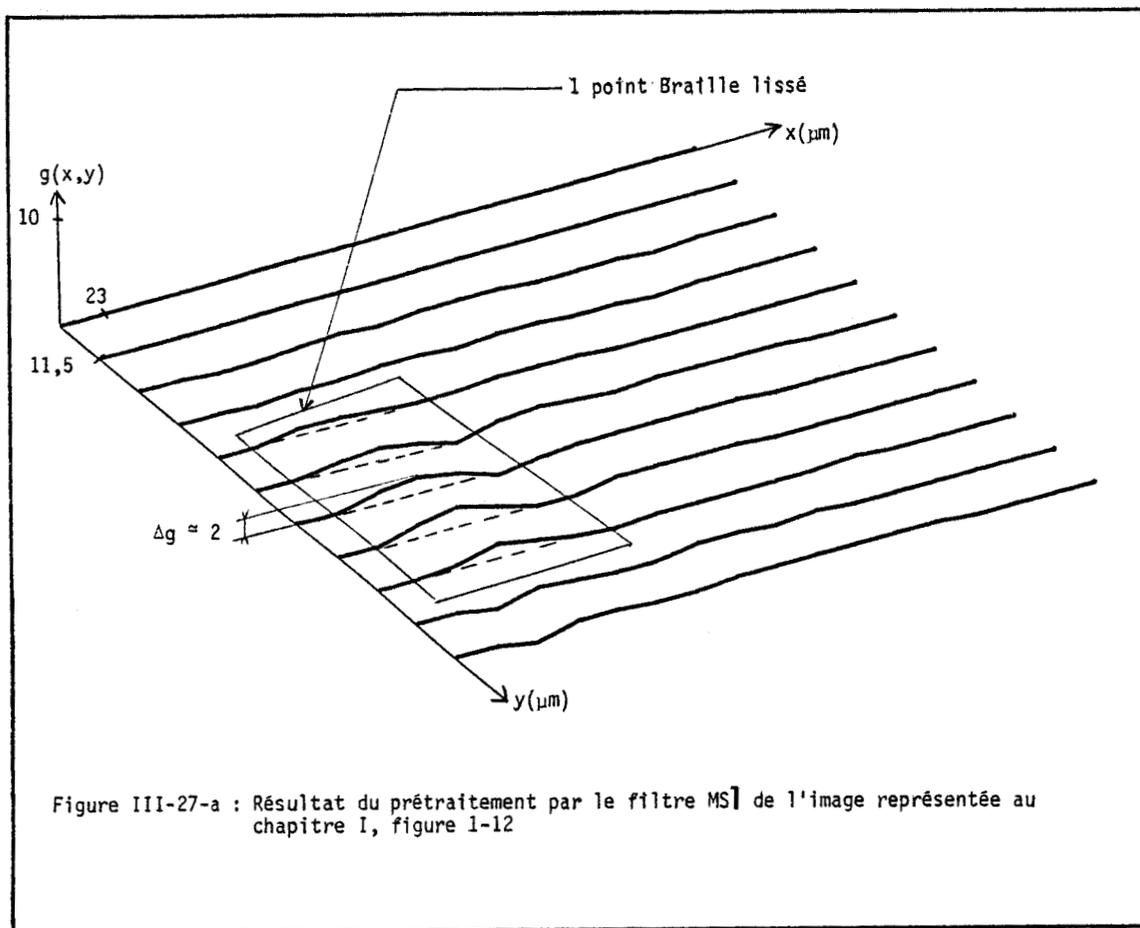
Figure III-25-d : Relief figure I-9 reconnu





```
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 01 00 00 00 00 00 00 00 00 00 00
00 00 00 00 01 01 00 00 00 01 01 00 00 00 00 00
00 00 00 00 01 01 00 00 01 01 01 00 00 00 00 00
00 00 00 00 01 01 00 00 01 01 01 00 00 00 00 00
00 00 00 00 01 00 00 00 00 01 01 00 00 00 00 00
00 00 00 00 01 00 00 00 00 01 01 00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
```

Figure III-26-d : Relief figure I-11 reconnu



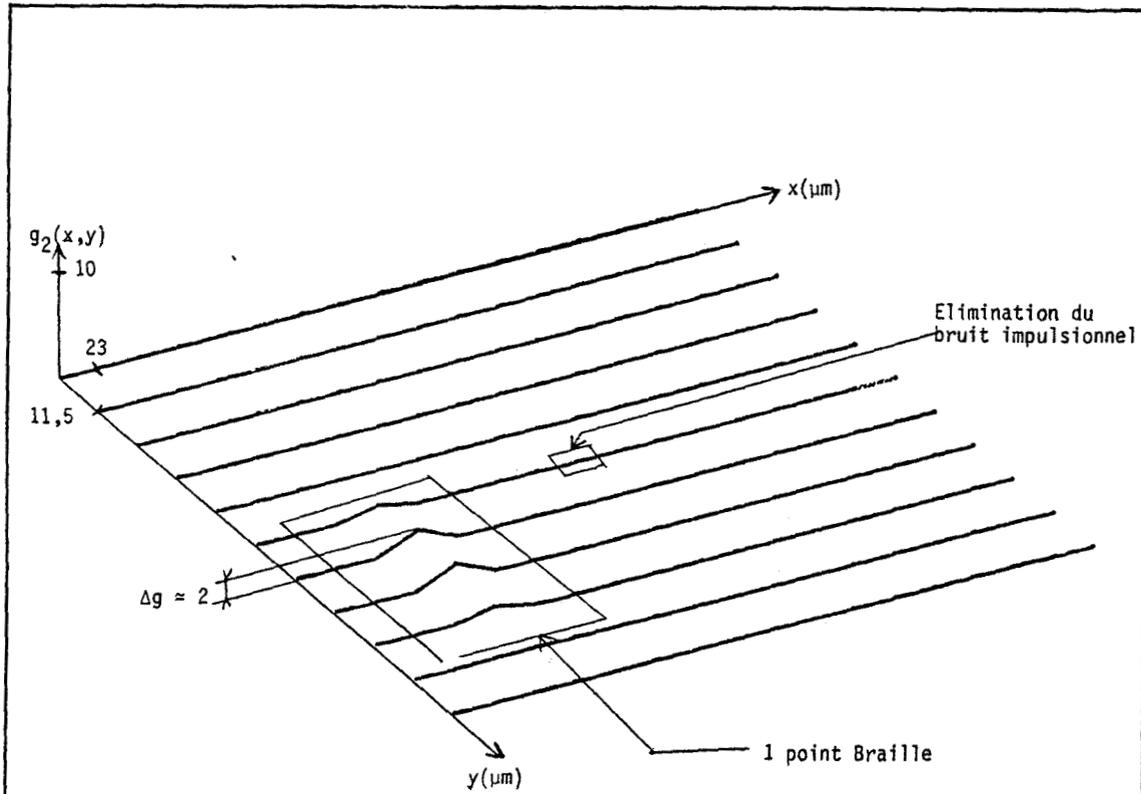


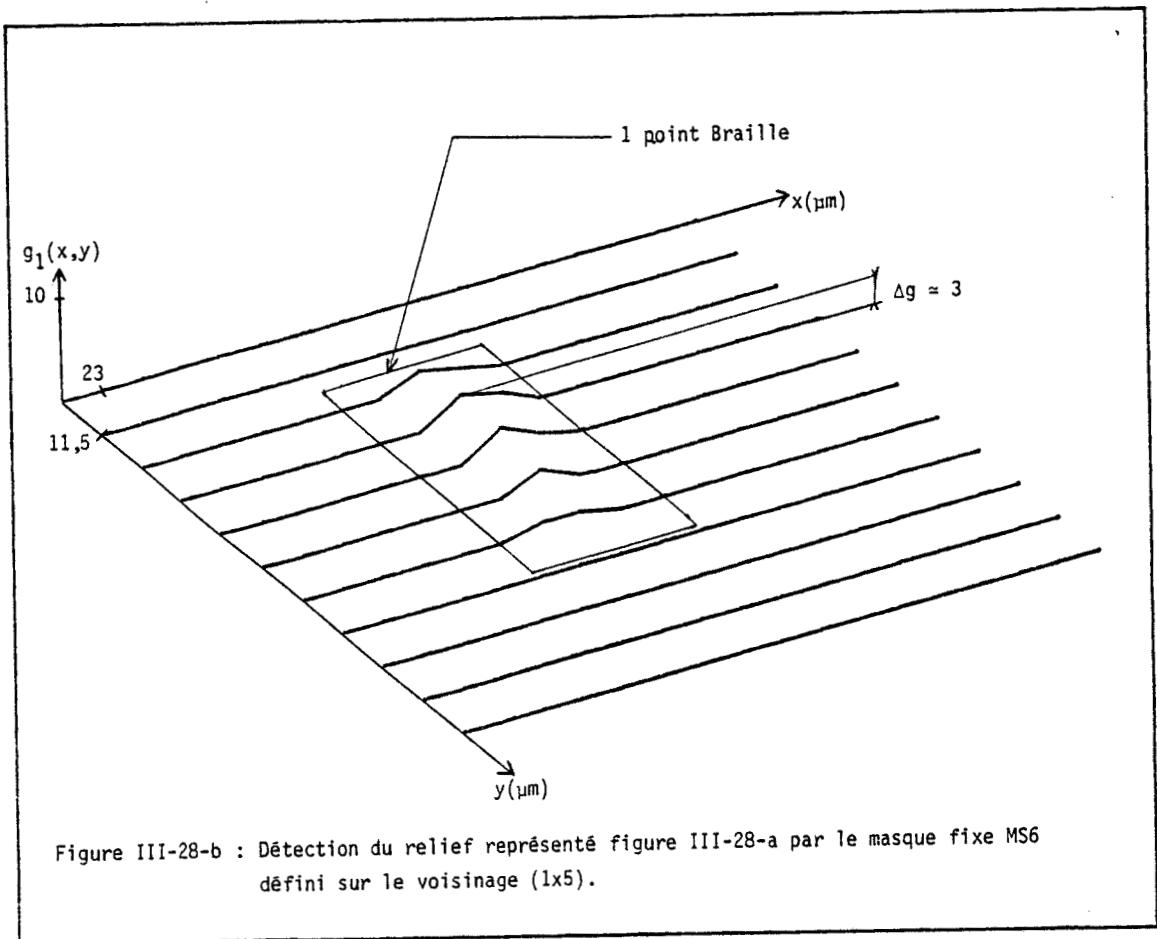
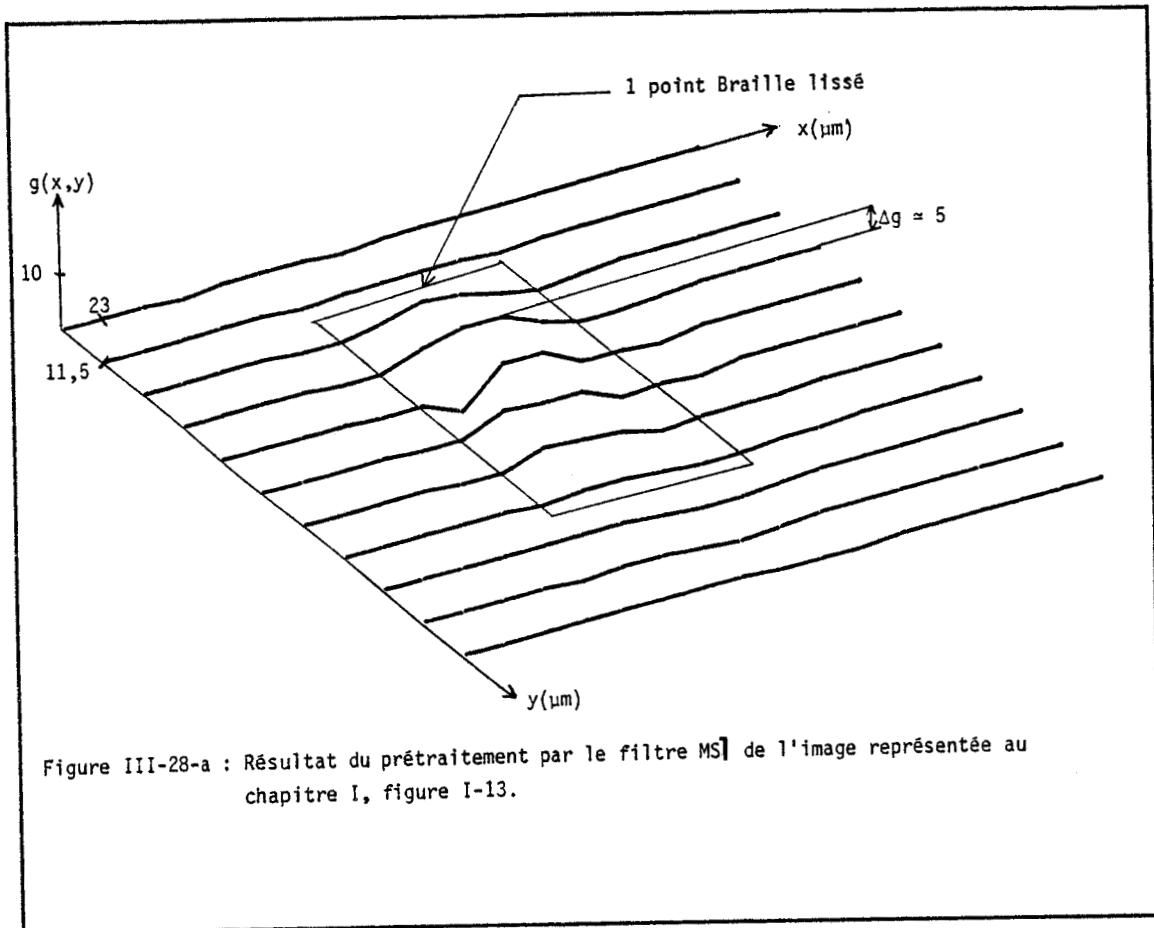
Figure III-27-c : Lissage de l'image figure III-27-b par le filtre médian vertical 3 points

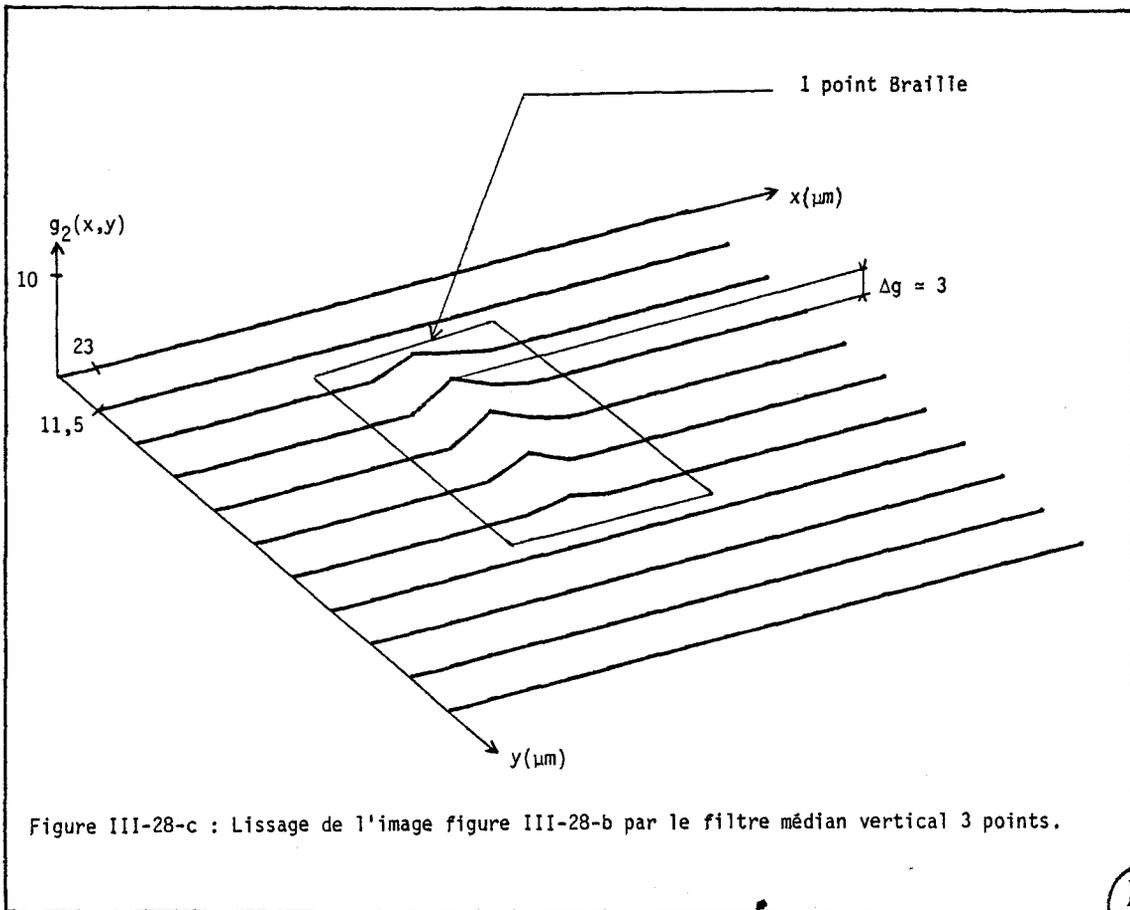


```

00-00-00-00-00-00 00-00-00-00-00-00 00 00 00 00-
00-00-00-00-00-00-00-00-00-00-00-00-00 00 00-00-
00-00-00-00-00-00-00-00-00-00-00-00-00 00 00-00-
00-00-00-00-00-00 00-00-00-00-00-00-00 00 00 00-00-
00-00-00-00-00-00-00-00-00 00 00-00 00 00 00 00-
00-00-00 01 00 00 00-00-00-00-00-00-00 00 00 00-
00-00-00 01 00 00 00-00-00-00 00 00-00 00 00 00-
00 00-00 01 00 00 00 00-00 00 00 00 00 00 00 00-
00-00-00 00 00 00 00-00-00 00 00 00 00 00 00 00-
    
```

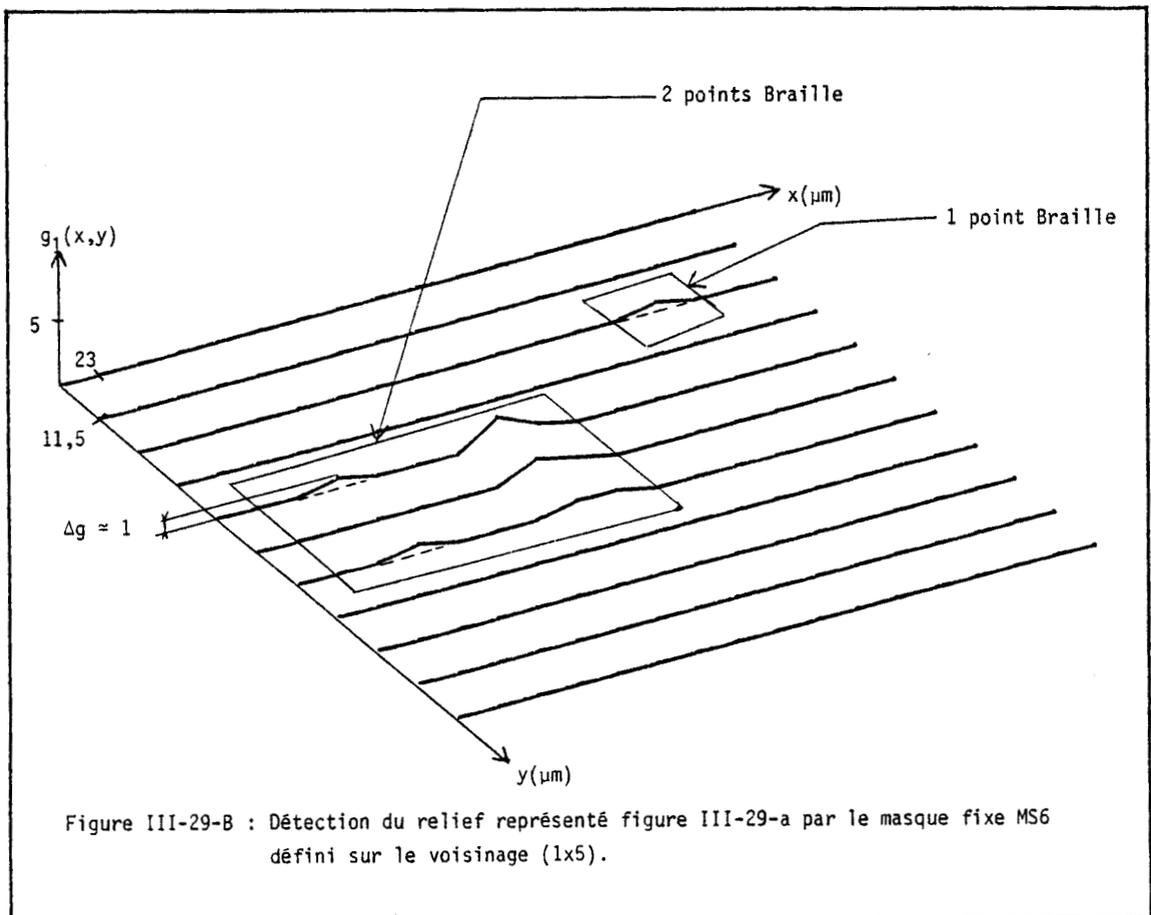
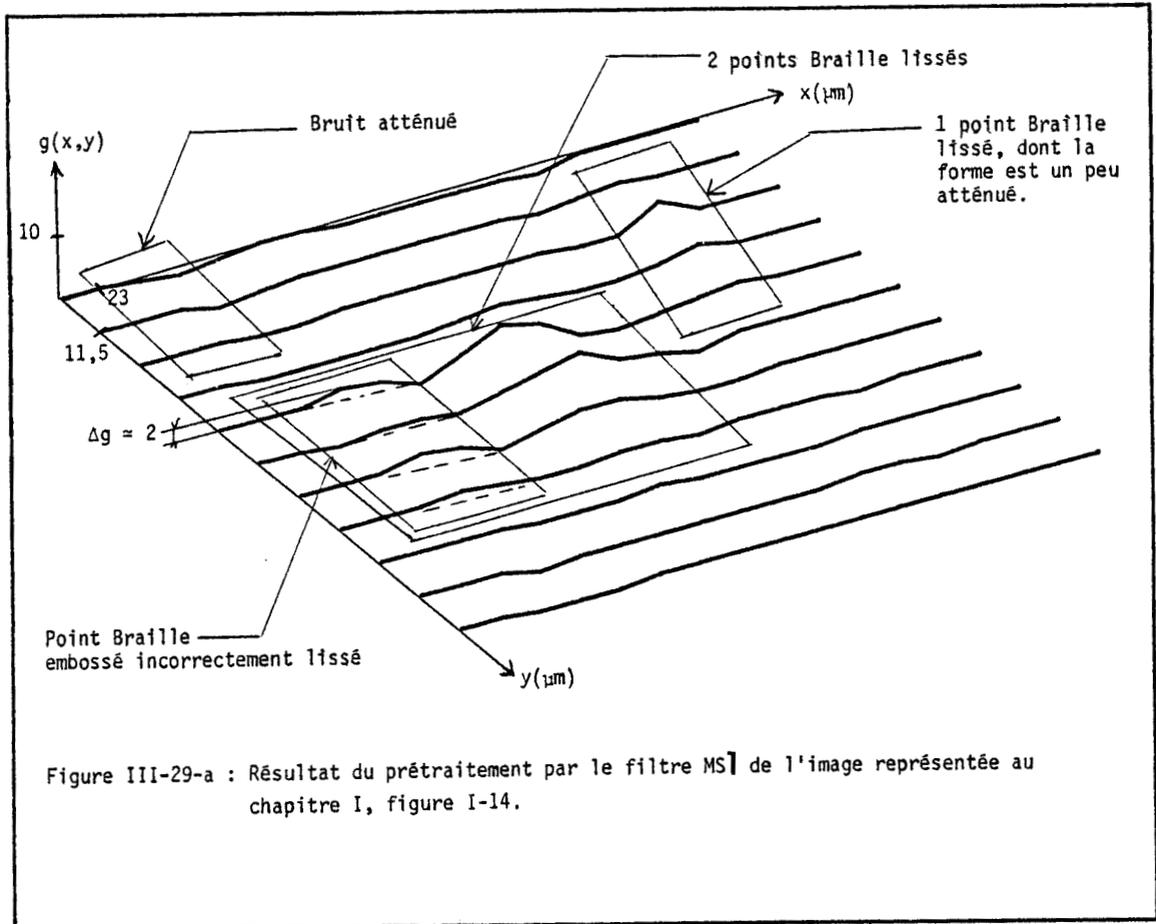
Figure III-27-d : Relief figure I-12 reconnu





```
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 01 01 00 00 00 00 00 00 00
00 00 00 00 00 00 00 01 01 00 00 00 00 00 00 00
00 00 00 00 00 00 00 01 01 00 00 00 00 00 00 00
00 00 00 00 00 00 00 01 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 01 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
```

Figure III-28-d : Relief figure I-13 reconnu



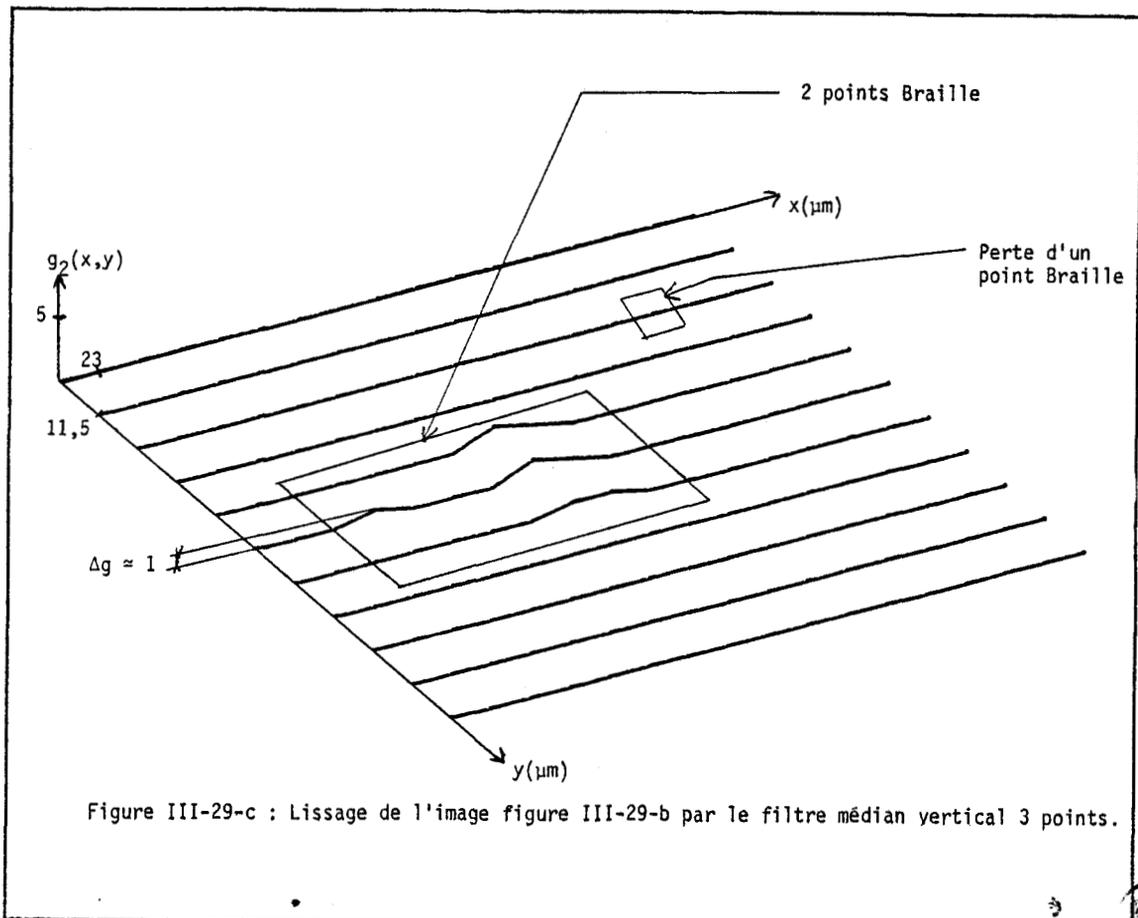


Figure III-29-c : Lissage de l'image figure III-29-b par le filtre médian vertical 3 points.

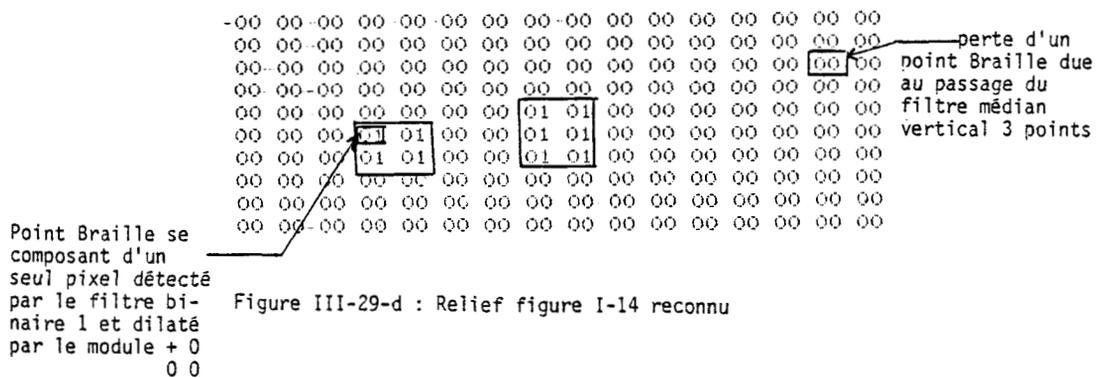


Figure III-29-d : Relief figure I-14 reconnu

Cependant, pour pouvoir estimer l'efficacité de détection du relief de la méthode de filtrage ci-dessus, nous avons effectué une étude statistique permettant de définir le taux de réussite réel d'extraction du relief.

Elle consiste à appliquer la méthode de filtrage représentée figure 30 aux caractères Braille du tableau "1" (chapitre I) appartenant aux deux procédés de fabrication du relief (manuel et mécanique).

Nous définissons le taux réel de reconnaissance du relief des deux procédés de fabrication par :

$$T_{RR} = \frac{\text{Nombre total de caractères Braille reconnus}}{\text{Nombre total de caractères Braille}}$$

Cette étude statistique a été menée dans les mêmes conditions que celle du chapitre I, sauf que nous nous sommes intéressés seulement à la détection du relief qui est liée au filtrage.

Les résultats que nous présentons dans le tableau 2 ne tiennent pas compte des défauts d'alignement. Les défauts d'alignement étant liés à la reconnaissance des caractères Braille.

Procédé de fabrication du relief	Nombre total d'échantillons (graphème)	Taux d'extraction du relief réel $T_{RR}$
"Embossage mécanique"	$N_T = 2834$	99,14 %
"Embossage manuel"	$N_T = 3256$	95,21 %

Tableau 2 : Taux de réussite d'extraction du relief Braille de la méthode de détection du relief représentée figure 30.

La méthode de filtrage présentée ci-dessus permet la détection de l'information relief avec un taux d'extraction de caractères Braille voisin de 99 % dans le cas d'embossage "mécanique" et voisin de 95 % dans le cas d'embossage "manuel".

Dans ces conditions, il est probable qu'après transcription en noir, l'analyse du contexte permette de restituer les 100 % du document. Nous nous sommes bornés à cette démarche sans aborder celle de l'analyse du contexte.

#### VII-4- ROBUSTESSE DES MASQUES FIXES FACE AU BRUIT

La méthode des moindres carrés a été présentée sous forme déterministe sans se préoccuper des propriétés statistiques du bruit. Lorsque le bruit est gaussien, de moyenne nulle, de variance  $\sigma^2$  et indépendant des variables  $(x,y)$ , on peut caractériser dans le plan statistique l'estimé  $\hat{b}$ . Pour cela, il suffit de calculer la variance de l'estimation  $\hat{b}$ .

Pour s'assurer de l'efficacité et la robustesse du masque en face du bruit, nous allons calculer le taux de réduction du bruit. Nous supposons que la répartition du bruit suit une loi normale  $N(0, \sigma^2)$ . Dans ce cas, on peut démontrer que la variance de la sortie modèle  $\hat{z}$  est donnée par<sup>[38]</sup>:

$$\hat{\sigma}^2 = h^T [U^T U]^{-1} h \sigma^2 \quad (32)$$

Pour le masque (1x5), la matrice U et le vecteur d'entrée du pixel central sont donnés par (10) et (12). En remplaçant (10) et (12) dans (32), nous obtenons :

$$\hat{\sigma}^2 = 0,49 \sigma^2 \quad (33)$$

Ainsi, le taux de réduction du bruit du masque (1x5) est de 0,49. Des calculs semblables donnent respectivement pour les masques (3x3) et (5x5) :

$$\hat{\sigma}^2 = 0,56 \sigma^2 \quad (34)$$

et

$$\hat{\sigma}^2 = 0,28 \sigma^2 \quad (35)$$

L'impact modeste du masque (1x5) sur le bruit est dû au faible nombre de degrés de liberté (N-K), avec N le nombre d'observations et K le nombre de paramètres. Il faut insister sur le fait que (32) est valable seulement lorsque le modèle supposé (2) est correct. Si ce modèle est inadapté aux données, l'écart type  $\hat{\sigma}$  peut augmenter considérablement. Dans le cas où la taille du masque de régression est plus grande que la surface à détecter, le modèle du second degré n'est plus applicable. Cette remarque trouve son importance lorsqu'on utilise un grand masque pour parvenir à une bonne suppression du bruit.

Dans le cas du relief Braille, le voisinage (1x5) de définition du modèle (2) est le mieux adapté à la forme de l'information Braille. Il permet la détection du relief avec un très bon taux de réussite et une diminution du bruit appréciable.

L'estimation  $\hat{b}$  étant une variable aléatoire, il ne faut pas s'attendre à ce qu'elle soit égale au véritable modèle b.

#### VIII- CONCLUSION

Dans ce paragraphe, nous avons présenté une méthode destinée à détecter les objets décrits par une fonction polynômiale définie sur le voisinage (1x5) altérée par du bruit.

Dans un premier temps, nous avons mis en valeur les surfaces décrites par des polynômes du second degré. Ensuite, nous avons détecté les objets et effectué deux opérations qui consistent à lisser, fermer les trous des objets présentant des discontinuités dans les formes et dilater les points Braille très faiblement contrastés.

La phase de mise en valeur et la phase de détection sont effectuées au moyen de masques de régression fixes. Le masque de mise en valeur supprime le bruit. Les segments du second degré sont laissés intacts. La sortie du filtre de détection est positive lors de la présence d'un point Braille.

Cette méthode de filtrage permet la détection du relief avec un taux de réussite d'extraction de caractères Braille voisin de 99 % dans le cas d'embossage "mécanique" et voisin de 95 % dans le cas d'embossage "manuel".

Nous rappelons aussi qu'il existe une approche de type "syntaxique" dans laquelle la sémantique des images à traiter peut s'introduire par les règles de réécriture<sup>[29][30][31][32]</sup>. Ces méthodes vont trouver l'efficacité de leur application surtout dans la partie reconnaissance des caractères Braille. La partie reconnaissance de caractères Braille n'est pas abordée dans ce mémoire.

Dans le quatrième chapitre, nous allons décrire les différentes particularités matérielles du nouveau transcripteur Braille-noir.

CHAPITRE IV : PARTICULARITÉS DE L'ARCHITECTURE DU NOUVEAU  
TRANSCRIPTEUR BRAILLE/NOIR.

I - INTRODUCTION

Le but de notre réalisation est de concevoir un système capable de remplacer le lecteur optique de relief Braille<sup>[3][4][5]</sup>, fait d'un capteur monté à la place de la plume d'une table traçante et son architecture microinformatique associée, par un système d'acquisition constitué d'une caméra CCD et d'une mémoire image, le tout piloté par ordinateur (figure IV-1).

Nous décrivons dans ce chapitre les particularités de la conception et de la réalisation matérielle d'un prototype capable en particulier de répondre aux nécessités de fonctionnement de l'automate de transcription Braille/noir.

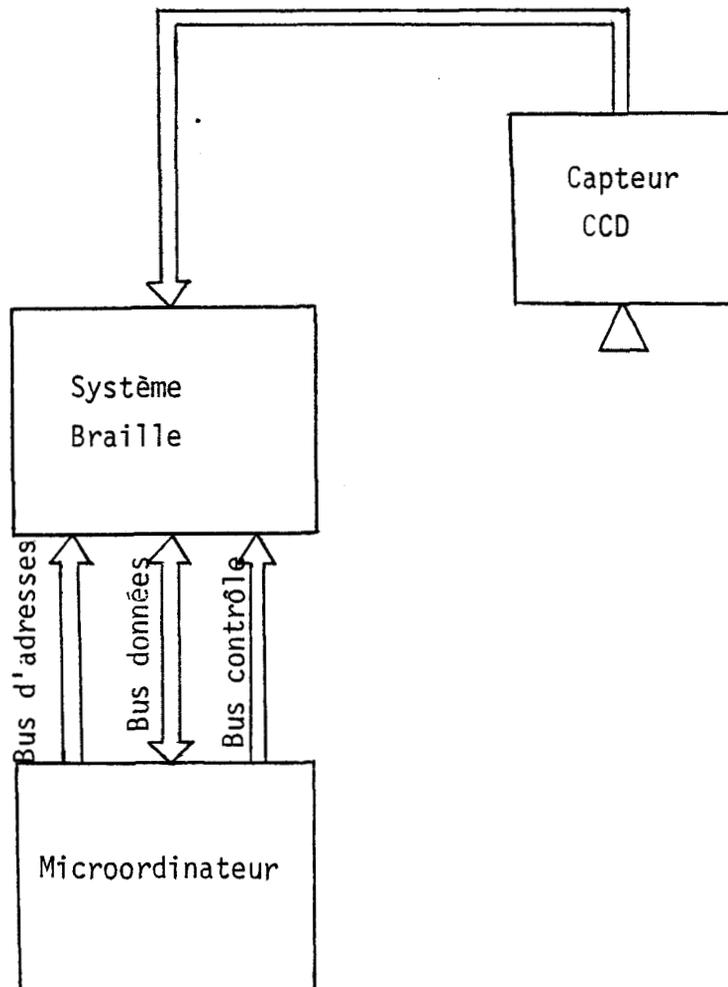


Figure IV-1

## II - ETUDE DE LA FONCTION FENETRAGE ET SYNCHRONISATION

---

Le capteur CCD matriciel de la caméra restitue un signal analogique échantillonné de 576 lignes chacune comportant 384 points élémentaires. Cette caméra fournit deux trames entrelacées comportant au total 221184 pixels, chaque pixel étant codé sur 6 bits.

La fréquence d'arrivée des pixels est rythmée par l'horloge point  $F_p$  de 7,375 MHz. Chaque pixel, numérisé sur 6 bits, est donc présent sur le bus de sortie de la caméra pendant un temps approximatif de 135 ns.

Pour mémoriser l'image formée des deux trames entrelacées, il faut une capacité mémoire équivalente à 256 K mots de 6 bits ; cette image est stockée en 40 ms (standard CCIR, 25 images/seconde).

Comme le traitement du relief Braille va être effectué ligne par ligne, nous avons choisi de limiter à 32 K octets ou 64 K octets la capacité mémoire, ce qui va permettre l'acquisition de 3 ou 6 lignes Braille.

Ce choix permet de disposer cette mémoire de stockage à la place de l'extension mémoire d'un microordinateur de pilotage. Il nécessite d'un microordinateur d'acquérir l'image et de traiter l'information par zone limitée à trois lignes Braille, ce qui ne constitue pas un inconvénient majeur.

La fonction fenêtrage et synchronisation a pour but de définir une fenêtre de visualisation dans l'image totale. Cette fenêtre devra limiter l'acquisition en mémoire d'un nombre de pixels allant de 0 à 32768 points (le maximum étant de 32 K octets pour la mémoire image) ; la fenêtre pourra être déplacée sur l'image totale afin de mémoriser tous les pixels de l'image à traiter.

Le fenêtrage consiste donc à déterminer les dimensions d'un rectangle défini par 4 paramètres d'entrée (figure IV-2) ; ce rectangle correspondra à une partie de l'image ne pouvant pas excéder les limites de la mémoire image fixées à 32 K octets.

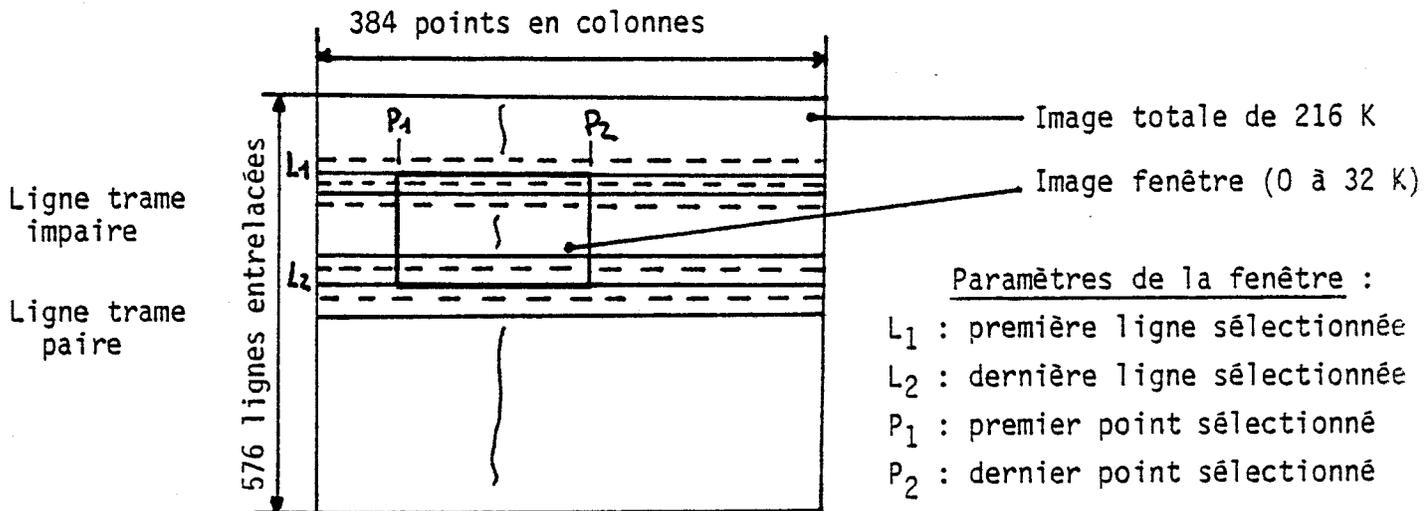


Figure IV-2

Pour la commande de sélection de la fenêtre image et prétraitement caméra, nous utilisons un interface se composant de deux PIA (Peripheral interface Adaptater-Processseur 6809 et 68000).

La figure IV-3 représente l'architecture de la carte fenêtrage, son interface et les chronogrammes correspondants.

### III - ETUDE ET REALISATION DE LA MEMOIRE IMAGE

La mémoire image a la structure représentée figure IV-4. Elle comporte 3 accès possibles :

- un accès direct en lecture pour la mémorisation des données provenant de la caméra,
- un accès permanent en lecture par le système de rafraîchissement de mémoire et recomposition du signal vidéo pour la visualisation des données mémorisées,
- un accès par le microordinateur de traitement en écriture et en lecture

La commutation de chacun de ces accès est assurée par l'interface repère 2 de fenêtrage et prétraitement.



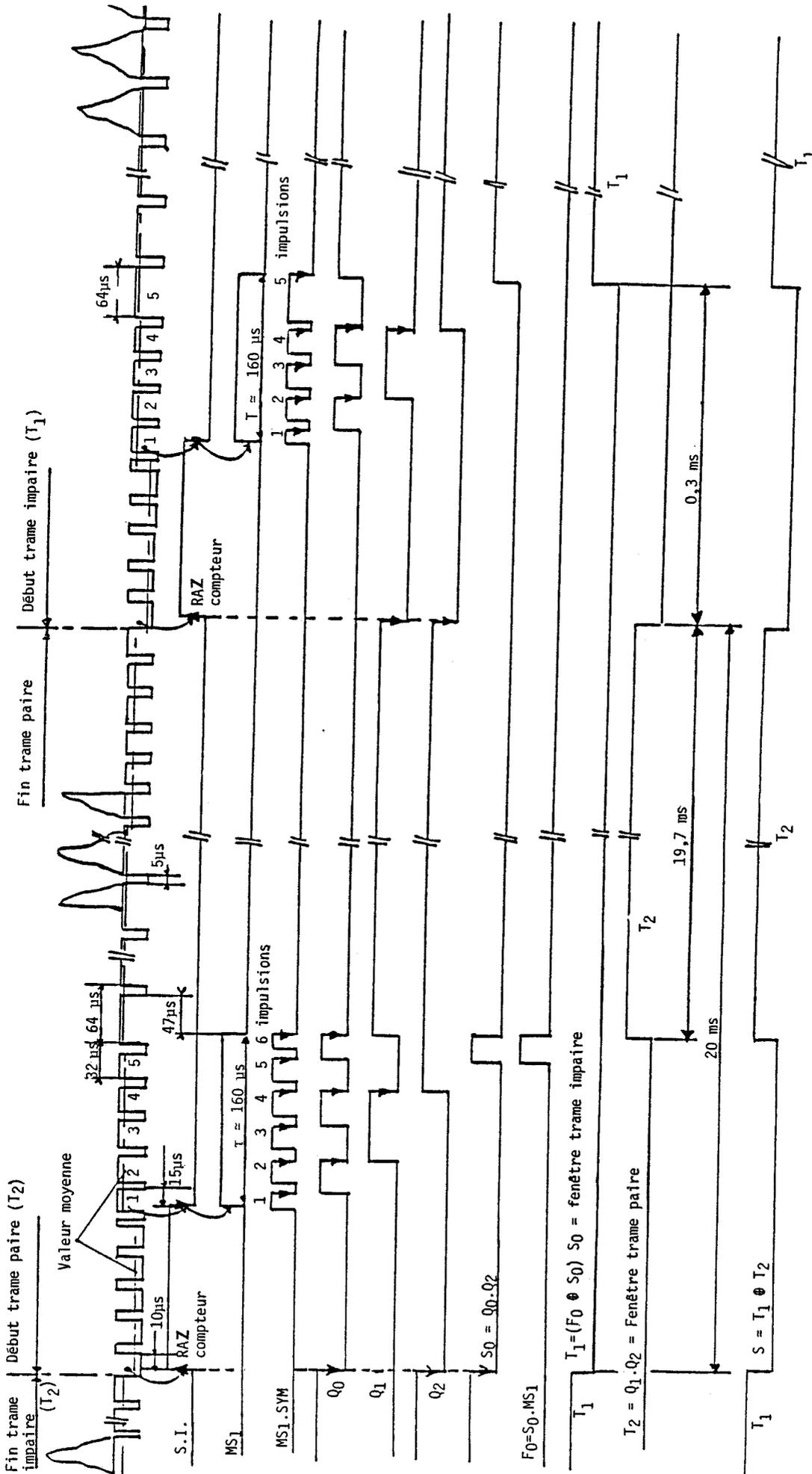


Figure IV-3-b : Chronogramme réalisant l'enveloppe de la trame impaire T1 et de la trame paire T2.



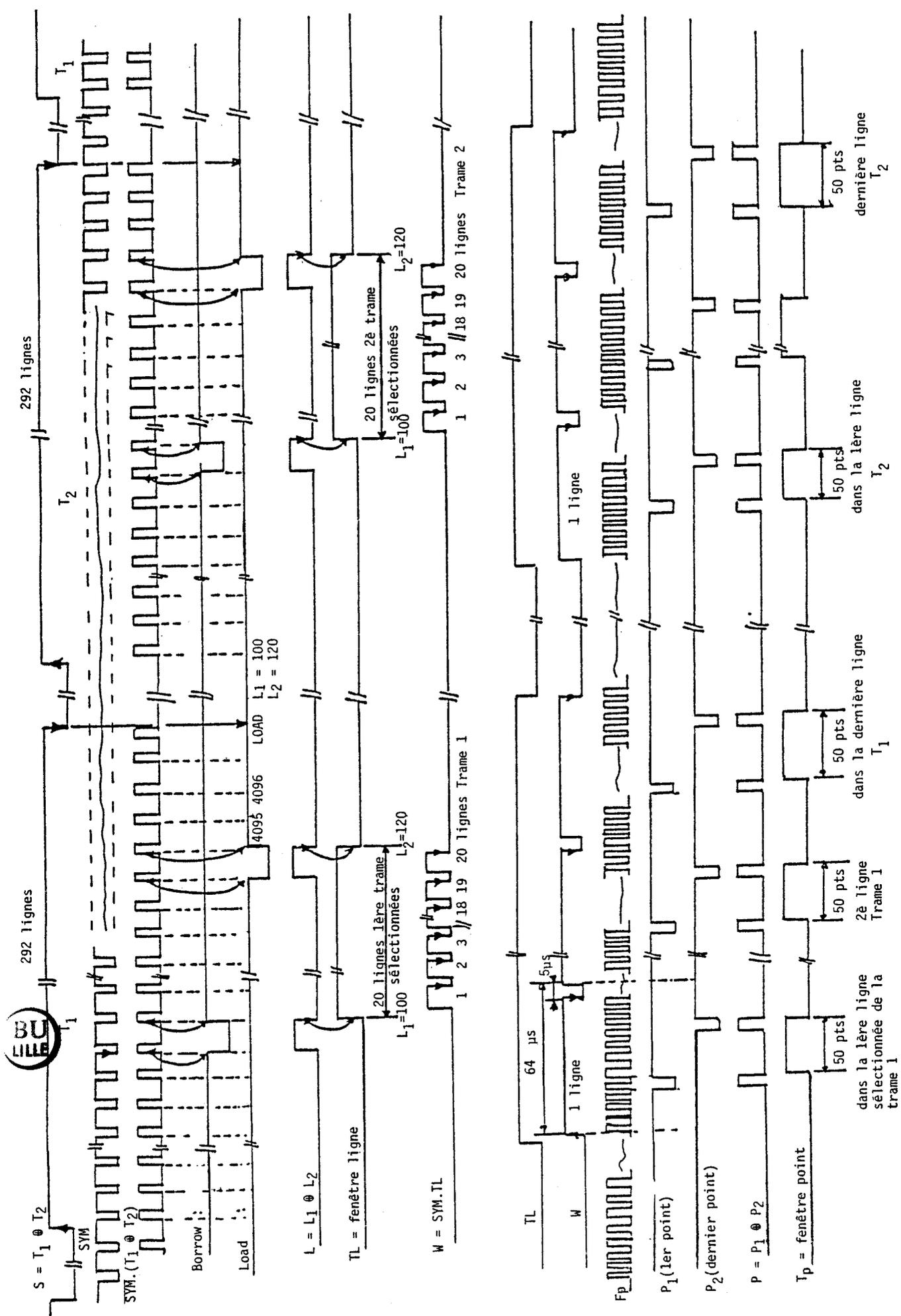


Figure IV-3-c : Chronogramme réalisant la fenêtre ligne (TL) et la fenêtre point (Tp)

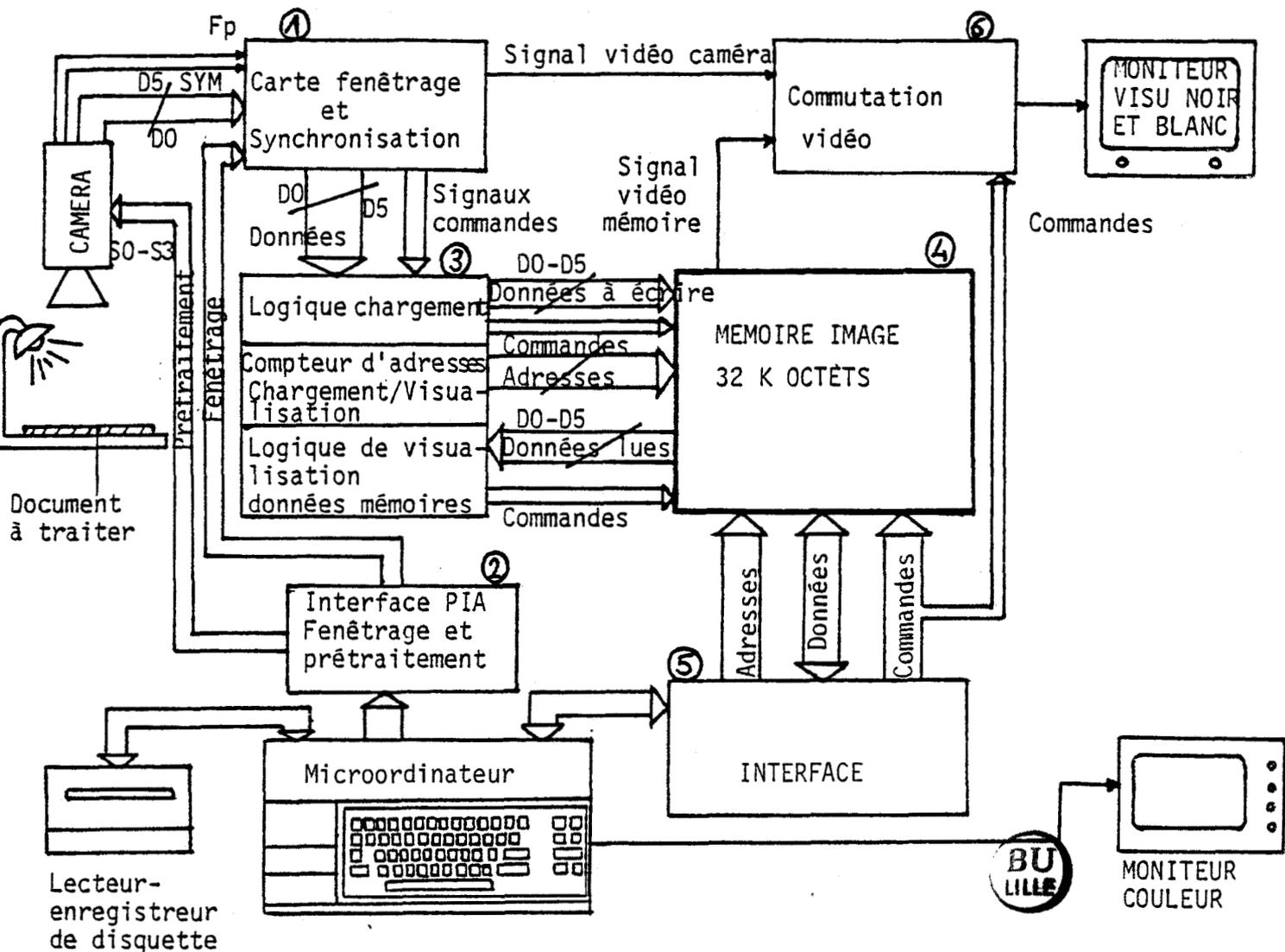


Figure IV-4

La fonction fenêtrage et synchronisation, ainsi que les commandes de prétraitement envoyées à la caméra sont assurées par les ensembles 1 et 2 .

Sur ce schéma fonctionnel apparaît deux sous-ensembles liés (repère 3) qui assurent les fonctions de chargement et de visualisation de l'image mémorisée et utilisent en commun le même système d'adressage.

Cette mémoire dont un prototype a été étudié et réalisé comporte un certain nombre de particularités que nous décrivons et qui permettent l'utilisation de composants électroniques courants.

L'étude qui suit analyse le fonctionnement de ces particularités.

### III-1- ANALYSE DES PARTICULARITES DE LA MEMOIRE IMAGE

#### III-1-1- Réalisation du signal chargement

Une image est constituée de deux trames entrelacées. Pour mémoriser l'image de la caméra, il faut écrire en mémoire tous les pixels contenus dans ces trames délivrées successivement (figure IV-5).

La trame impaire T1 est constituée des lignes impaires et la trame T2 est constituée des lignes paires.

Le chargement d'une image s'effectue pendant l'état 1 d'un signal "chgt" qui encadre l'émission de deux trames successives de la caméra.

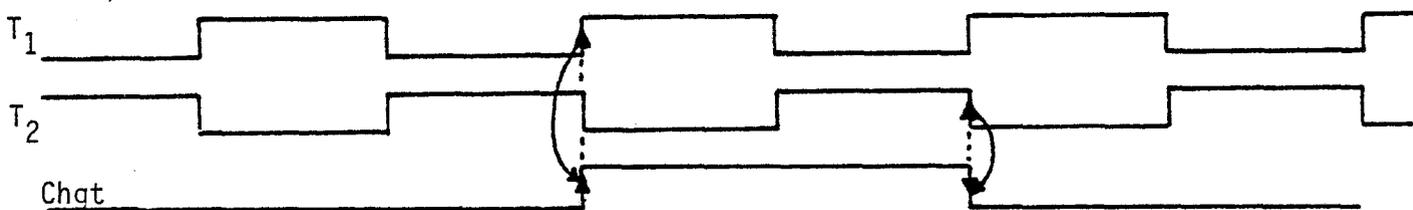


Figure IV-5 : signal chargement

Ce signal chgt est généré soit manuellement à l'aide d'un poussoir, soit par programmation à travers le bus de commande du microordinateur.

#### III-1-2- Problèmes liés au temps d'accès des mémoires par rapport à la fréquence d'arrivée des pixels

La fréquence point (Fp) de la caméra est de 7,375 MHz. Elle rythme, sur le front montant, l'arrivée de chaque pixel défini sur 6 bits (D0 à D5) (figure IV-6). Nous disposons donc de 135 ns pour mémoriser chaque pixel.

Or, le temps d'accès des mémoires courantes en 1985 est de l'ordre de 150 ns. Il est donc incompatible avec la fréquence d'arrivée des pixels qu'il faut mémoriser. La transformation série/parallèle schématisée ci-dessous a permis de résoudre ce problème.

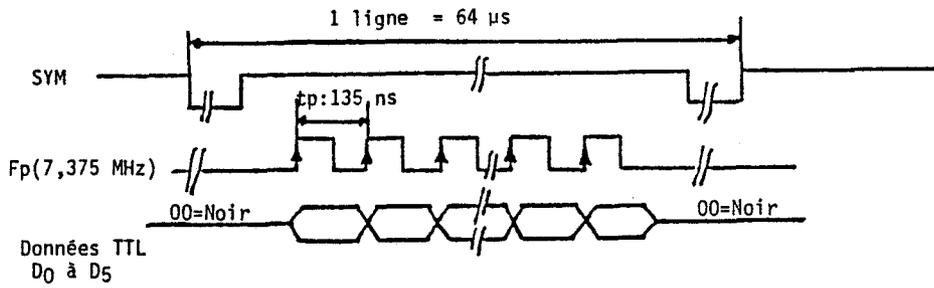


Figure IV-6 : Signaux fournis par la caméra

### III-1-3- Transformation série/parallèle des pixels

Pour s'affranchir du temps d'accès des mémoires, on réalise la structure de la figure IV-7. Chaque pixel est appliqué simultanément à 4 mémoires tampons à bascule D.

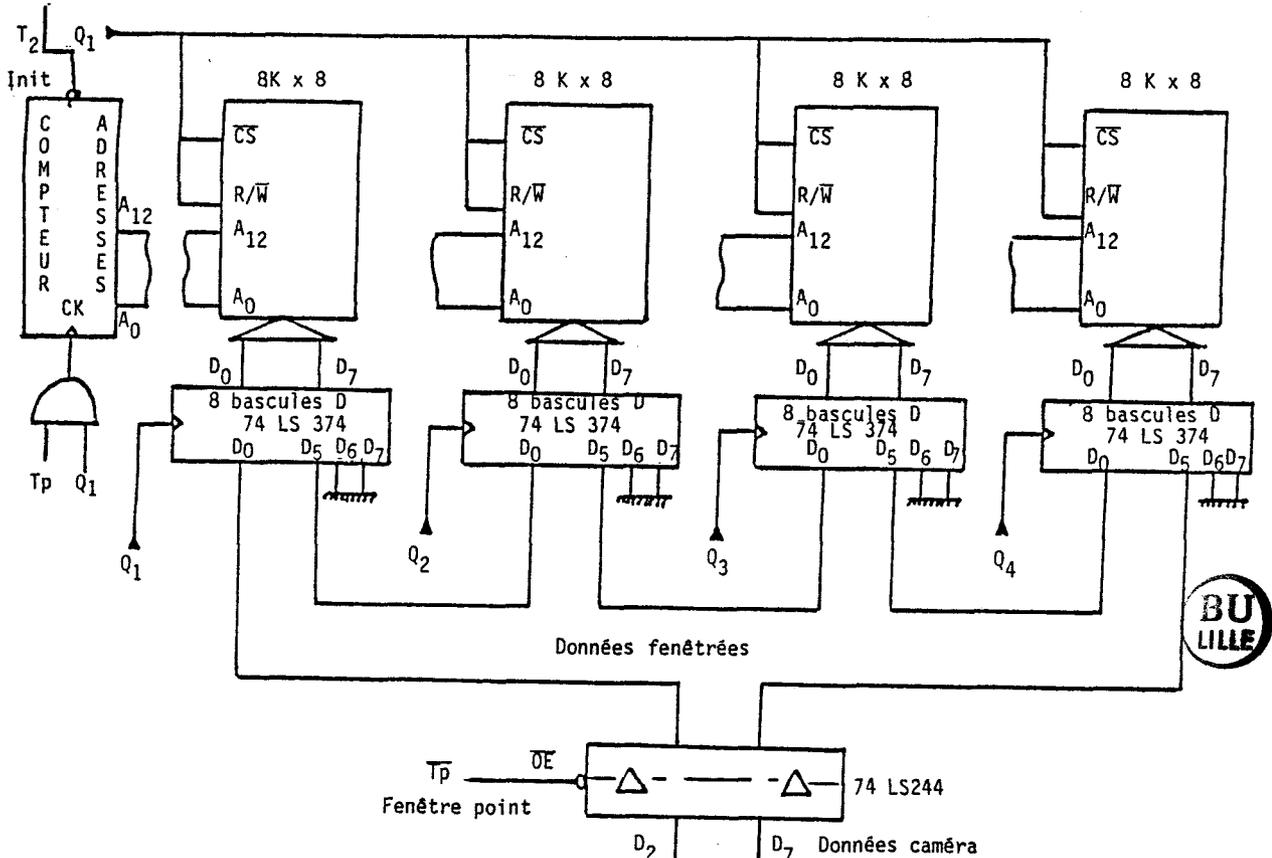


Figure IV-7 : Principe de la transformation série/parallèle des pixels afin de s'affranchir des temps d'accès des mémoires.

A partir de la fréquence point  $F_p$  de la caméra, nous réalisons une logique permettant de générer 4 horloges de fréquence  $F_p/4$  de même période et décalées entre elles d'une période  $t_p = \frac{1}{F_p} \approx 135$  ns (figure IV-8).

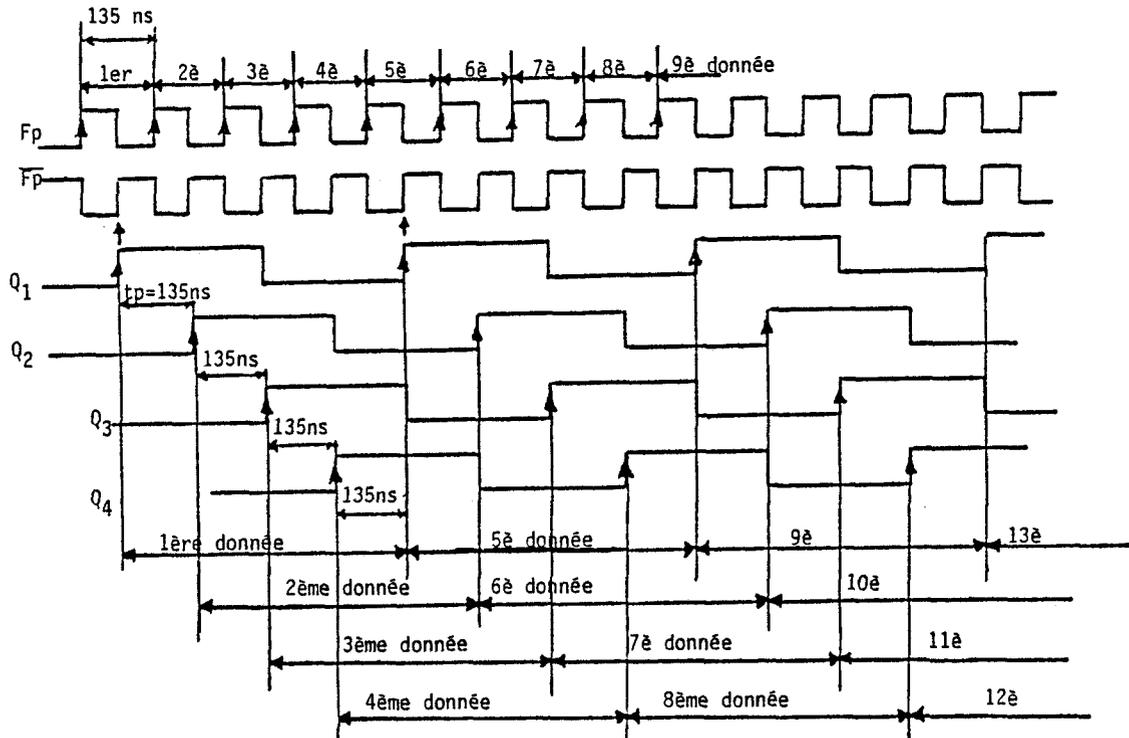


Figure IV-8 : Chronogramme des horloges  $Q_1$  à  $Q_4$  de mémorisation des données dans les 4 mémoires tampons.



Les signaux  $Q_1$ ,  $Q_2$ ,  $Q_3$ ,  $Q_4$  étant ainsi créés, il faut que chaque front montant de ces horloges valide la mémorisation de quatre pixels, prélevés tour à tour avec un retard  $t_p$ , dans quatre mémoires tampons à bascule D dont les sorties sont appliquées aux entrées  $D_0$  à  $D_7$  de chacune des quatre mémoires 8 K octets (figure IV-7).

Chaque pixel est maintenant présent en sortie de chaque mémoire tampon pendant un temps équivalent à  $4 \times t_p \approx 540$  ns (supérieur aux temps d'accès des mémoires qui est de l'ordre de 150 ns). La mémorisation des quatre mots de 6 bits est validée par le signal  $Q_1$  comme le montre la figure IV-9.



Pour que le microprocesseur puisse accéder en lecture et en écriture à chaque pixel consécutif d'une trame en incrémentant son adresse, il suffit de décoder les bits d'adresses poids faibles A0 et A1 et de créer les quatre signaux de sélection qui seront appliqués à leur plan mémoire respectif.

Les bits A2 à A12 du microordinateur seront respectivement appliqués aux bits A0-A10 des quatre plans mémoires (figure IV-11).

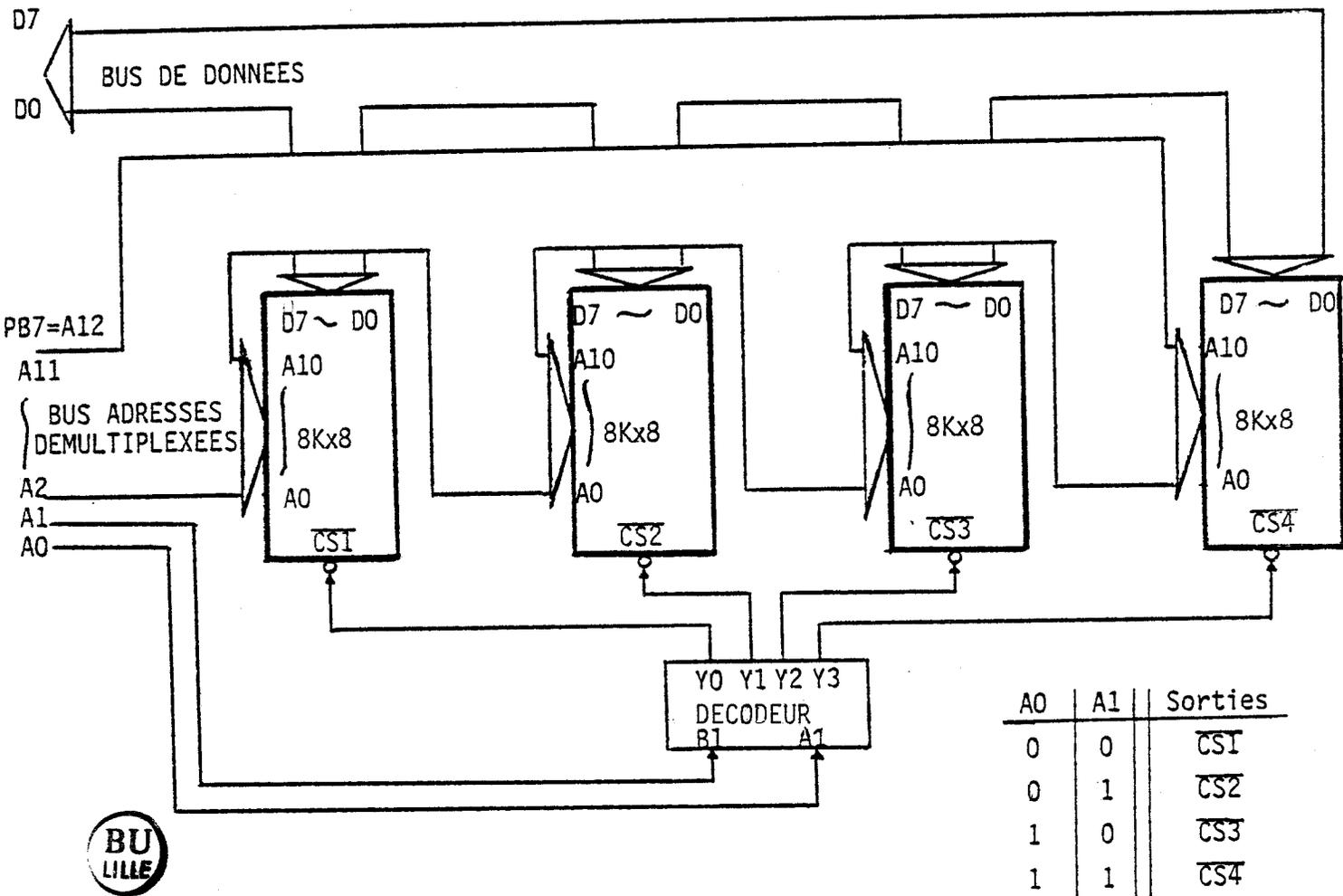


Figure IV-11 : Schéma de principe de l'accès du microordinateur aux données de la mémoire image.

Ce procédé d'adressage particulier permet de réduire considérablement le temps d'exécution des logiciels de traitement de l'image. Le nombre de commutations de banques mémoire étant réduit au minimum lors du traitement effectué par "le processeur d'entrelacement" des lignes par exemple.

Nous disposons donc en mémoire d'une image "désentrelacée". Avant de pouvoir traiter cette image, il faut "entrelacer" les lignes afin d'obtenir les pixels dans l'ordre voulu. Cette opération est effectuée par logiciel.

### III-1-4- Etude et réalisation de la visualisation de la mémoire image

Pour la mémorisation des données, une transformation série/parallèle a été effectuée. Les données fournies par la caméra arrivant périodiquement tous les 135 ns, cette transformation a permis de réduire l'afflux des données à chaque plan mémoire à un rythme de 540 ns. Les données ont été ainsi aiguillées vers quatre plans mémoire.

Pour la visualisation, le problème posé est le multiplexage des données lues dans la mémoire image afin de les retrouver ordonnées avec une période d'apparition de 135 ns. Le principe de visualisation de l'image mémorisée est représenté par la figure IV-12.

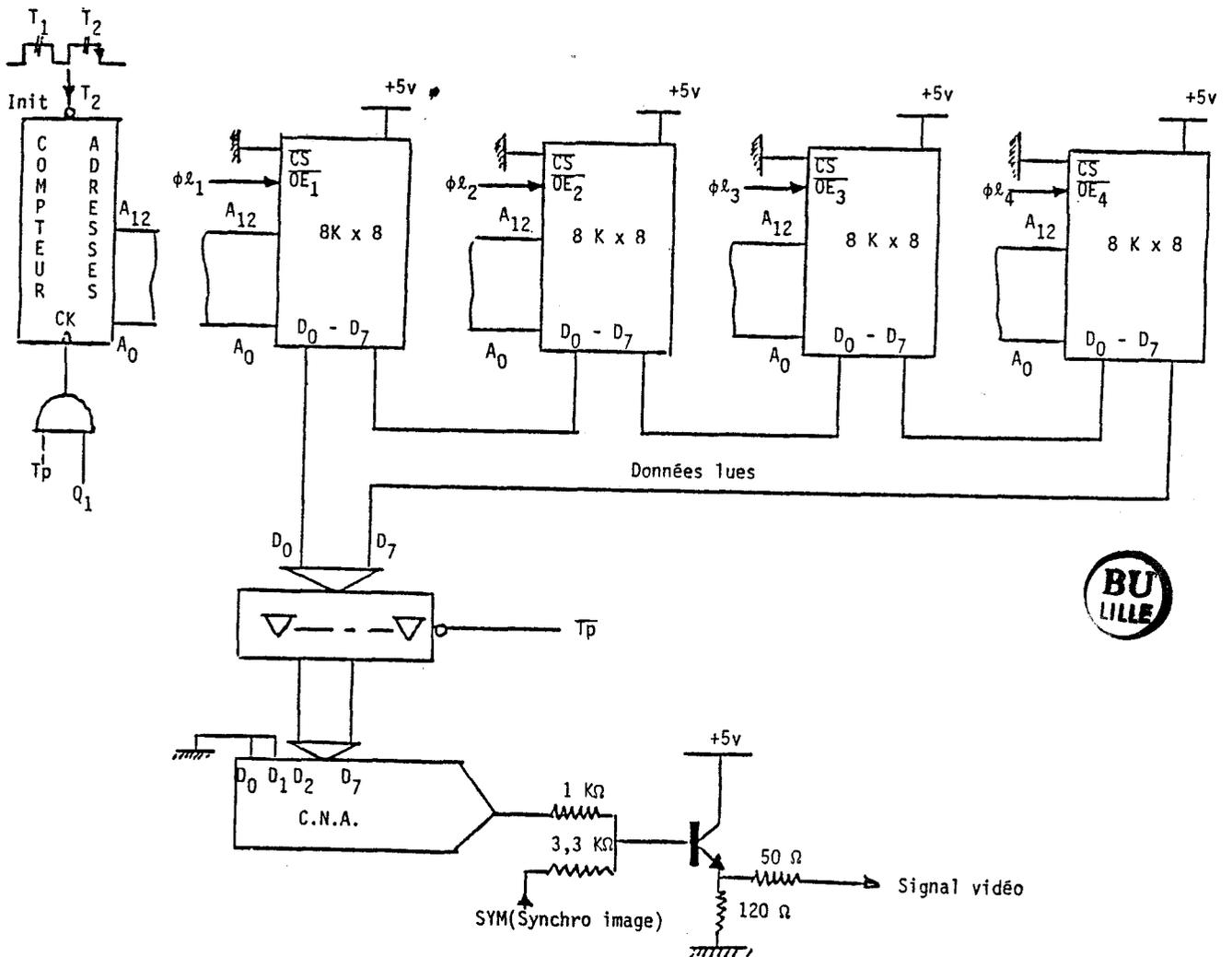


Figure IV-12 : Principe de visualisation de l'image mémorisée.

Nous allons créer quatre signaux de lecture :  $\phi_{l_1}, \phi_{l_2}, \phi_{l_3}, \phi_{l_4}$  qui vont être appliqués tour à tour aux entrées lecture-écriture des quatre plans mémoire comme le montre la figure IV-13.

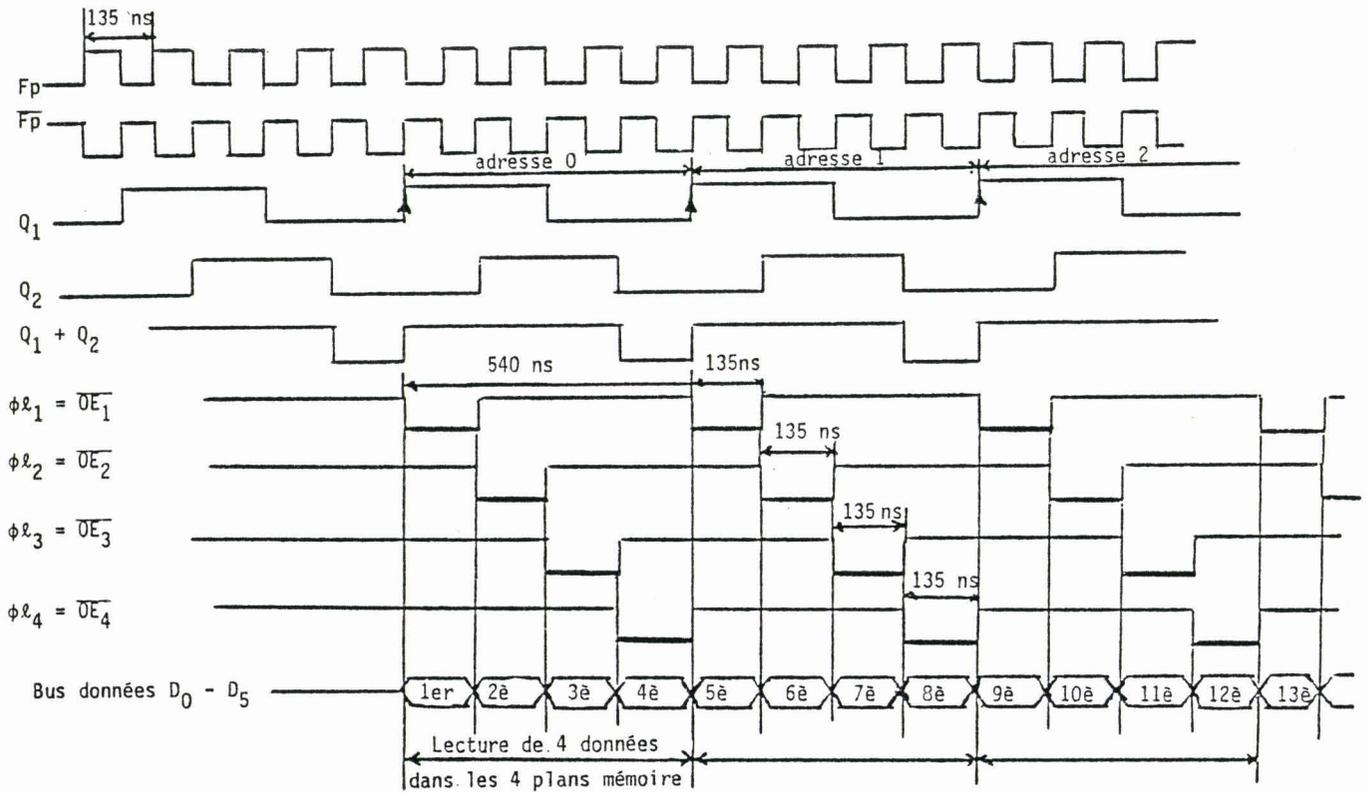


Figure IV-13 : Multiplexage des données lues tour à tour dans les quatre plans mémoire.

Les données des quatre plans mémoire sont mises en série vers le convertisseur numérique-analogique, ainsi l'image visualisée est obtenue.

Nous venons d'étudier les différentes fonctions de la mémoire.



La figure IV-14 représente l'architecture réalisée permettant la mémorisation, la visualisation et l'accès microordinateur.

La mémoire image est autonome, elle peut être couplée à n'importe quel ordinateur.

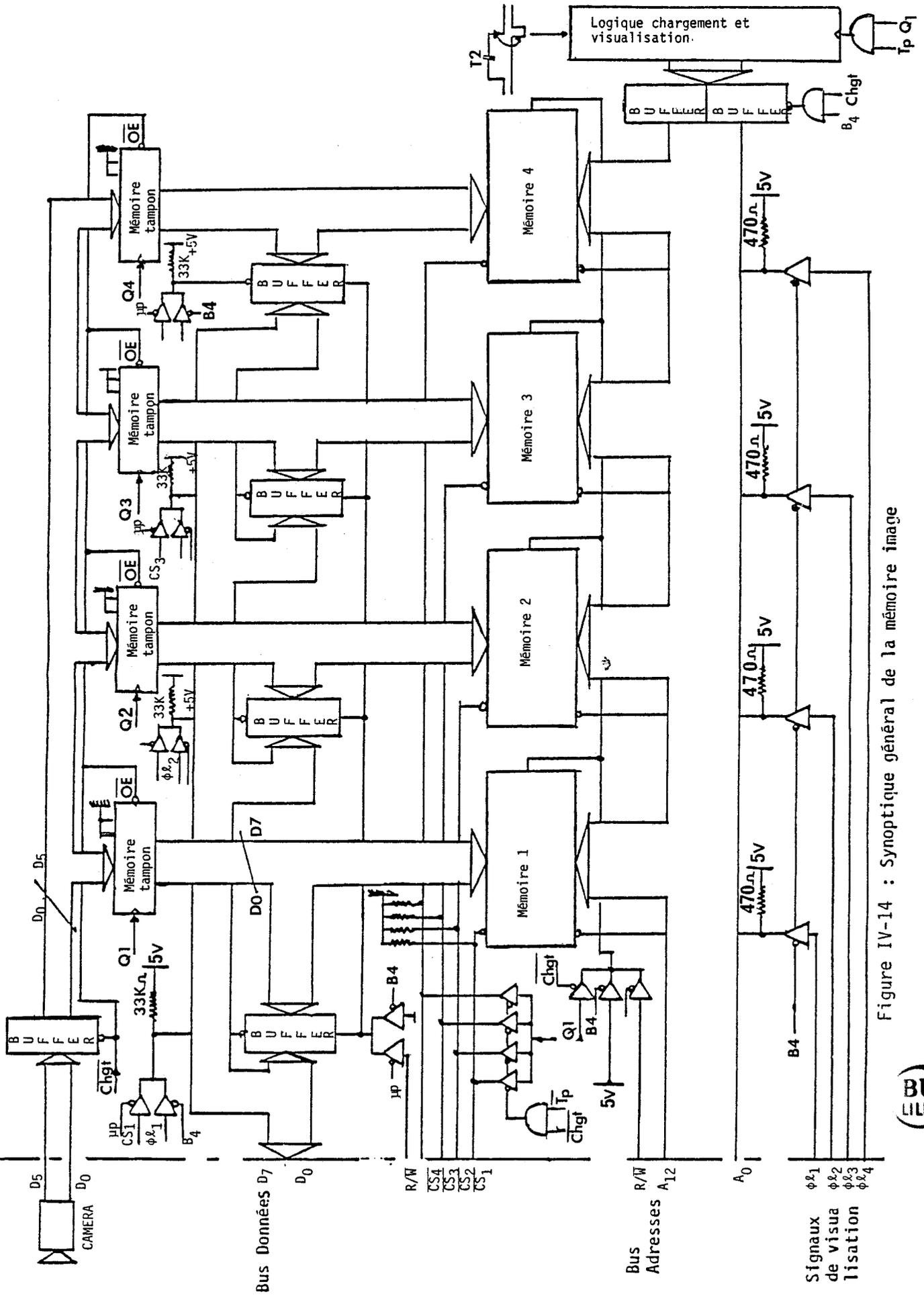


Figure IV-14 : Synoptique général de la mémoire image



#### IV - CONCLUSION

Dans ce chapitre, nous avons décrit les particularités de l'architecture du nouveau transcripateur Braille/noir réalisé. Ce système est muni aussi d'un noyau logiciel possédant toutes les fonctions nécessaires pour effectuer le traitement du relief Braille. Il reste à terminer l'étude de la reconnaissance des caractères Braille pour mettre en essai le nouveau transcripateur Braille/noir.



.../...

Enfin, dans un dernier temps, nous avons défini une méthode destinée à détecter les points Braille altérés. Les masques opérateurs de ces filtres sont calculés à l'aide de régression linéaire. Nous avons modélisé l'estimation de sa robustesse et appliqué systématiquement cette méthode de filtrage aux reliefs Braille altérés.

Les résultats obtenus montrent qu'à l'aide de cette méthode, l'estimation de la restitution des 95 % de l'information Braille est parfaitement vérifiée.

Dans ces conditions, il est probable qu'après transcription en texte noir, l'analyse du contexte permette de restituer les 100 % du document. Nous nous sommes bornés à cette démarche sans aborder celle de l'analyse du contexte.

Il reste à terminer l'étude de la reconnaissance des caractères Braille pour tester le nouveau transcripteur Braille/noir. Les techniques de reconnaissance et l'interactivité du système homme-machine vont permettre la restitution du document dans son intégralité. Ainsi, sera né un nouveau système de transcription Braille/noir des documents embossés manuellement.



.../...

- [15] NAHI , ASSEFI  
"Bayesien recursive image enhancement"  
IEEE Trans Comp. Vol. C 21 pp, 734-738, July 1972
- [16] HABIBI  
"Two dimensional bayesien estimate of images"  
Proc. IEEE, vol. 60, n° 7, pp. 878-883, July 1972
- [17] JAIN , ANGEL  
"Image restoration, modelling and reduction of dimensionality"  
IEEE trans on Comp, May 74  
Vol. 6 23, n° 5, pp. 470-476
- [18] S. TOHME  
"Squelettisation d'images binaires"  
Rapport interne ENST 1977
- [19] PAVEL  
"Squelettes de catégories d'images"  
Note aux comptes rendus de l'Académie des Sciences de Paris, 1976  
Série A, Tome 282, pages 751-754
- [20] ROSENFELD  
"Connectivity in digital pictures"  
JACM Vol. 17, n° 1, Janv.70 (146-160)
- [21] LEVINE  
"Feature extraction : a Survey"  
Proc. of IEEE, Vol. 57, n° 8, Août. 69
- [22] MONTANARI  
A note on minimal lenght Polygonal Approx.  
Comm. ACM 13, janv.70
- [23] FREEMAN  
"Computer processing of line drawing images"  
Comp. Survey Vol. 6, n° 1, Mars 74
- [24] A. KACED  
"Reconnaissance des formes et intelligence artificielle."  
"Le point sur les K formes"  
4ème congrès, janvier 1984.
- [25] B. BRILLAULT, H. MERCIER  
"Eléments de la théorie Morphologie mathématique sur des images binaires"
- [26] J. SERRA, S.R. STERNBERG  
Summer school in mathematical morphology
- [27] R. DERICKE  
"Mise en oeuvre de filtres à deux indices"  
Thèse de 3è cycle, Paris IX, 6 juin 1982
- [28] J.F. ABRAMATIC  
"Approximation de filtres biindiciels et traitement numérique des images"  
Thèse d'état, Paris VI, 18 décembre 1980

- [29] J. CAMILLERAPP, R. CUZON  
"Utilisation du contexte pour la détection d'objets faiblement contrastés"  
Congrès AFCET-IRIA Toulouse septembre 79, T2, pp. 217-224
- [30] K.S. FU  
"Syntactic (Linguistic) Pattern recognition"  
Digital Pattern recognition Springer Verlag 76, pp. 95-134
- [31] A. KACED, R. MOHR, J.J. SHENELL, J.P. HATON  
"Application de méthodes de traitements numériques d'images à l'examen  
de chaussées routières"  
Congrès AFCET-IRIA Toulouse, Septembre 79, T2, pp. 174-184
- [32] R.A. KIRSCH  
"Resynthesis of biological images from tree-Structure decomposition data"  
IFIP Working conference on graphic languages Vancouver, May 1972, North-  
Holland, pp.1-19
- [33] A. SCHER, R.F. DIAS VELASCO, A. ROSENFELD  
"Some new image smoothings techniques"  
IEEE, Trans. On Systems, Man and Cybernetics, Vol. SMC-10, n° 3,  
March 80, (153-159)
- [34] W.K. PRATT  
"Digital image processing"  
A Wiley Interscience Publication, 1978
- [35] V. CANTONI, I.DE LOTTO, M. FERRETTI  
"A template-matching operator for edge points detection in digital pictures"  
Signal processing 4, n° 5-6, Oct. 82, (349-360)
- [36] BASSEVILLE  
1."Edge detection using sequential methods for change in level"  
2."Sequential detection of change in mean"  
IEEE Trans. Vol. A SSP 29 Nr. 1, Février 1981
- [37] M. LABARRERE, J.P. KRIEF, B. GIMONET  
"Le filtrage et ses applications"
- [38] N.R. DRAPER and H. SMITH  
"Applied Regression Analysis. Wiley, New-York, 1966
- [39] L. REWO  
"Enhancement and detection of convexe objects using regression models",  
August,10, 1982



## RESUME

Ce travail est consacré à l'étude des méthodes numériques de détection par acquisition du relief Braille à l'aide d'une caméra et de traitement d'images dans une mémoire de trame dans le but de réaliser un nouveau système de transcription du texte Braille en texte noir.

Tout d'abord, nous mettons en évidence différents problèmes posés par la reconnaissance du relief en présentant une étude statistique permettant une classification des caractères Braille en fonction de la qualité d'embossage du relief compte tenu des conditions de mesure.

Dans une seconde partie, nous rappelons différentes méthodes numériques qui permettent de fournir une solution théorique au problème de détection du relief Braille. Nous mettons en oeuvre une première méthode de filtrage utilisant des compas directionnels (Prewitt, Kirsh ...) et des techniques de morphologie mathématique.

Mais pour s'affranchir des fluctuations des niveaux de gris de la luminosité dues aux propriétés photométriques des documents et pouvoir établir un seuil optimal de décision pour reconnaître l'information Braille, nous définissons dans une troisième partie un modèle adapté à la forme du signal Braille à l'aide de régression linéaire. Il permet de restituer 95 % des documents dans le cas d'embossage manuel du relief et 99 % de l'information dans le cas d'embossage mécanique du relief. Pour s'assurer de l'efficacité d'extraction du relief des masques fixes à régression linéaire, nous avons testé et estimé à l'aide du modèle leur robustesse face au bruit.

Enfin, dans une quatrième partie, nous décrivons les particularités de l'architecture du nouveau système de transcription du texte Braille en texte noir.

Mots clés : Relief Braille, filtrage bidimensionnel, Analyse d'images, Reconnaissance des formes.

