



# Traitement, Synthèse, Technologie et Applications

BIARRITZ — Mai 1984 —

PRETRAITEMENT OPTIQUE POUR UNE METHODE D'ANALYSE DE LA TEXTURE

OPTICAL PREPROCESSING FOR A TEXTURE ANALYSIS METHOD

L. DETTWILLER, R. MERCIER, P. CHAVEL

Institut d'Optique, Université de Paris-Sud, B.P. 43 - 91406 ORSAY Cedex, France  
(laboratoire associé au C.N.R.S.)

## RESUME

Dans la méthode d'analyse et de classification des textures appelée "texture energy measure", proposée par K.I. Laws [1], de bonnes performances sont obtenues après un traitement en trois étapes : l'image originale  $A$  est d'abord convoluée par un jeu de  $n$  masques ( $n = 5$  à  $10$ ) locaux de taille de l'ordre de  $3 \times 3$  pixels (maximum  $7 \times 7$ ). Les images résultantes  $B_i$ ,  $i = 1$  à  $n$ , sont ensuite élevées au carré puis lissées. On obtient ainsi  $n$  images traitées  $D_i$  dont les valeurs en chaque pixel constituent les composantes d'un vecteur  $n$ -dimensionnel caractérisant la texture. La classification supervisée ou non de ce vecteur par les méthodes usuelles fournit l'information de texture. Dans cette méthode, l'essentiel du temps de calcul est consacré aux opérations locales de bas niveau : convolution par les masques, élévation au carré, lissage. Ces opérations peuvent être effectuées par voie optique de façon pratiquement instantanée en parallèle sur toute l'image : compte tenu de la structure des masques, les convolutions peuvent être réalisées à l'aide d'un jeu de cristaux biréfringents dédoubleurs placés entre polariseurs et analyseurs convenablement orientés ; l'élévation au carré se produit automatiquement par détection quadratique, et l'intégration est effectuée lors de la digitalisation par un récepteur étendu. De cette façon, le nombre de points à digitaliser est diminué et l'ordinateur n'est plus utilisé que pour l'étape de plus haut niveau du traitement, la décision de classification.

Le principe de la méthode est décrit et des résultats préliminaires sont présentés.

## SUMMARY

K.I. Laws [1] proposed a texture analysis and classification method called "texture energy measure". Good results are obtained with this method after three processing steps : first, the original textured scene  $A$  is convolved by a set of  $n$  local convolution masks ( $n$  of the order of 5 to 10) ; these masks are of typical size  $3 \times 3$  pixels (maximum  $7 \times 7$ ). The resulting images  $B_i$ ,  $i = 1$  to  $n$ , are then squared and low-pass-filtered, yielding  $n$  processed images  $D_i$ . Each pixel of the  $D_i$ 's can be considered as a  $n$ -component vector characterizing texture. Texture information is then obtained by supervised or unsupervised classification. In that method, most of the computation time is devoted to low-level, local calculations : convolution by the masks, squaring, low-pass filtering. These operations can be performed optically virtually instantaneously, in real time on the whole image. The structure of the convolution masks allows their implementation by suitable combinations of birefringent crystals placed between polarizers and analysers. The squaring results automatically from quadratic detection, and the low-pass filtering effect can be obtained by digitizing the preprocessed images with an extended resolution area. The number of points to be digitized can then be reduced and the computer is used only for the high level step of the method, classification.

We describe the principles of the whole processing and present preliminary results.



## INTRODUCTION

Une grande partie de l'information contenue dans une image s'y trouve sous forme de textures qui permettent à l'oeil de reconnaître un objet indépendamment de sa forme. Une texture monochrome peut être décrite par des propriétés statistiques, comme les probabilités de co-occurrence de plusieurs niveaux de gris.

Or, pour une image de  $N \times N$  pixels à  $G$  niveaux de gris, il y a  $N^2 G^2$  probabilités de deux niveaux de gris seulement, soit  $4,3.10^9$  environ pour  $N = G = 256$ . Ce nombre énorme de données, engendrant des temps de calcul très longs, est le principal obstacle au traitement de la texture par voie purement numérique.

Même en se restreignant aux calculs associés au spectre de puissance des textures, transformation de Fourier, corrélation, convolution, l'analyse des textures nécessite des calculs numériques très lourds. Or, ces opérations s'effectuent aisément par voie optique. Nous présentons donc une méthode hybride d'analyse de la texture, où tout le traitement de bas niveau mais à grand nombre de points est réalisé optiquement, puis le traitement de plus haut niveau comme la segmentation ou la classification est réalisé numériquement.

Cette méthode reprend l'idée exploitée par Laws [1] de caractériser la texture à partir de moyennes locales effectuées sur une série d'images prétraitées. Après en avoir rappelé succinctement les points essentiels, nous décrirons les principes utilisés dans la mise en oeuvre hybride optique/numérique et présentons un résultat préliminaire.

## 1. PRESENTATION DE LA METHODE D'ANALYSE DE LA TEXTURE.

Mise en oeuvre par son auteur sous forme purement numérique [1], cette méthode procède en trois étapes.

### 1.1. Première étape : filtrage

L'image originale est tout d'abord convoluée par une série de réponses percussionnelles finies représentées par des matrices  $m \times n$  ; dans la référence [1],  $m$  et  $n$  prennent les valeurs 1, 3, 5, 7, le couple (1,1) étant exclu. Nous nous limiterons ici au cas  $3 \times 3$ , avec lequel Laws a obtenu de très bons résultats de classification. Les matrices  $3 \times 3$  en question sont au nombre de neuf et s'obtiennent en prenant tous les produits "tensoriels" des trois matrices lignes  $1 \times 3$  :

$$\begin{aligned} L &= \begin{pmatrix} 1 & 2 & 1 \\ -1 & 0 & 1 \\ -1 & 2 & -1 \end{pmatrix} \\ E &= \begin{pmatrix} -1 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 2 & -1 \end{pmatrix} \\ S &= \begin{pmatrix} 1 & 2 & 1 \\ -1 & 0 & 1 \\ -1 & 2 & -1 \end{pmatrix} \end{aligned}$$

avec leurs trois transposées. On obtient donc les neuf matrices suivantes :

$$L * {}^T L = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{pmatrix} \quad L * {}^T E = \begin{pmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{pmatrix} \quad L * {}^T S = \begin{pmatrix} -1 & -2 & -1 \\ 2 & 4 & 2 \\ -1 & -2 & -1 \end{pmatrix}$$

$$E * {}^T L = \begin{pmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix} \quad E * {}^T E = \begin{pmatrix} +1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{pmatrix} \quad E * {}^T S = \begin{pmatrix} 1 & 0 & 1 \\ -2 & 0 & 2 \\ 1 & 0 & -1 \end{pmatrix}$$

$$S * {}^T L = \begin{pmatrix} -1 & 2 & -1 \\ -2 & 4 & -2 \\ -1 & 2 & -1 \end{pmatrix} \quad S * {}^T E = \begin{pmatrix} 1 & -2 & 1 \\ 0 & 0 & 0 \\ -1 & 2 & -1 \end{pmatrix} \quad S * {}^T S = \begin{pmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & 2 & 1 \end{pmatrix}$$

Cet ensemble de 9 matrices forme une base de l'espace des matrices  $3 \times 3$ . On constate que la somme des coefficients d'une quelconque de ces matrices est nulle, sauf pour  $L * {}^T L$ .

Cette première étape donne donc, à partir de l'image originale, 9 images filtrées.

### 1.2. Deuxième étape : statistiques du premier ordre

Ensuite, sur une fenêtre de  $15 \times 15$  pixels par exemple, sont calculées différentes statistiques du premier ordre : la variance, l'écart type, la moyenne des valeurs absolues, la moyenne des niveaux de gris positifs ou la moyenne des niveaux de gris négatifs. La pratique montre que, quand elles opèrent sur l'ensemble des 9 images filtrées, les efficacités de ces statistiques appliquées à la classification des textures sont très voisines.

Remarquons que toute image convoluée par une matrice dont la somme des coefficients est nulle a une moyenne nulle ; c'est le cas de 8 images filtrées sur 9. Pour ces 8 images filtrées, la variance est donc égale à la moyenne des carrés des niveaux de gris, c'est-à-dire à une "énergie" moyenne de l'image.

### 1.3. Analyse agrégative

Pour une texture donnée, l'une des statistiques (comme la variance) calculée sur les 9 images filtrées donne, dans l'espace à 9 dimensions, un point caractéristique de la texture considérée. En fait, comme les statistiques sont calculées sur une fenêtre de dimensions finies ( $15 \times 15$  pixels par exemple), elles varient avec la position de la fenêtre sur le champ de texture étudié. On n'obtient donc pas un point, mais un nuage de points d'autant plus serré que la fenêtre est grande.

En appliquant l'analyse agrégative à un champ de texture composite, on obtient plusieurs nuages de points, un pour chaque texture, et s'ils sont bien séparés on peut classifier les textures. La résolution de la segmentation est d'autant moins bonne que la fenêtre où sont calculées les statistiques est plus grande, alors que la précision de la segmentation augmente avec la finesse des nuages, donc la dimension de la fenêtre ; d'où la nécessité de trouver compromis précision résolution.

Cette méthode ne peut être mise en oeuvre que pour des images à faible nombre de pixels ( $512 \times 512$ ) à cause du temps de calcul (de 30 à 50 min) sur PDP 10 - KL 10 pris à 80 % pour les deux premières étapes. D'où l'intérêt de la méthode que nous présentons, qui réalise optiquement les deux premières étapes, donc pendant le temps de propagation de la lumière, quel que soit le nombre de pixels de l'image traitée.

## 2. PRETRAITEMENT OPTIQUE.

### 2.1. Décomposition des matrices de filtrage en matrices élémentaires

Il est évident que les matrices ligne  $L$ ,  $E$  et  $S$  s'obtiennent par convolution de matrices élémentaires  $(1 \ 1)$  et  $(-1 \ 1)$  :

$$\begin{aligned} (1 \ 1) * (1 \ 1) &= (1 \ 2 \ 1) = L \\ (-1 \ 1) * (1 \ 1) &= (-1 \ 0 \ 1) = E \\ (-1 \ 1) * (-1 \ 1) &= (1 \ -2 \ 1) = -S \end{aligned}$$

donc toutes les matrices  $3 \times 3$  utilisées par Laws pour obtenir les images filtrées dans la première étape, peuvent s'exprimer comme produit de convolution des matrices élémentaires  $(1 \ 1)$ ,  $(-1 \ 1)$  et leurs transposées. Trois produits de convolution entre quatre matrices élémentaires sont alors nécessaires, ce qui donne  $2^4 = 16$  combinaisons possibles, mais à cause de la commutativité de la convolution  $(-1 \ 1) * (1 \ 1) = (1 \ 1) * (-1 \ 1)$ , on n'obtient que 9 matrices différentes au lieu de 16.

PRETRAITEMENT OPTIQUE POUR UNE METHODE D'ANALYSE DE LA TEXTURE  
OPTICAL PREPROCESSING FOR A TEXTURE ANALYSIS METHOD

L. DETTWILLER, R. MERCIER, P. CHAVEL

2.2. Réalisation optique du filtrage

En traitement analogique, l'analogue de la convolution d'une image digitalisée par (1 1) ou (-1 1) est la convolution par les réponses percussionnelles élémentaires

$$R_1(\vec{r}) = \delta_D(x) + \delta_D(x - a)$$

ou

$$R_{-1}(\vec{r}) = \delta_D(x) + \delta_D(x - a)$$

où  $\vec{r} = \begin{pmatrix} x \\ y \end{pmatrix}$  est le vecteur repérant un point courant de l'image, et  $\delta_D$  représente la distribution de Dirac ; la distance  $a$  correspond au pas d'échantillonnage de l'image digitale.

Or, ces réponses percussionnelles peuvent être obtenues à l'aide de cristaux biréfringents en éclairage cohérent. En effet, une lame à faces parallèles d'un matériau uniaxe donne d'un même objet deux images de Gauss translattées. Si l'axe optique du milieu est orienté à environ 45° des faces, les deux images sont décalées au maximum tout en se trouvant pratiquement dans le même plan (voir figure 1).

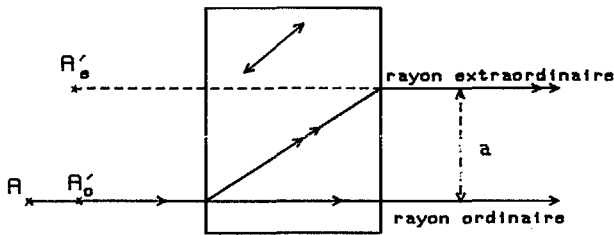


Figure 1. Dédoublment d'image par une lame biréfringente à faces parallèles.

Pour que les amplitudes de ces images soient égales, et se trouvent en phase ou en opposition de phase, la solution la plus simple consiste à mettre avant la lame un polariseur P orienté à 45° de ses lignes neutres pour l'incidence normale, à choisir une lame introduisant un déphasage de  $\pi$  entre les deux rayons, et enfin à placer après la lame un analyseur A perpendiculaire ou parallèle au polariseur (voir figure 2). De plus, comme le déphasage entre les rayons

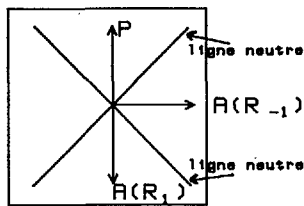


Figure 2. Orientations du polariseur P, des lignes neutres du cristal et de l'analyseur A pour obtenir les réponses  $R_1$  et  $R_{-1}$ .

ordinaire et extraordinaire dépend de l'incidence, l'objet placé avant la lame doit être éclairé par une onde plane.

La convolution par  $R_1$  ou  $R_{-1}$  peut aussi être vue comme le filtrage des fréquences spatiales de l'objet (qui diffractent chacune une onde plane) par la figure d'interférences à l'infini de la lame cristalline, entre polariseur et analyseur. Cette figure, qui joue le rôle d'une fonction de transfert de modulation, a l'aspect de franges d'Young. Ces considérations s'inscrivent dans le cadre général de la théorie du filtrage des fréquences spatiales par interférence [2].

Pour réaliser la convolution de l'objet par les matrices 3 x 3 de Laws, il faut mettre en cascade quatre lames identiques entre polariseurs, deux

donnant un décalage horizontal et deux un décalage vertical. Le polariseur qui précède la première lame étant fixé à 45° des lignes neutres de toutes les lames, il suffit de mettre chacun des autres polariseurs parallèle ou perpendiculaire au premier pour obtenir les 9 images filtrées par les 9 matrices 3 x 3.

2.3. Réalisation optique des statistiques du premier ordre

Réalisées en éclairage cohérent, les convolutions ci-dessus portent donc sur l'amplitude objet. Or, le récepteur d'images est sensible à l'éclairement, donc au carré de l'amplitude image, intégré sur l'aire de sa cellule de résolution. Cette opération d'enregistrement revient à calculer la moyenne du carré de l'amplitude de l'image filtrée sur des fenêtres (les cellules de résolution du récepteur) ce qui, pour les images obtenues par convolution par une matrice de somme nulle, coïncide avec la variance qu'utilisait Laws.

2.4. Montage préliminaire et problèmes expérimentaux

Le montage utilisé se trouve représenté sur la figure 3. Avant de recevoir les images filtrées avec une

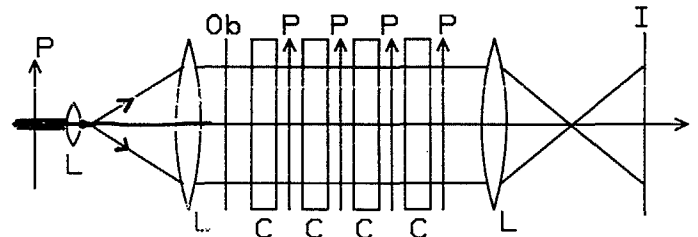


Figure 3. Schéma du montage ; P = polariseur, L = objectif, C = cristal, Ob = objet, I = image. Eclairage par laser He-Ne.

caméra de télévision permettant d'acquérir directement une image numérisée pour en faire l'analyse agrégative, nous mettons pour une expérience préliminaire une chambre photographique et nous numérisons et analysons les images filtrées après développement.

Les cristaux utilisés donnent un décalage de 39  $\mu\text{m}$  chacun. Ces 39  $\mu\text{m}$  définissent la taille du "pixel". Dans l'expérience de principe décrite ici, nous réalisons un objet test que montre la figure 4, formé non pas de

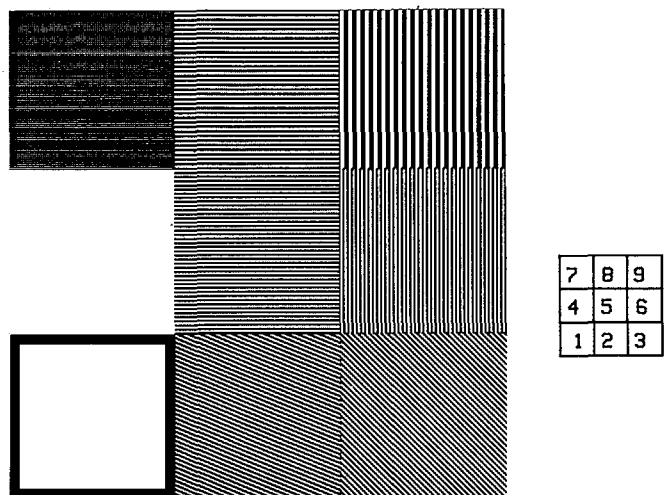


Figure 4. Objet test (négatif agrandi) et numérotation des zones.



PRETRAITEMENT OPTIQUE POUR UNE METHODE D'ANALYSE DE LA TEXTURE  
OPTICAL PREPROCESSING FOR A TEXTURE ANALYSIS METHOD

L. DETTWILLER, R. MERCIER, P. CHAVEL

textures naturelles mais de mires de fréquences et d'orientations différentes, mais dont les ordres de grandeur des pas sont adaptées au décalage des cristaux. La taille de l'objet est 10,2 x 10,2 mm.

Le problème expérimental majeur est celui de la lumière parasite. Elle provient de la diffusion par la lentille collimatrice, qui crée une sorte de fond continu, et des réflexions sur les faces des cristaux qui rajoutent des images filtrées supplémentaires. Le carré noir sur l'objet positif sert à estimer l'importance du bruit par rapport au signal des plages voisines.

### 3. CLASSIFICATION SUR LES IMAGES PRETRAITEES

Une fois développée, l'image prétraitée (résultant des deux premières étapes) est analysée et numérisée avec un microdensitomètre. La fenêtre d'analyse, ramenée sur l'objet, est un carré de 0,16 mm de côté : autrement dit, la seconde étape de la méthode, le calcul de statistiques locales sur une fenêtre, est réalisée automatiquement par intégration du signal sur une surface de 4 x 4 pixels lors de la numérisation. Les pixels de l'image numérisée sont donc 16 fois moins nombreux qu'avant digitalisation.

Les niveaux de gris enregistrés sur chaque pixel de l'image digitalisée pour les 9 images prétraitées numérisées forment les coordonnées d'un ensemble de points d'un espace à 9 dimensions. Chaque pixel d'une des mires de l'objet est représenté par un point de l'espace à 9 dimensions, qui s'inscrit dans un nuage caractéristique de la partie de l'objet traité. Sur la figure 5 a, on voit les projections des nuages sur un plan correspondant à deux des neuf coordonnées.

Les nuages qui sont confondus sur la figure 5 a peuvent être mieux séparés sur d'autres plans de coordonnées. La figure 5 b montre la séparation des quatre nuages qui étaient confondus dans la figure 5 a. La figure 5 c montre les deux nuages confondus en 5 a (les moins bien séparés), projetés sur leur meilleur plan de coordonnées. Rappelons que ces résultats sont obtenus avec une intégration sur une fenêtre de 4 x 4 pixels, alors que Laws utilisait 15 x 15 pixels. La figure 5 d montre l'amélioration de la séparation, par intégration sur 8 x 8 pixels.

On est alors en mesure de réaliser une classification sur l'image. Pour cela, on calcule au préalable les barycentres des nuages, et on cherche de quel barycentre les points représentatifs de chaque pixel se rapprochent le plus.

La figure 6 montre le résultat de la classification. Les zones de l'objet sont très bien reconnues, seules apparaissent parfois des fluctuations aux bords des zones. Entre la zone 5 et la zone 8, on observe cependant une diffusion d'une zone vers l'autre, due à la moins bonne séparation de leurs nuages : en effet, les mires de ces deux zones ont des traits parallèles et leurs pas ne diffèrent que de 12 % ; cette différence qui n'est pas évidente à l'oeil suffit à segmenter de façon approximative ces zones, par notre méthode.

### CONCLUSION

Cette méthode s'inscrit dans l'axe de l'évolution actuelle du traitement optique d'images, où les possibilités de l'optique sont utilisées en complément du traitement numérique : on les associe ici dans un traitement hybride où l'optique est utilisée pour effectuer rapidement et en parallèle les nombreuses convolutions nécessaires dans un prétraitement. Notre méthode, avec un convertisseur d'images à l'entrée, des cristaux électro-optiques commandés par ordinateur et une caméra reliée au système de traitement numéri-

que, devrait permettre l'analyse de textures en temps réel.

### REFERENCES

- [1] K.I. LAWS, Textured image segmentation, 1980, thèse, University of Southern California.
- [2] L. DETTWILLER and P. CHAVEL, "Optical spatial frequency filtering by interferences", *J. Opt. Soc. Am. A* (janvier 1984).

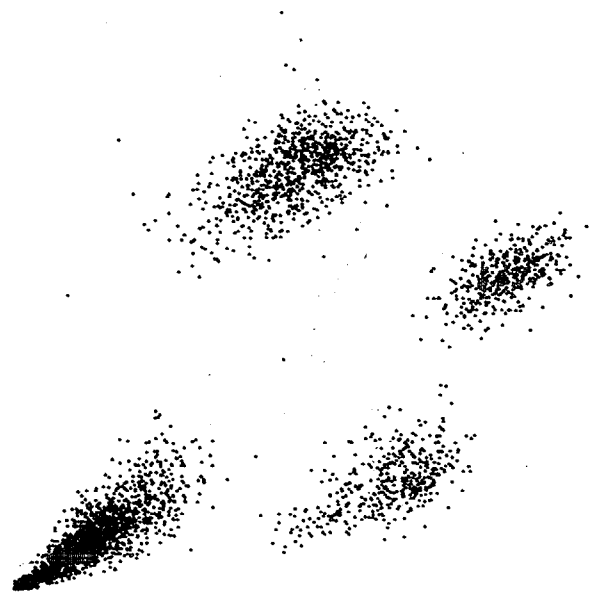


Figure 5 a - Huit nuages de textures projetés dans un plan de coordonnées : quatre des huit nuages se chevauchent en bas à gauche, deux en haut ; les deux autres sont bien séparés.

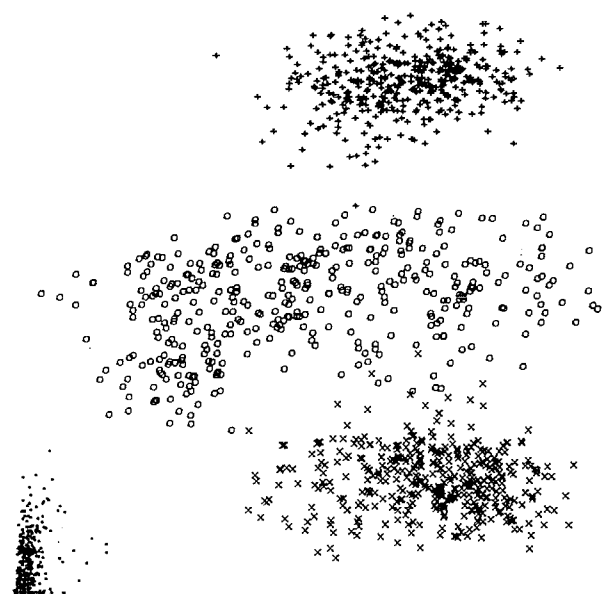


Figure 5 b - Projection dans un autre plan de coordonnées des quatre nuages confondus en bas à gauche de 5 a.





PRÉTRAITEMENT OPTIQUE POUR UNE METHODE D'ANALYSE DE LA TEXTURE  
 OPTICAL PREPROCESSING FOR A TEXTURE ANALYSIS METHOD

L. DETTWILLER, R. MERCIER, P. CHAVEL

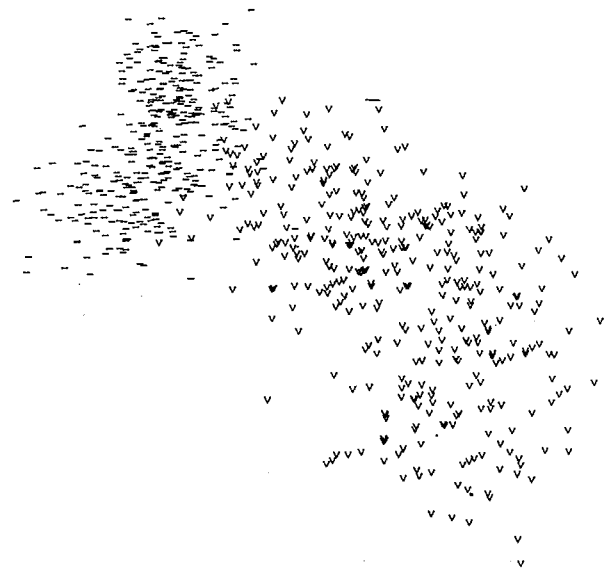
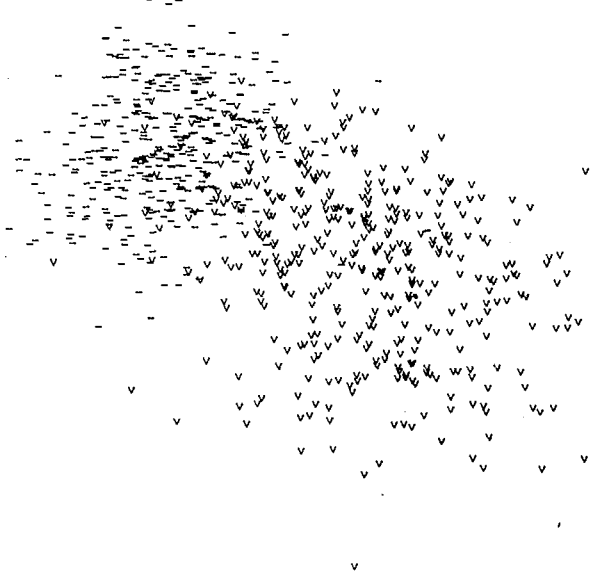


Figure 5 c - Projection dans un autre plan de coordonnées des deux nuages confondus en haut de 5 a.

Figure 5 d - Les deux nuages de 5 c après intégration des images prétraitées sur 8 x 8 pixels (au lieu de 4 x 4 pour les figures 5 a à c).

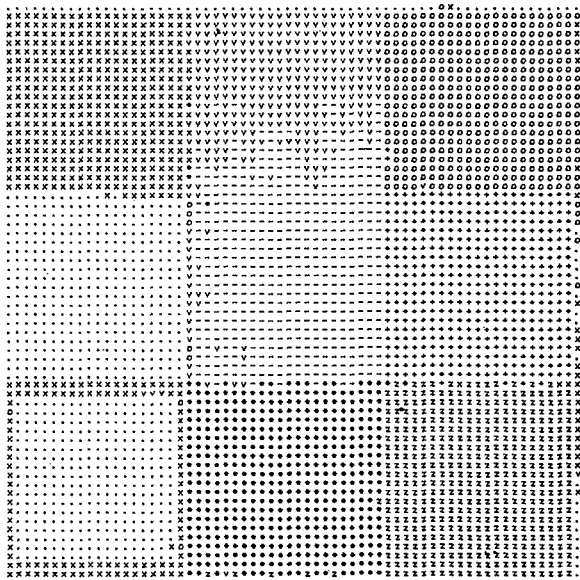


Figure 6 - Résultats de la classification supervisée : chaque pixel a été attribué à une classe parmi huit, chaque classe est représentée par un symbole différent, qui coïncide avec celui utilisé en 5 b, 5 c et 5 d.