

## Segmentation optimale d'images de texture par transformations adaptées multiples

---

### *Optimal Texture Image Segmentation Using Multiple Matched Transforms*

par P. PROVENT, J. LEMOINE et E. PETIT

Laboratoire d'Etudes et Recherches en Instrumentation Signaux et Systèmes  
Université Paris-XII Val de Marne  
61, avenue du Général de Gaulle, 94010 Créteil Cedex  
Tél. : 45 17 14 93

#### Résumé

Dans cet article nous présentons une méthode statistique d'analyse de textures fondée sur la Transformée de Karhunen-Loeve (T.K.L.). La TKL permet une description compacte et optimale de la texture. Des bancs de filtres adaptés à chaque texture de l'image (filtres RIF) sont déterminés durant une phase d'apprentissage.

La segmentation automatique d'une image de texture est réalisée par classement automatique des pixels sur la base des valeurs prises par leur vecteur d'attributs de texture. Le banc de filtres qui assure une répartition optimale de "l'énergie de texture" dans le voisinage du pixel considéré détermine sa classe d'affectation. La règle de décision choisie est fondée sur le principe du maximum de vraisemblance (Maximum A Posteriori).

Le choix généralement critique des attributs de texture les plus pertinents se trouve résolu puisqu'il découle directement de la méthode proposée. Ainsi, la méthode permet de traiter des textures aussi bien déterministes qu'aléatoires.

**Mots clés :** Analyse de texture, classification, transformation de Karhunen-Loeve, filtrage adapté.

#### Abstract

*This paper deals with a new method for texture analysis based on the Karhunen Loeve Transform (KLT). KLT allows an optimal and compact description of textures. During a training step, filter sets matched to each analyzed texture are computed.*

*Automatic image segmentation is performed classifying each pixel according to its texture attribute vector. A pixel belongs to the class corresponding to the filter set which leads to an optimal share out of its "texture energy". The decision rule is based on the Maximum Likelihood principle.*

*The critical problem of choosing the most relevant texture attributes is automatically solved by the proposed method. Thus, different texture classes can be processed.*

**Key words :** Texture analysis, classification, Karhunen-Loeve transform, matched filtering.

## 1. Introduction

La recherche de mesures permettant de quantifier l'information structurelle de la texture (attributs de texture) est fondamentale pour la caractérisation de la surface des objets. Ces mesures doivent d'une part fournir une caractérisation fidèle d'une grande variété de textures (pertinence des attributs) et, d'autre part, permettre d'en différencier plusieurs aisément (fort pouvoir discriminant).

La notion de texture intervient dans de nombreux problèmes tels que la classification, la segmentation, l'analyse et la synthèse d'images. D'après Gagalowicz [1] une texture doit être considérée comme : "une structure spatiale constituée par l'organisation de primitives ayant chacune un aspect aléatoire. Une texture est alors une structure hiérarchique à deux niveaux".

Le choix d'une méthode d'analyse dépend du type de texture à traiter [2].

- Les méthodes structurelles d'analyse s'intéressent aux textures déterministes et correspondent à une vision macroscopique de celles-ci (haut niveau). Dans ce cas, on cherche à isoler des motifs élémentaires (texels [3]) et à y associer un ensemble de règles de placement [4]. Des textures très structurées de type mur de brique, grillage, trame de textile, ... sont ainsi analysées. On les appelle généralement macrotextures.

- Les méthodes statistiques sont dédiées aux textures dites aléatoires [5]. Il est difficile dans ce cas d'isoler des texels de manière stable. On parlera alors de vision microscopique de texture (bas niveau) et on supposera pour l'analyse que la texture est la réalisation d'un processus aléatoire bidimensionnel. Les caractéristiques texturales sont évaluées à partir des statistiques d'ordre N d'un

groupe de pixels pris au travers d'une fenêtre d'observation qui peut être réduite au pixel. Les textures naturelles font partie de cette catégorie : sable, forêts et cultures vues à très grande distance ... Elles sont appelées microtextures.

Le système visuel humain possède deux modes principaux d'interprétation : préattentif et focal. La perception des textures s'effectue essentiellement au niveau préattentif. D'après Julesz [6] le système visuel humain procède à une analyse que l'on peut qualifier de "mesure de densité d'unités élémentaires". De ce principe se dégage une méthodologie générale de l'analyse de texture opérant en deux étapes (figure 1) :

- Premièrement : des mesures locales sont extraites par observation au travers d'une fenêtre de voisinage (caractérisation d'unités élémentaires),
- Deuxièmement : des mesures globales sont déterminées par groupement des mesures précédentes. Elles constituent les attributs de texture et devront fournir une information quantitative compacte, au même titre que l'observation d'une texture se traduit par une perception visuelle homogène.

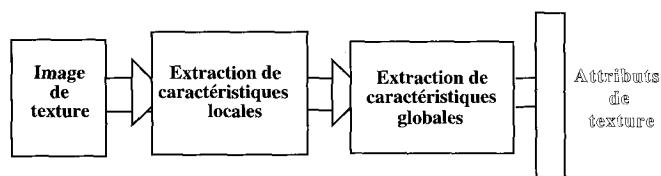


Figure 1 : Méthodologie de l'analyse de texture

## 2. Analyse de texture par filtrage linéaire

L'article que nous présentons portera spécifiquement sur la caractérisation des microtextures. Les microtextures se caractérisent essentiellement par leurs propriétés de voisinage. Ainsi, l'analyse revient à quantifier les relations structurelles qui lient les pixels localement. Le modèle le plus simple est fondée sur une mesure de la dépendance spatiale linéaire entre pixels vus au travers d'une fenêtre microscopique.

### 2.1. DESCRIPTION DE L'ANALYSE DE TEXTURE PAR FILTRAGE LINÉAIRE

L'utilisation de techniques de filtrage linéaire bidimensionnel pour l'analyse de texture a été introduite par Laws [7] et Wermser [8]. Laws a utilisé un banc de filtres à Réponse Impulsionnelle Finie (filtres RIF). Chacun de ces filtres permet d'extraire de l'image une composante structurelle particulière de la texture. Celle-ci est orientée suivant la direction du gradient privilégiée par le filtre. La figure 2 présente schématiquement la procédure d'extraction des attributs de texture par filtrage linéaire.

Laws définit un vecteur d'attributs de texture  $E = \{E_1, E_2, \dots, E_N\}$ , nommé "vecteur d'énergie de texture" calculé à partir des valeurs prises par les variables  $y_n$  décrivant les

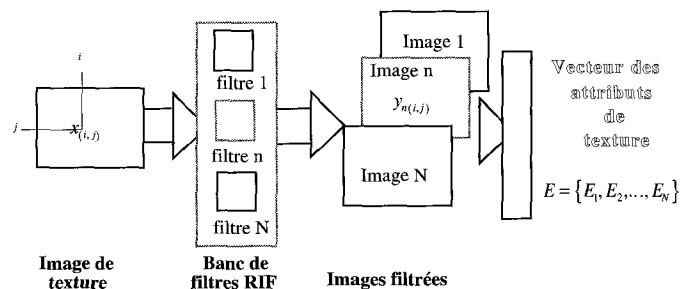


Figure 2 : Extraction d'attributs de texture par filtrage linéaire

images filtrées. Une composante  $E_n$  de  $E$  est donnée par :

$$E_N = \frac{1}{I \cdot J} \sum_{i=1}^I \sum_{j=1}^J (y_{n(i,j)} - \bar{y}_n)^2 \quad (1)$$

$n$  est le numéro de l'image filtrée,  $i$  et  $j$  sont les variables d'espace. La valeur moyenne d'une image filtrée est déterminée par :

$$\bar{y}_n = \frac{1}{I \cdot J} \sum_{i=1}^I \sum_{j=1}^J (y_{n(i,j)})$$

La contribution énergétique d'une image filtrée à l'énergie totale sera d'autant plus grande que les réponses aux filtres appliqués le seront. Les plus fortes réponses permettent de déterminer les directivités prépondérantes des textures analysées. Ainsi, deux textures a priori différentes auront une répartition différente de leur "énergie de texture" pour l'ensemble des filtres de Laws. Il est ainsi théoriquement possible de les différencier par une mesure de distance dans l'espace des attributs de textures.

Les inconvénients majeurs de cette technique résident premièrement, dans le choix empirique des filtres RIF, deuxièmement, par le fait qu'elle n'assure pas une conservation de l'énergie totale de l'image et troisièmement, parce que les images filtrées sont corrélées (redondance d'information).

### 2.2. GÉNÉRALISATION DE L'ANALYSE DE TEXTURE PAR FILTRAGE LINÉAIRE

A la différence de Laws qui utilise un banc de filtres introduit de manière empirique, M. Unser [9] propose une formulation plus générale de l'analyse de texture par filtrage linéaire qui utilise le formalisme vectoriel suivant :

Soit  $X$  un vecteur colonne de mesure,  $X = \{x_1, x_2, \dots, x_N\}^T$  où les  $x_n$  sont les pixels appartenant à une fenêtre d'observation de forme quelconque (figure 3).

Le vecteur de mesure s'exprime alors sur l'espace bidimensionnel de l'image par :

$$X_{(i,j)} = \{x_{(i-k,j-l)}, x_{(i-k+1,j-l)}, \dots, x_{(i+k,j+l)}\}^T$$

Ainsi, l'analyse de texture par filtrage linéaire s'exprime comme une transformation linéaire par la relation suivante :

$$Y = A \cdot X \quad (2)$$

$\mathbf{A} = \{A_1, A_2, \dots, A_N\}^T$  est la matrice de passage du vecteur de mesure  $X$  au vecteur transformé  $Y$ .  $A_n$  est un vecteur colonne de la matrice  $\mathbf{A}$ .

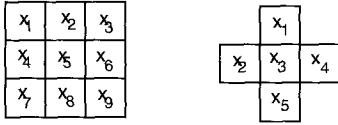


Figure 3 : Exemples de fenêtres d'observation

La première contrainte imposée aux différentes transformations possibles est la conservation énergétique totale de la texture. Compte tenu de la définition du vecteur  $X$ , l'énergie de la texture se retrouve dans les termes diagonaux de sa matrice de covariance spatiale  $\mathbf{C}_X$ .

Il doit en être de même avec la matrice de covariance  $\mathbf{C}_Y$  du vecteur transformé  $Y$ . Le critère de conservation énergétique s'exprime donc comme suit :

$$\text{Trace}(\mathbf{C}_X) = \text{Trace}(\mathbf{C}_Y) \quad (3)$$

Une deuxième contrainte impose l'orthogonalité des vecteurs de la base de transformation afin d'assurer une décorrélation des composantes  $y_n$  de  $Y$ .

Plusieurs transformées classiques comme la DST (Discret Sin Transform), la DCT (Discret Cosin Transform) ou celle d'Hadamard ont été testées par Ade [10] en comparant leur efficacité en contrôle de qualité de la fabrication de textiles. Ces transformations ont l'inconvénient d'être des bases fixées une fois pour toute et ne sont donc pas, de ce fait, adaptées à la nature propre des textures.

### 3. Analyse statistique de texture par transformation de Karhunen-Loeve

Parmi les différentes transformations répondant aux critères précédemment décrits nous avons opté pour la Transformation de Karhunen-Loeve.

#### 3.1. ANALYSE STATISTIQUE DE TEXTURE

Les propriétés statistiques des microtextures sont étudiées par des observations au travers de la fenêtre d'observation  $X_{(i,j)}$ . Les réalisations du vecteur de mesure  $X$  sont recueillies en déplaçant la fenêtre d'observation par glissement et avec recouvrement à l'intérieur d'une région de texture.

Si l'image est décrite par un modèle statistique où chaque variable  $x_n$  du vecteur de mesure  $X$  est considérée comme aléatoire, alors, une texture sera entièrement caractérisée par sa fonction de densité de probabilité conjointe d'ordre  $N$ ,  $p(X)$ , ou de manière équivalente par l'ensemble de ces moments.

On considérera par la suite que les caractéristiques texturales sont spatialement invariantes par translation (processus stationnaire)

et on admettra qu'elles sont statistiquement bien représentées. Cette dernière condition sera réalisée si le nombre d'observations est grand. La région d'analyse doit être suffisamment large pour caractériser entièrement la texture.

La région de taille minimale qui satisfait cette condition est appelée *résolution de la texture*.

Les grandeurs statistiques retenues dans notre approche sont les moments d'ordre deux. On parlera par la suite indifféremment de moments centrés d'ordre deux ou de matrice de covariance en considérant des textures à valeur moyenne nulle.

Soit  $\mathbf{X}$  la matrice d'observation du vecteur  $X$ , la condition précédente s'écrit :

$$\hat{E}(\mathbf{X}) = 0$$

On vérifie donc l'égalité ci-dessous :

$$\mathbf{C}_X = \hat{E}(\mathbf{X} \cdot \mathbf{X}^T) = \hat{E} \left( (\mathbf{X} - \hat{E}(\mathbf{X})) (\mathbf{X} - \hat{E}(\mathbf{X}))^T \right)$$

La matrice de covariance spatiale  $\mathbf{C}_X$  exprime les dépendances linéaires entre les variables  $x_n$  du vecteur de mesure, donc par transposition à l'image, celles entre les pixels  $x_{(i,j)}$  de la fenêtre d'observation.

#### 3.2. VERS UNE MÉTHODE OPTIMALE POUR L'ANALYSE DE TEXTURE

La Transformation de Karhunen-Loeve est une transformation linéaire unitaire bidimensionnelle définie par l'équation suivante :

$$Y = \Phi \cdot X$$

Cette transformation respecte l'égalité :

$$\Lambda \cdot \Phi = \Phi \cdot \mathbf{C}_X \quad (4)$$

avec,

$$\Lambda = \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \dots & \\ 0 & & & \lambda_N \end{bmatrix} \text{ matrice des valeurs propres de } \mathbf{C}_X$$

et,

$\Phi = \{\Phi_1, \Phi_2, \dots, \Phi_N\}^T$  matrice des vecteurs propres colonne de  $\mathbf{C}_X$

La TKL est la seule transformation qui assure une décorrélation parfaite des variables  $y_n$  quelque soit la fonction de densité de probabilité  $p(X)$ .

Soit :

$$\text{cov}(y_i, y_j) = 0 \quad \forall i \neq j \quad (5)$$

avec,

$$\Phi_i \cdot \Phi_j^T = 0 \quad \forall i \neq j \quad \text{et} \quad \|\Phi_i\| = 1$$

De plus, elle peut être considérée comme optimale dans le sens où elle conduit à la meilleure concentration de l'énergie (variance

de  $X$ ) dans les variables  $y_n$ . De fait, l'énergie est maximum pour  $y_1$  et décroît quand l'ordre de la variable augmente.

$$\text{var}(y_1) > \text{var}(y_2) > \dots > \text{var}(y_N)$$

On démontre que la variance de  $y_n$  peut être estimée par la valeur propre  $\lambda_n$  :

$$\text{var}(y_n) = \widehat{E}(y_n^2) = Y_n \cdot Y_n^T$$

or,

$$C_Y = \Lambda = \Phi \cdot C_X \cdot \Phi^T = Y \cdot Y^T$$

d'où

$$\text{var}(y_n) = \lambda_n \quad (6)$$

Puisque le vecteur  $X$  est pris au travers d'une fenêtre bidimensionnelle déplacée en glissant sur l'image, on peut établir la correspondance entre la transformation linéaire d'un vecteur  $X$  par un vecteur propre  $\Phi_n$  et le produit de convolution bidimensionnel de la fenêtre d'observation  $X_{(i,j)}$  par un filtre RIF  $\Phi_{n(i,j)}$ . Ce filtre est formé des coefficients du vecteur propre  $\Phi_n$  réordonnés suivant leur position dans la fenêtre d'observation  $X_{(i,j)}$  comme l'illustre la figure 4 pour une fenêtre carrée de  $3 \times 3$  pixels.

$$\Phi_n = \{\phi_{n1}, \phi_{n2}, \dots, \phi_{n9}\}^T \Leftrightarrow \Phi_{n(i,j)} = \begin{bmatrix} \phi_{n1} & \phi_{n2} & \phi_{n3} \\ \phi_{n4} & \phi_{n5} & \phi_{n6} \\ \phi_{n7} & \phi_{n8} & \phi_{n9} \end{bmatrix}$$

un vecteur propre  $\Leftrightarrow$  un filtre RIF

Figure 4 : Un vecteur propre mis sous la forme d'un filtre RIF bidimensionnel

En conséquence, le produit scalaire de la transformation s'exprime comme un produit de convolution bidimensionnelle noté  $**$  :

$$y_n = \Phi_n^T \cdot X \Leftrightarrow y_{n(i,j)} = X_{(i,j)} ** \Phi_{n(i,j)}$$

L'image décrite par la variable  $y_{n(i,j)}$  est appelée *image ou canal filtré* [11].

### 3.3. INTÉRÊT DE LA TKL POUR L'ANALYSE DE TEXTURE

L'intérêt principal de la TKL réside dans le fait qu'aucune hypothèse n'est posée sur les fonctions de densité de probabilité  $p(X)$ . Elle s'applique donc sans restriction à de nombreuses textures.

La TKL présente pour l'analyse de texture deux aspects intéressants :

#### 1) Adaptabilité à la structure de la texture

La TKL conduit pour une texture donnée à un ensemble de filtres RIF unique. Chaque image filtrée étant de plus décorrélée des autres (5).

L'application de la TKL à des images de synthèse (figure 5) montre sa capacité à analyser des textures en les décomposant

en plusieurs composantes structurelles. On peut constater par exemple que le premier filtre extrait de l'image "grille" (figure 5) reflète fidèlement la structure de cette image. Il est aisé de constater que c'est le filtre RIF qui extrait le maximum de la variance de l'image. Il comporte neuf composantes car la fenêtre d'observation utilisée est carrée de taille  $3 \times 3$  pixels. Les huit autres filtres ne sont pas représentés, leur contribution à l'énergie totale est insignifiante (nulle dans ce cas).

La part de la variance portée par une variable  $y_{n(i,j)}$  est souvent exprimée en pourcentage d'inertie [12]. Le pourcentage d'inertie porté par une image filtrée est défini par :

$$I(y_{n(i,j)}) = \lambda_n / \text{Trace}(\Lambda)$$

L'image "damier" conduit aussi à un seul filtre RIF.

L'image "diagonale" conduit à deux filtres orthogonaux, répartissant l'énergie de la texture en 55,56% pour le filtre numéro 1 et 44,44% pour le filtre numéro 2.

La TKL permet donc une caractérisation fidèle des microtextures à partir d'un modèle linéaire d'étude de la dépendance locale entre pixels.

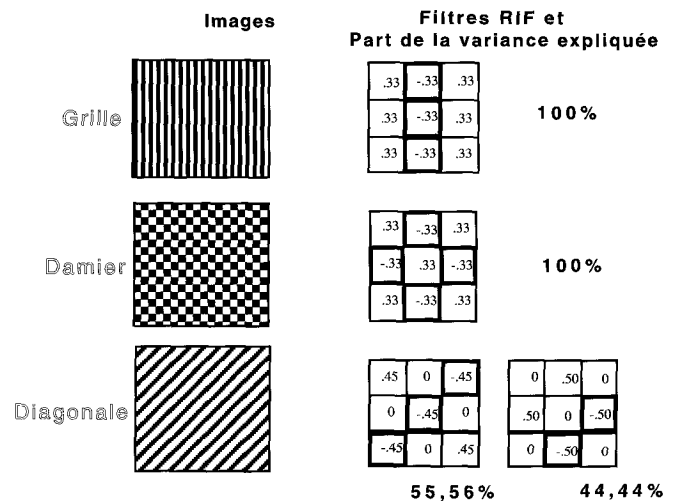


Figure 5 : Textures de synthèse et filtres RIF correspondants.

#### 2) Optimalité de la transformation

La TKL est optimale dans le sens où elle assure une concentration énergétique maximale de "l'énergie de texture" dans les premières images filtrées, et une conservation globale de cette énergie.

On vérifie l'équation :

$$\text{Trace}(C_X) = \text{Trace}(C_Y)$$

Si le vecteur  $X$  comporte  $N$  variables alors son énergie totale vaut :

$$\text{Trace}(C_X) = \sum_{n=1}^N \hat{\sigma}_{x_n}^2$$

Puisque les variables  $x_n$  couvrent, à cause du balayage de l'image par la fenêtre d'observation, une même zone de l'image, alors elles

sont identiquement distribuées. Ainsi, les moments associés à ces variables sont tous identiques.

$$E(x_1^{k_1}) = E(x_2^{k_1}) = \dots = E(x_N^{k_1}) \quad \forall k_1 \in N^* \quad (7)$$

On peut donc écrire :

$$Trace(\mathbf{C}_X) = N \cdot \hat{\sigma}_{x_n}^2$$

Ainsi, l'énergie totale de la texture est égale à  $N$  fois l'énergie d'une composante  $x_n$  du vecteur de mesure  $X$ .

L'énergie totale après transformation s'obtient facilement à partir de la matrice des valeurs propres. Ce qui s'écrit :

$$Trace(\mathbf{C}_Y) = Trace(\mathbf{\Lambda}) = Trace(\mathbf{\Phi} \cdot \mathbf{C}_X \cdot \mathbf{\Phi}^T) \quad (8)$$

Or, comme

$$Trace(\mathbf{\Phi} \cdot \mathbf{C}_X \cdot \mathbf{\Phi}^T) = Trace(\mathbf{C}_X \cdot \mathbf{\Phi}^T \cdot \mathbf{\Phi})$$

et,

$$\mathbf{\Phi} \cdot \mathbf{\Phi}^T = \mathbf{I}$$

On vérifie l'égalité (3), puisque  $\mathbf{\Lambda}$  est une matrice diagonale. Ce qui donne :

$$N \cdot \hat{\sigma}_{x_n}^2 = \sum_{n=1}^N \lambda_n$$

La TKL est une transformation sans perte d'information (matrice  $\mathbf{\Phi}$  unitaire) qui permet une caractérisation fidèle des microtextures grâce à des bancs de filtres adaptés, ainsi qu'une représentation optimale des textures.

### 3.4. DÉFINITION DES ATTRIBUTS DE TEXTURE

La détermination des filtres optimaux adaptés est fondée sur l'analyse de l'ensemble des statistiques d'ordre 2 du vecteur de mesure. Malheureusement, ces statistiques sont insuffisantes pour différencier plusieurs textures [3].

Une caractérisation plus complète de la fonction de densité de probabilité  $p(X)$  peut être obtenue en évaluant la fonction de densité de probabilité  $p(Y)$  dans l'espace transformé. Ces fonctions sont liées par la relation suivante :

$$p_Y(Y) = \mathbf{\Phi}^{-1} \cdot p_X(X = \mathbf{\Phi}^{-1} \cdot Y)$$

On pourra aisément constater que la transformation  $\mathbf{\Phi}$  modifie sensiblement les grandeurs statistiques associées à chaque composante  $x_n$  du vecteur de mesure. En effet, celles-ci sont fortement affectées par les relations de dépendance spatiale entre les pixels pris en compte lors de la transformation. Ainsi, si elles étaient identiques dans l'espace d'origine, elles différeront dans l'espace transformé.

Là où les statistiques du premier ordre  $p(x_n)$  du vecteur de mesure  $X$  sont insuffisantes (7) pour caractériser la fonction de densité de probabilité  $p(X)$ , (ou de manière équivalente l'ensemble des moments  $E(x_n^{k_1})$   $k_1 \in N^*$ ), les statistiques du premier ordre

$p(y_n)$  du vecteur transformé  $Y$  seront par contre généralement suffisantes. Elles correspondent à une caractérisation partielle de  $p(X)$  au moyen de projections dans l'espace propre généré par la TKL.

On choisira donc pour caractériser les textures de calculer l'ensemble des histogrammes des images filtrées.

L'histogramme associé à une image filtrée, quantifiée sur  $N_g$  niveaux de gris, vaut :

$$p(y_n) = \{P(y_n = q) / q = 0, \dots, N_g - 1\}$$

Une caractérisation équivalente peut-être obtenue à partir des moments associés à chaque image filtrée. Un moment d'ordre  $k_1$  vaut  $E(y_n^{k_1})$ .

Souvent on se limite pour des raisons pratiques aux moments centrés de degré inférieur à cinq. Dans notre cas nous utiliserons uniquement le moment centré d'ordre deux  $\hat{\sigma}_{y_n(i,j)}^2$  estimé spatialement dans une image filtrée à l'intérieur d'une fenêtre de dimension  $K \times L$  :

$$\hat{\sigma}_{y_n(i,j)}^2 = \frac{1}{K \cdot L} \sum_{k=-K/2}^{K/2} \sum_{l=-L/2}^{L/2} (y_n(i+k, j+l) - \bar{y}_{(i,j)})^2 \quad (9)$$

C'est une mesure de l'énergie moyenne locale dans un voisinage  $K \times L$  qui correspond à la définition de l'énergie de texture proposée par Laws (1).

### 3.5. EXEMPLE D'ANALYSE DE TEXTURES NATURELLES

A titre d'illustration de la méthode nous présentons les résultats de l'analyse de plusieurs textures naturelles (figure 6). Ces images ont été étudiées au travers d'une fenêtre carrée de  $3 \times 3$  pixels.

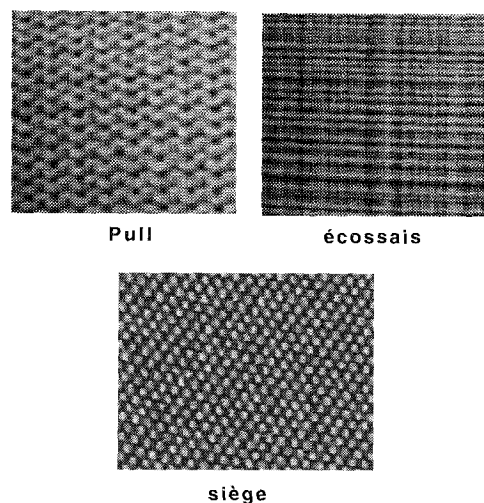


Figure 6 : Trois textures naturelles

La figure 7 présente quatre des neuf filtres RIF extraits des images "écossais" et "pull". Ces filtres représentent, dans les trois cas, près de 98% de l'inertie totale. Ils montrent que chacune de ces

textures présente des structures spatiales marquées suivant des orientations privilégiées.

Le premier filtre de la texture "écossais" (gradient horizontal) permet d'extraire 49% de l'énergie de texture. Le second (moyen-nour) indique que 38% de l'énergie est contenue dans des régions de niveau de gris constant.

La TKL génère des bancs de filtres adaptés à chaque texture : généralement des gradients directionnels horizontaux, verticaux, diagonaux et composés.

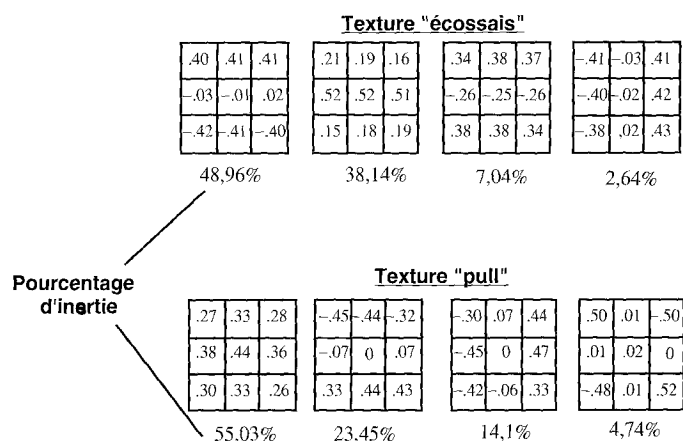


Figure 7 : Textures naturelles et filtres RIF correspondants.

La figure 8 montre les quatre premières images filtrées de la texture "écossais". Chaque image met en valeur une directivité particulière de la texture orientée dans la direction du gradient. Ses structures partielles sont isolées. La perception du contraste de l'image diminue progressivement lorsque l'indice du filtre augmente (diminution de la part de la variance portée par l'image).

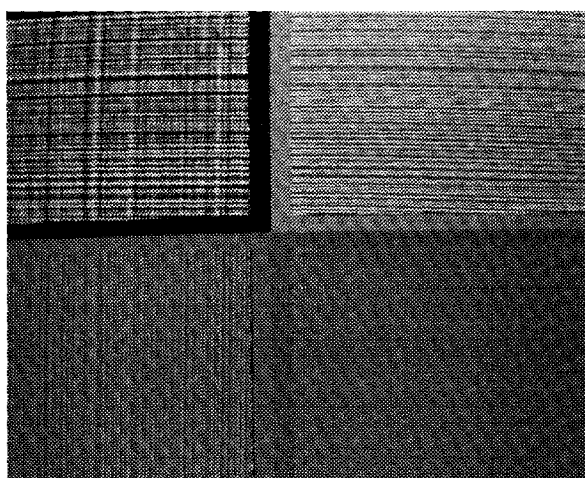


Figure 8 : Images des quatre premières images filtrées de la texture "écossais".

La figure 9a présente les histogrammes des quatre images filtrées de la texture "écossais", représentées à la figure 8, et la figure 9b ceux obtenus avec l'image de texture "pull" de la figure 7.

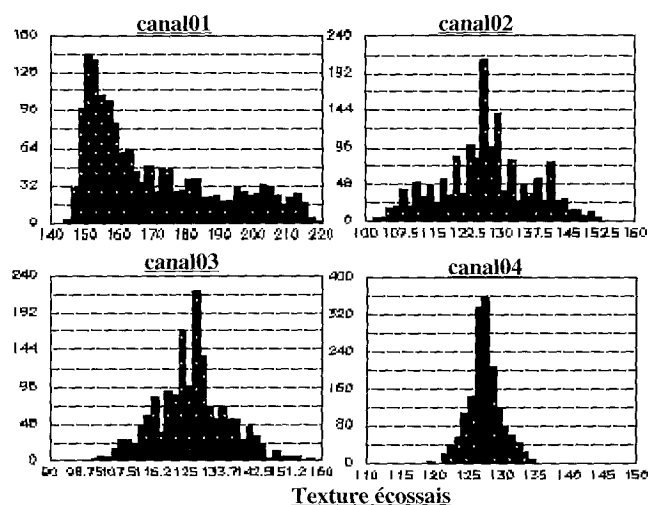


Figure 9a : Histogrammes des 4 canaux filtrés de la texture "écossais".

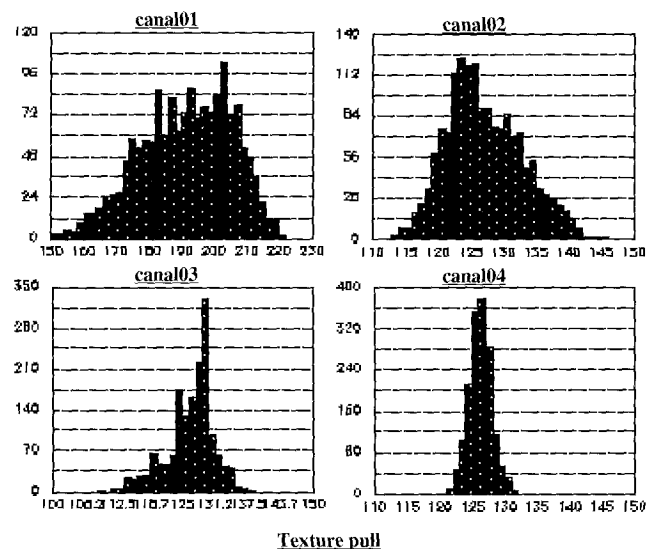


Figure 9b : Histogrammes des 4 canaux filtrés de la texture "pull".

La distribution des niveaux de gris est différente pour chacune de ces images, et la dispersion des niveaux de gris diminue lorsque l'indice d'image diminue (la variance diminue en conséquence).

## 4. Segmentation supervisée d'images de texture

### 4.1. PRINCIPE DE LA SEGMENTATION SUPERVISÉE

La segmentation d'images texturées est abordée comme un problème de classement. Chaque classe est formée par l'ensemble des pixels ayant des attributs de texture similaires. L'indice de similarité d'un individu (pixel) par rapport à une classe de texture particulière  $q$  peut être mesuré par la distance (géométrique) le séparant des individus moyens représentatif de ces classes dans l'espace des attributs de texture.

Les méthodes de classement se différencient des méthodes de classification par le fait qu'elles nécessitent des connaissances a priori. Nous verrons comment disposer de ces informations avec la TKL.

Pour réaliser la segmentation de l'image par classement, il convient d'en définir la procédure (figure 10).

On distingue deux étapes [13] :

- Phase d'apprentissage : des vecteurs prototypes représentatifs de chaque classe de texture sont déterminés (centres de gravité),
- Phase de reconnaissance : les pixels sont regroupés en classes de propriétés voisines suivant les valeurs prises par leurs attributs de texture.

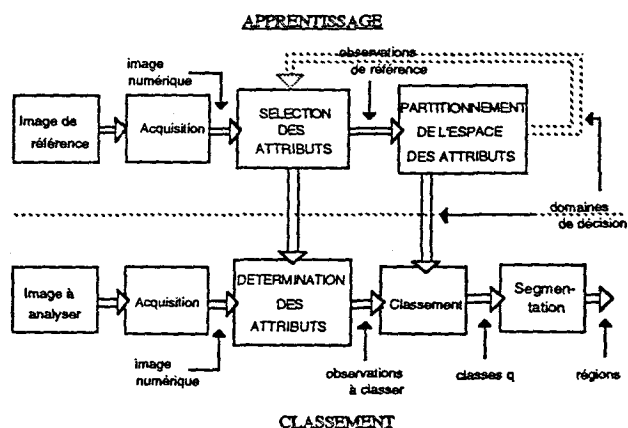


Figure 10 : Schéma fonctionnel d'une procédure de segmentation par classement

Plusieurs points doivent être résolus pour mettre en oeuvre cette procédure :

- Quels sont les attributs de texture à retenir pour caractériser les pixels de l'image à segmenter ?
- Quels sont les individus représentatifs de chaque classe de l'image et comment les déterminer ?
- Comment partitionner l'espace des attributs pour obtenir une bonne discrimination des classes de texture ?
- Quel indice de similarité prendre et quelle règle de décision choisir ?

Un moyen de répondre à la plupart de ces questions est d'étudier la distribution du vecteur des attributs de texture et de la modéliser.

## 4.2. DISTRIBUTION DES ATTRIBUTS DE TEXTURE

### 4.2.1. Mesure d'énergie de texture

Dans le modèle de la figure 1 les attributs de texture sont calculés après extraction des microtextures. Nous utiliserons l'estimateur de la variance locale dans les images filtrées défini en (9) comme attribut de texture. Nous montrerons par la suite que cet estimateur est non biaisé, et que sa valeur moyenne est égale à la valeur

propre associée au filtre RIF ayant servi à filtrer l'image. Pour cela posons :

$$\widehat{E}(\widehat{\sigma}_{y_n(i,j)}^2) = \frac{1}{G.H} \sum_{g=-G/2}^{G/2} \sum_{h=-H/2}^{H/2} \widehat{\sigma}_{y_n(i+g,j+h)}^2 \quad (10)$$

Espérance mathématique de la variance locale estimée spatialement sur le domaine  $G \times H$  de l'image (figure 11).

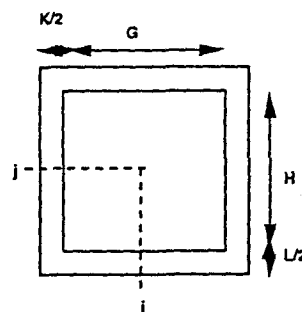


Figure 11 : Fenêtre de calcul de la variance locale moyenne.

Si l'image est à valeur moyenne nulle alors on peut écrire :

$$\widehat{E}(\widehat{\sigma}_{y_n(i,j)}^2) = \frac{1}{G.H} \sum_{g=-G/2}^{G/2} \sum_{h=-H/2}^{H/2} \frac{1}{K.L} \sum_{k=-K/2}^{K/2} \sum_{l=-L/2}^{L/2} (y_n(i+k,g,j+l+h))^2$$

Ou encore :

$$\widehat{E}(\widehat{\sigma}_{y_n(i,j)}^2) = \frac{1}{K.L} \sum_{k=-K/2}^{K/2} \sum_{l=-L/2}^{L/2} \frac{1}{G.H} \sum_{g=-G/2}^{G/2} \sum_{h=-H/2}^{H/2} (y_n(i+k,g,j+l+h))^2$$

Si  $G$  et  $H$  définissent entièrement le domaine de l'image, et si les pixels en dehors de ce domaine sont de niveau de gris nul, donc :

$$\sum_{g=-G/2}^{G/2} \sum_{h=-H/2}^{H/2} (y_n(i+k,g,j+l+h))^2 = \widehat{\sigma}_{y_n(i+k,j+l)}^2$$

Où  $\widehat{\sigma}_{y_n(i+k,j+l)}^2$  représente la variance moyenne de l'image sur le domaine  $G \times H$ . Soit en l'occurrence les valeurs propres  $\lambda_n$ . Les versions translattées de  $\widehat{\sigma}_{y_n(i+k,j+l)}^2$  pour les valeurs de  $k$  et  $l$  dans la limite fixée par la figure 11 sont identiques. D'où,

$$\frac{1}{K.L} \sum_{k=-K/2}^{K/2} \sum_{l=-L/2}^{L/2} \widehat{\sigma}_{y_n(i+k,j+l)}^2 = \frac{1}{K.L} \sum_{k=-K/2}^{K/2} \sum_{l=-L/2}^{L/2} \lambda_n = \lambda_n$$

Cet estimateur est sans biais et correspond bien à la variance moyenne de l'image définie par la TKL. Soit :

$$\widehat{E}(\widehat{\sigma}_{y_n(i,j)}^2) = \lambda_n$$

La figure 12 illustre la concordance de cet estimateur vers les valeurs propres calculées sur les trois images de textures naturelles de la figure 6.

	Texture "pull"		Texture "siège"		Texture "écossais"	
Canal	$\hat{E}(\hat{\sigma}_{y_n}^2)$	$\lambda_i$	$\hat{E}(\hat{\sigma}_{y_n}^2)$	$\lambda_i$	$\hat{E}(\hat{\sigma}_{y_n}^2)$	$\lambda_i$
1	4871	4876	9187	9182	9611	9678
2	2064	2087	4123	4204	7683	7538
3	1338	1342	3213	3160	1516	1392

Figure 12 : Comparaison variances locales moyennes/valeurs propres.

En conséquence, on peut dire que la distribution des attributs de texture sera centrée sur les valeurs propres de la TKL. Un léger biais peut intervenir si les pixels en dehors du domaine  $G \times H$  ne sont pas à valeur nulle (minimisé en choisissant  $G \times H \gg K \times L$ ).

#### 4.2.2. Distribution des attributs de textures naturelles

La figure 13 présente la distribution des attributs (variances locales dans les canaux (9)) de la texture "écossais" pour une fenêtre  $K \times L$  de  $7 \times 7$  pixels.

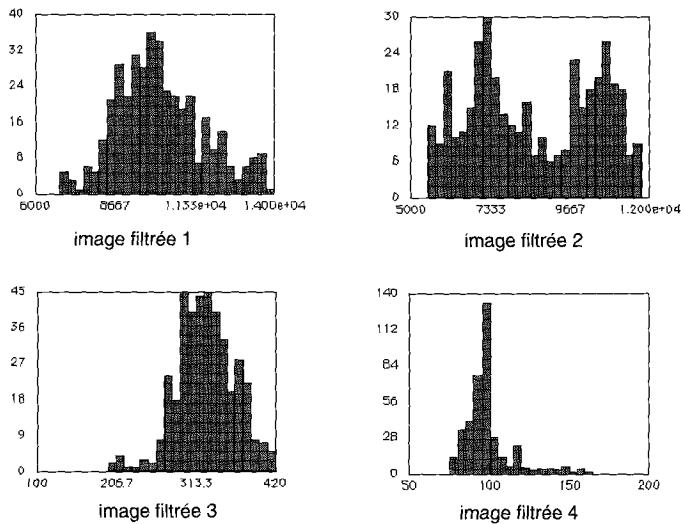


Figure 13 : Histogrammes des attributs de la texture "écossais".

Ces distributions sont bien centrées sur les valeurs propres de la TKL. Trois histogrammes sur quatre sont unimodaux, on pourra donc facilement les approcher par une loi normale. L'expression analytique de la fonction de densité de probabilité  $p(Z)$  du vecteur des attributs de texture s'écrit alors :

$$p(Z) = \frac{1}{(2\pi)^{\frac{N}{2}} |\mathbf{S}_Z|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(Z - \bar{Z})^T \mathbf{S}_Z^{-1}(Z - \bar{Z})\right) \quad (11)$$

$Z = \{\hat{\sigma}_{y_1}^2, \hat{\sigma}_{y_2}^2, \dots, \hat{\sigma}_{y_N}^2\}^T$ , est le vecteur des attributs de texture,  $\mathbf{S}_Z$  sa matrice de covariance.

L'approximation normale (11) est valable pour  $p(Z)$  tant que l'hypothèse sur les variables  $y_n$  (aléatoires et identiquement distribuées) reste vraie. En effet, le théorème de la limite centrale [14] garantit une convergence asymptotique vers une loi normale si la sommation de ces variables porte sur un nombre suffisamment grand de valeurs, soit en l'occurrence si la taille de la fenêtre  $K \times L$  est grande. La taille minimum de cette fenêtre définit la résolution de texture.

La figure 14 montre l'influence de cette fenêtre sur la distribution de la troisième variable du vecteur des attributs  $Z$  de la texture "siège".

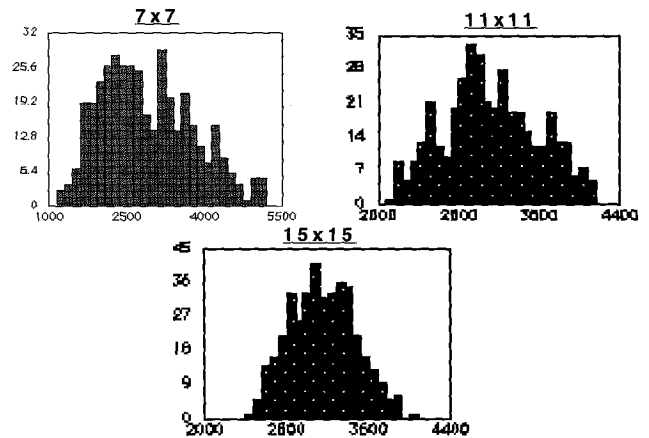


Figure 14 : Effet de la taille du moyenneur sur la distribution des attributs de la texture "siège".

Plus la taille du moyenneur sur les attributs de texture est grande, plus la densité  $p(Z)$  tend à être unimodale, plus elle sera facile à approcher par une loi normale. L'attribut retenu est une mesure globale calculée à partir de caractéristiques locales décrivant les microtextures. Ceci est en adéquation avec la définition de la notion de texture que nous avons adoptée (figure 1).

Avec la TKL, la problématique souvent délicate du choix des attributs de texture les plus pertinents se trouve donc résolue. Il suffit simplement de retenir l'estimateur de la variance locale dans les images filtrées.

### 4.3. SEGMENTATION D'IMAGES TEXTURÉES PAR TRANSFORMATIONS ADAPTÉES MULTIPLES

#### 4.3.1. Principe de la méthode

Nous avons précédemment vu que la TKL fournit pour toute texture un banc de filtres RIF adapté à sa structure spatiale. A fortiori, on peut supposer que les distributions des attributs de texture seront différentes elles aussi.

Pour affecter chaque pixel à une classe de texture particulière, on va postuler que son énergie de texture se répartit de manière optimale lorsque le voisinage du pixel est transformé par son banc de filtres adapté et que toute autre transformation produira une répartition forcément différente de l'optimum.



Cette hypothèse a été vérifiée expérimentalement sur des images de textures naturelles.

En exemple, sont présentés les histogrammes des attributs de texture des images "pull" et "siège" transformées par le banc de filtres adapté extrait de la texture "pull" (figure 15).

On notera par la suite  $\Phi^{(q)}$  le banc adapté à la texture  $q$ , et  $L^{(q)}$  le vecteur de ses valeurs propres associées.

$$L^{(q)} = \{\lambda_1^{(q)}, \lambda_2^{(q)}, \dots, \lambda_N^{(q)}\}^T \quad (12)$$

La variable  $q$  prend  $Q$  modalités suivant le nombre de textures de l'image.

Il apparaît clairement que les individus "pull" se distribuent autour des valeurs propres  $L^{(pull)}$  de la base "pull", alors que ceux de la texture "siège" se répartissent autour de valeurs a priori quelconques lorsqu'ils sont exprimés dans la base "pull".

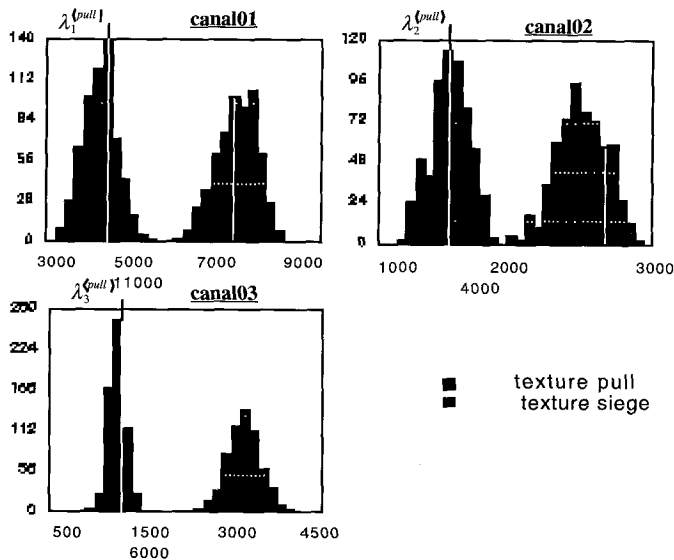


Figure 15 : Attributs des textures "pull" et "siège" filtrés par le banc de filtres  $\Phi^{(pull)}$  (voir aussi page 56).

Nous avons ensuite appliqué l'estimateur de la variance locale (10) à trois images de texture après transformation par les trois bancs de filtres adaptés (figure 16). Les estimations concordent avec les valeurs du vecteur  $L^{(q)}$  lorsqu'une texture  $q$  est filtrée par son banc de filtres propre  $\Phi^{(q)}$  (cases ombrées) mais diffèrent foncièrement lorsque l'image est transformée par un autre banc (cases claires). Ceci confirme le postulat émis précédemment qui supposait que la répartition optimale de l'énergie d'une texture  $q$  est assurée par sa transformation adaptée  $\Phi^{(q)}$ .

Ainsi, l'affectation d'un individu inconnu à une classe de texture revient à trouver la transformation  $\Phi^{(q)}$  qui assure une répartition optimale de son énergie de texture. La discrimination entre les classes sera d'autant meilleure que les individus moyens représentatifs de chaque classe seront différents. Dans notre approche, ils sont connus et ont pour valeur les vecteurs prototypes  $L^{(q)}$  formés par l'ensemble des valeurs propres issues des transformations  $\Phi^{(q)}$ .

Canal	base $\Phi^{(pull)}$			base $\Phi^{(siège)}$			base $\Phi^{(écossais)}$			
	1	2	3	1	2	3	1	2	3	
$\lambda_i$	4876	2087	1342	9182	4204	3160	9678	7538	1392	
variance	*pull*	4871	2064	1338	4929	1341	2067	2066	4288	709
moyenne	*siège*	8945	3198	4107	9187	4223	3213	3229	6903	1605
des images	*écossais*	6212	9120	748	6120	882	9024	9611	7653	3516

Figure 16 : Estimation des attributs moyens de 3 textures naturelles.

Fort de cette constatation, on établit la règle de décision suivante : On calcule la distance  $d$  séparant le vecteur des attributs de texture du pixel  $X_{(i,j)}$  et les vecteurs  $L^{(g)}$  des classes  $g$ . Le minimum pour  $d$  détermine sa classe d'affectation.

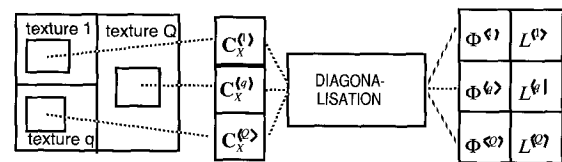
Soit :

$Z_{(i,j)}^{(q)}$  appartient à la classe de texture  $q$  si

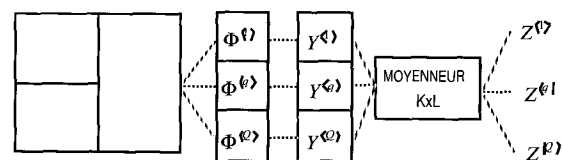
$$d(Z_{(i,j)}^{(q)}, L^{(q)}) = \min_g (d(Z_{(i,j)}^{(g)}, L^{(g)})) \quad g \in \{1, \dots, Q\} \quad (13)$$

Les vecteurs  $L^{(q)}$  sont les centres de gravité des  $q$  classes de texture.

Nous venons donc de définir le principe général de la segmentation d'une image de texture par transformations multiples adaptées. Le schéma synoptique d'extraction des attributs est présenté à la figure 17.



Calcul des bancs de filtres adaptés



Transformation de l'image et calcul des attributs

Figure 17 : Attributs de texture par transformations adaptées multiples

- Premièrement, on sélectionne des échantillons représentatifs des  $q$  textures de l'image (fenêtres d'apprentissage des caractéristiques textuelles) afin de déterminer les bancs de filtres adaptés  $\Phi^{(q)}$ .
- Deuxièmement, on applique les bancs de filtres  $\Phi^{(q)}$  à l'image, et on calcule les attributs de texture  $Z_{(i,j)}^{(q)}$ .

#### 4.3.2. Choix du critère de décision

Dans le schéma de la figure 10, la partition de l'espace des attributs durant la phase d'apprentissage repose sur le choix du critère

de décision. Chaque région définie dans l'espace des attributs correspond une classe de texture. Nous avons testé deux types d'approche : géométrique et statistique.

### 1) Critère géométrique : distance Euclidienne

Pour les règles de décision fondées sur une mesure de distance on dispose d'informations a priori : les centres de gravité des classes. On identifie la classe d'appartenance d'un pixel en mesurant ses degrés de similarité par rapport à l'ensemble des classes.

La distance Euclidienne est couramment utilisée et s'exprime dans notre cas comme suit :

$$d(Z_{(i,j)}^{(q)}, L^{(q)}) = \left( \sum_{n=1}^N (z_{n(i,j)}^{(q)} - \lambda_n^{(q)})^2 \right)^{\frac{1}{2}} \quad (14)$$

La règle de décision (13) est fondée sur la minimisation de cette distance.

La mise œuvre de cette procédure de classement est rapide et efficace. Avec  $Q$  classes et un vecteur d'attribut de taille  $N$  elle nécessite  $Q.N.(N-1)$  additions et  $2.Q.N$  multiplications [15].

Afin que le poids des variables  $z_{n(i,j)}^{(q)}$  dans la règle de décision soit identique on choisira de les réduire. Ainsi, le pouvoir discriminant d'une composante de  $Z$  n'est plus dépendant de son inertie propre. La nouvelle distance est :

$$d(Z_{(i,j)}^{(q)}, L^{(q)}) = \left( \sum_{n=1}^N \left( \frac{z_{n(i,j)}^{(q)} - \lambda_n^{(q)}}{\sigma_{z_n^{(q)}}} \right)^2 \right)^{\frac{1}{2}} \quad (15)$$

### 2) Critère statistique : critère de Bayes

La théorie de la décision, souvent abordée statistiquement, tient compte des statistiques des variables décrivant les images.

Trouver la classe d'affectation d'un pixel revient à déterminer ses probabilités d'appartenance a posteriori par rapport à l'ensemble des classes  $q$ , notées  $P(q/Z)$ .

La règle de décision fondée sur le critère de Bayes est appelée test du maximum a posteriori (MAP), car on affecte un pixel à la classe  $q$  pour laquelle  $P(q/Z)$  est maximum.

$$Z \in q \Rightarrow P(q/Z) = \max_g (P(g/Z)) \quad g \in \{1, \dots, Q\}$$

Nous supposons que les classes sont équiprobablement réparties. Ce qui sera toujours le cas lorsque l'on segmentera une image de textures naturelles. En effet, on ne peut savoir a priori la proportion de pixels appartenant à chacune d'elles. La probabilité a priori de chaque classe est donc :

$$P(q) = \frac{1}{Q} \quad q \in \{1, \dots, Q\}$$

Si les fonctions de densité de probabilité conditionnelles  $p(Z/q)$  du vecteur des attributs de texture sont normales, elles s'expriment analytiquement par :

$$p(Z/q) = \frac{1}{(2\pi)^{\frac{N}{2}} |\mathbf{S}_Z^{(q)}|^{\frac{1}{2}}} \exp(-d(Z^{(q)}, \bar{Z}^{(q)})) \quad (16)$$

avec  $d$  : distance de Mahalanobis du vecteur  $Z$  par rapport à la classe  $q$ , s'exprimant comme :

$$d(Z^{(q)}, \bar{Z}^{(q)}) = \frac{1}{2} (Z^{(q)} - \bar{Z}^{(q)})^T \mathbf{S}_Z^{(q)-1} (Z^{(q)} - \bar{Z}^{(q)})$$

$\bar{Z}^{(q)}$  est le vecteur moyen défini en (12), et  $\mathbf{S}_Z^{(q)}$  la matrice de covariance intraclasse des classes  $q$ .

Dans le cas de distributions normales, le critère de Bayes est optimal [16] et peut s'exprimer sous forme logarithmique. Affecter un pixel à une classe  $q$  revient à trouver le minimum de l'expression suivante :

$$\min_g (\ln(\mathbf{S}_Z^{(g)}) + d(Z^{(g)}, L^{(g)})) \quad g \in \{1, \dots, Q\}$$

Cette expression fait intervenir la distance de Mahalanobis par rapport aux classes  $g$  et un terme constant  $\ln(\mathbf{S}_Z^{(g)})$ .

Les paramètres  $L^{(q)}$  et  $\mathbf{S}_Z^{(q)}$  des distributions  $p(Z/q)$  sont des informations a priori calculées par apprentissage sur des régions de texture de l'image. A chaque texture  $q$  sera associé un banc de filtres  $\Phi^{(q)}$ , un vecteur prototype  $L^{(q)}$  et une matrice de covariance  $\mathbf{S}_Z^{(q)}$ .

#### Remarque :

La distance de Mahalanobis ne peut être définie que si les matrices  $\mathbf{S}_Z^{(g)}$  sont inversibles. De plus, le calcul de la distance de Mahalanobis est coûteux en nombre d'opérations élémentaires. On peut le réduire [15] en décomposant les matrices de covariance intraclasses en vecteurs propres  $\Phi_{Z_n}^{(g)}$  et valeurs propres  $\Lambda_{Z_n}^{(g)}$ . Le test de Bayes devient :

$$\min_g \left( \sum_{n=1}^N \ln(\lambda_{Z_n}^{(g)}) + \sum_{n=1}^N \frac{1}{\lambda_{Z_n}^{(g)}} \left( \Phi_{Z_n}^{(g)T} (Z^{(g)} - L^{(g)}) \right)^2 \right) \quad g \in \{1, \dots, Q\} \quad (17)$$

Le premier terme de cette expression est constant. Il sera évalué en initialisation de l'algorithme. Il faut  $Q.N(N+2)$  multiplications et  $N^2Q$  additions pour classer un pixel.

## 4.4. EXEMPLE DE SEGMENTATION D'IMAGES DE TEXTURES NATURELLES

### 4.4.1. Considérations générales sur la méthode

La segmentation d'images de texture dans le cadre des transformations adaptées multiples suit les procédures décrites figures 10 et 17.

• *Apprentissage* : Les bancs de filtres adaptés de chaque texture de l'image sont déterminés ainsi que leur vecteur prototype, à partir d'une petite région représentative (typiquement  $32 \times 32$  pixels), puis les attributs de texture après filtrage de l'image, et enfin, leurs fonctions de densité de probabilité conditionnelle suivant le modèle (16). Puisque celles-ci sont supposées normales elles sont donc entièrement décrites par leur vecteur moyen et leur matrice de covariance.

• *Classification* : Nous avons testé les deux critères de décision : géométrique (15) et statistique (17).

Des groupes de pixels de classe d'appartenance connue sont classés par ces deux méthodes. Ceci permet d'éprouver les performances des règles de décision ainsi que la pertinence des attributs de texture. Comme en analyse factorielle discriminante (AFD), on quantifie la qualité de la classification en comptant le nombre de pixels "bien classés". Cette vérification est essentielle pour débiter la classification de l'image entière. On peut en effet modifier certains paramètres de la méthode pour améliorer le classement.

Plusieurs paramètres peuvent être modifiés :

### 1) Forme et taille de la fenêtre d'analyse (figure 3)

Nous avons choisi une fenêtre carrée car celle-ci est simple à mettre en oeuvre. Sa taille influe sur la finesse de l'analyse. En effet, plus elle est large, plus les relations structurelles seront finement étudiées. En contre partie, plus la complexité de la méthode augmente; le nombre de variables devient important. Nous avons travaillé avec des fenêtres de  $3 \times 3$  pixels.

### 2) Taille du moyennage $K \times L$ sur les attributs de texture (9)

La distribution des attributs de texture dépend fortement du moyennage sur les attributs de texture (figure 14). Nous avons vu que l'approximation gaussienne est d'autant meilleure que la sommation porte sur un nombre important d'échantillons. Cependant, dans le cadre de la segmentation des images de texture, une remarque importante s'impose. A la frontière de deux textures voisines, il y a recouvrement des fenêtres, donc la valeur de l'attribut est un mélange des propriétés de ces deux textures. La confusion entraînée augmente en fonction de  $K \times L$ . Il faut donc choisir la taille minimale qui n'occasionne pas d'apparition d'hétérogénéités à l'intérieur des textures (des confusions peuvent survenir dans les classes). Ce qui s'explique par le fait que les histogrammes des attributs ne sont plus unimodaux et aussi parce que la dispersion des valeurs est plus grande. La limite inférieure pour le moyennage est fixée par la résolution de la texture. Ce compromis est difficile à régler de manière certaine. Seul un critère visuel est actuellement utilisé. On juge de la qualité de la discrimination sur des images test dont les frontières des régions texturées sont préalablement connues.

### 3) Nombre d'images filtrées pour la règle de décision

Dans le cas général, la TKL conduit à une concentration maximale de l'énergie de texture dans les premières images filtrées. Ainsi paraît-il logique de ne faire contribuer à la règle de décision que les premières images à condition qu'elles représentent à elles seules un fort pourcentage de l'énergie totale.

Le pouvoir discriminant des images à faible inertie semble minime.

Avec cette hypothèse on peut, comme pour l'analyse en composantes principales (ACP), réduire le nombre de variables nécessaires pour quantifier l'information texturale. Mais ceci ne doit pas se faire au détriment d'une perte de la qualité de la segmentation.

A ce titre deux options sont possibles :

- fixer le nombre d'images filtrées à retenir ( $M \leq N$ )

- fixer le pourcentage d'inertie minimum

$$(I = \sum_{m=1}^M \lambda_m / \text{Trace}(\Lambda) \leq 1).$$

Nous l'avons vérifié sur plusieurs images test.

## 4.4.2. Résultats

Les images de textures naturelles sont composées de plusieurs textures discrétisées à partir d'une caméra de résolution  $512 \times 512$  pixels. Nous présentons les résultats de la segmentation de deux images test de résolution  $256 \times 256$  pixels quantifiées sur 256 niveaux de gris. La première image est composée de deux textures naturelles : siège et écossais (figure 18); la seconde de trois textures : pull, écossais et siège (figure 19).

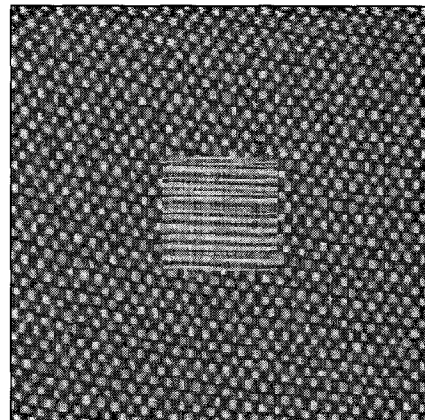


Figure 18 : Image de deux textures naturelles : écossais au centre, siège autour

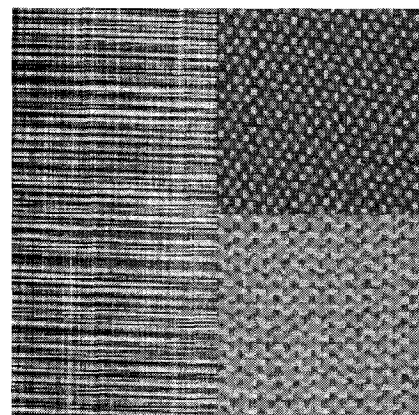


Figure 19 : Image de trois textures naturelles : écossais à gauche, siège et pull à droite

• *Classement de la base d'apprentissage*

La figure 20 donne le résultat en terme de nombre ou pourcentage de bien classés. Ces pixels sont des échantillons représentatifs pris à l'intérieur de chaque texture (pas d'individu aux frontières). Le vecteur de mesure est de taille  $3 \times 3$ , le moyennage sur les attributs est de  $7 \times 7$  pixels.

Les résultats sont meilleurs pour la distance de Mahalanobis (95,8% de bien classés globalement) que pour la distance Euclidienne (90,5%). Un taux inférieur à 100% laisse supposer qu'à l'intérieur des régions certaines zones seront mal classées. La dispersion du vecteur des attributs est trop importante.

Pour la réduire, il suffit d'augmenter la taille du moyennage. On obtient un pourcentage de 100% lorsque celui-ci est de  $13 \times 13$  pixels (figure 21). Au delà de cette limite aucune amélioration, ni dégradation ne se produit.

*classe d'affectation*

		pull	siège	écos	%
<i>classe d'origine</i>	pull	1509	0	91	94,3
	siège	1	1234	365	77
	écos	0	0	1600	100

**Distance euclidienne**

*classe d'affectation*

		pull	siège	écos	%
<i>classe d'origine</i>	pull	1472	128	0	92
	siège	20	1580	0	98,75
	écos	0	52	1548	96,75

**Distance de Mahalanobis**

Figure 20 : Classement de l'ensemble d'apprentissage, moyennage sur  $7 \times 7$  pixels.

		pull	siège	écos	%
<i>classe d'origine</i>	pull	1600	0	0	100
	siège	0	1600	0	100
	écos	0	0	1600	100

Figure 21 : Classement avec un moyennage sur  $13 \times 13$  pixels.

Cette vérification de la qualité de la discrimination permet d'engager la segmentation de l'image entière.

• *Segmentation d'images de texture*

Les figures 22 et 23 montrent les images après classement. Afin de juger de la segmentation, les contours détectés des textures ont été superposés à l'image originale. La réduction du nombre d'images filtrées pour le calcul des deux distances à un pourcentage d'inertie total minimum de 98 % n'entraîne aucune dégradation du résultat. Dans le cas de ces images nous n'avons retenu que les quatre premières images filtrées.

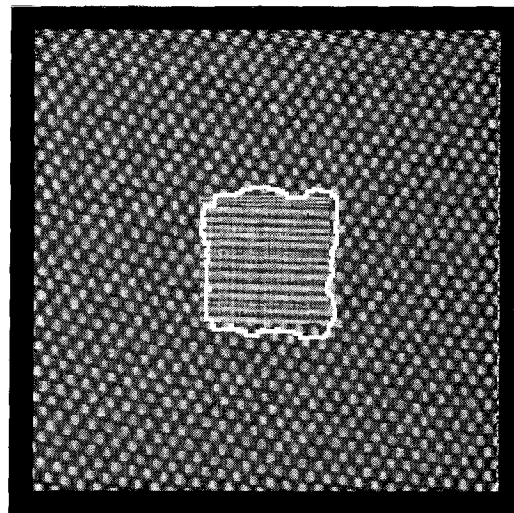


Figure 22 : Segmentation avec la distance Euclidienne (moyennage  $13 \times 13$  pixels)

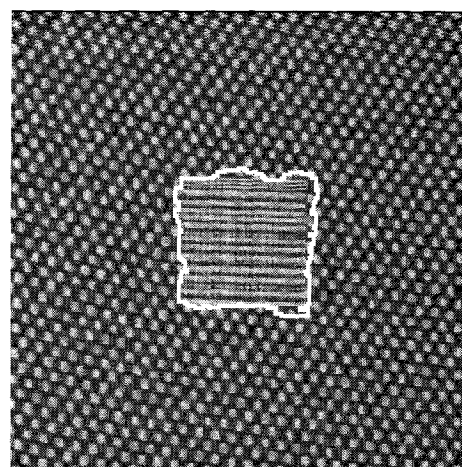


Figure 23 : Segmentation avec la distance de Mahalanobis (moyennage  $13 \times 13$  pixels)

Les figures 24 et 25 montrent le résultat de la segmentation pour une image composée de trois textures naturelles.

Les mesures réalisées avec les images test ont montré la supériorité de l'approche statistique par rapport à l'approche géométrique pour le classement.

Le critère optimal de Bayes s'avère toujours meilleur, même si la distribution des attributs de texture n'est pas parfaitement gaussienne.

## 5. Conclusion

La méthode décrite dans cet article a été appliquée à la segmentation d'images d'échographies cardiaques [15], mais peut évidemment s'appliquer à de nombreuses images texturées (télé-détection satellite, imagerie médicale, contrôle industriel, ...). Elle conduit

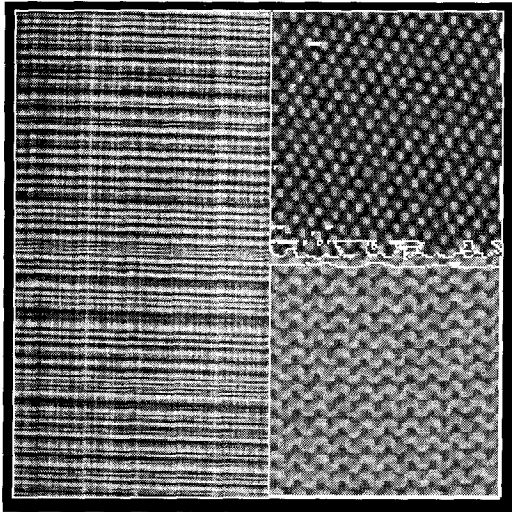


Figure 24 : Segmentation avec la distance Euclidienne (moyenneur 13x13 pixels)

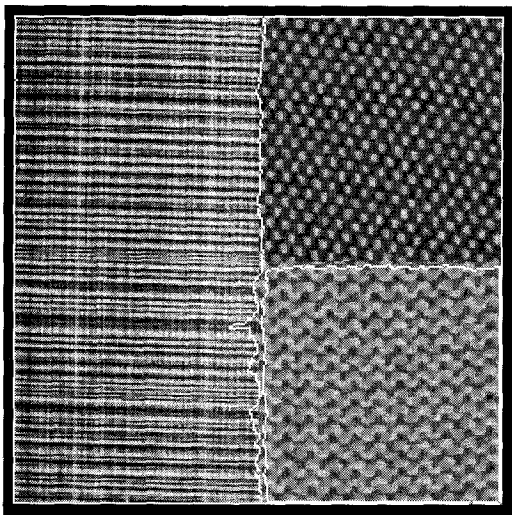


Figure 25 : Segmentation avec la distance de Mahalanobis (moyenneur 13x13 pixels)

à une solution optimale pour la segmentation supervisée d'images de texture.

La phase d'apprentissage de la méthode est basée sur la Transformation de Karhunen-Loeve. Les caractéristiques structurelles de chacune des textures constituant l'image à segmenter sont déterminées par une analyse statistique dans un voisinage fixé par une fenêtre d'observation appelée "fenêtre de voisinage".

La TKL permet une représentation optimale et condensée de l'information structurelle des textures, quel qu'en soit la nature. Elle fournit un système d'analyse de textures optimal sous forme de banc de filtres adaptés.

L'énergie moyenne d'une texture est conservée par le filtrage et se distribue autour des valeurs propres de sa transformation adaptée. Cette propriété essentielle constitue la base de la méthode de classification. Chaque pixel de l'image sera affecté à la classe

de texture pour laquelle le banc de filtres adaptés assure une répartition optimale de son énergie de texture.

Il est généralement possible de réduire les bases de transformation aux premiers filtres totalisant un fort pourcentage de l'énergie de texture, sans pour autant détériorer la qualité de la segmentation.

La TKL apporte une justification de l'utilisation des techniques de filtrage linéaire pour l'analyse de texture. Elle est d'une mise en œuvre relativement aisée et conduit à des solutions rapides pour la segmentation des images. De plus, elle résout la problématique souvent délicate du choix des attributs de texture les plus pertinents.

Manuscrit reçu le 6 janvier 1993.

## BIBLIOGRAPHIE

- [1] A. GAGALOWICZ. *Vers un modèle de textures*. Thèse de doctorat d'état, I.N.R.I.A., Mai 1983
- [2] M. HARALICK. *Statistical and structural approaches to texture*. Proceedings of the IEEE n° 5, 1979, Vol 67, pp 786-804
- [3] B. JULESZ. *Textons, the fundamental elements in preattentive vision and perception of textures*. The Bell System Technical Journal n° 6, 1983, Vol 62, pp 1619-1645
- [4] K. S. FU, S. Y. LU. *Stochastic tree grammar inference for texture synthesis and discrimination*. Computer Graphics and Image Processing n° 3, 1979, Vol 9, pp 246-266
- [5] P.A. DONDES, A. ROSENFELD. *Pixel classification based on gray level and local "Busyness"*. IEEE Transactions on PAMI n° 1, Vol 4, 1982, pp 79-83
- [6] B. JULESZ. *Experiments in the visual perception of texture*. Scientific American 233, 1975, pp 796-804
- [7] K.L. LAWS. *Texture energy measure*. Proceedings on Image Understanding Workshop, 1979, pp 47-51
- [8] D. WERMSEER. *Unsupervised segmentation by use of a texture gradient*. Proceeding of the IEEE, 1984, pp 1114-1116
- [9] M. UNSER. *Description statistique de textures : application à l'inspection automatique*. Thèse n° 534, Ecole Polytechnique Fédérale de Lausanne, 1984
- [10] F. ADE, N. LINS, M. UNSER. *Comparison of various filter sets for defect detection in textiles*. Proceedings of 7th International Conference on Pattern Recognition, Montreal, 1984, pp 429-431
- [11] F. ADE. *Characterization of textures by "Eigenfilters"*. Signal Processing 5, 1983, p 451-457
- [12] T. FOUCART. *Analyses factorielles : programmation sur micro-ordinateur*. Masson Editeur, Paris, 1981
- [13] J. POSTAIRE. *De l'image à la décision*. Dunod informatique, Paris, 1987
- [14] J. BASS. *Éléments de calcul de probabilités*. Masson & Cie, Paris, 1974
- [15] P. PROVENT. *Segmentation d'images par analyse statistique de textures : application à la segmentation des images d'échographie cardiaque*. Thèse de doctorat de l'Université Paris XII-Val de Marne, n° 1974, Septembre 1991
- [16] K. FUKUNAGA. *Introduction to statistical Pattern Recognition*. Academic Press, New York, 1972

## LES AUTEURS

Pierre PROVENT

Docteur de l'Université Paris XII-Val de Marne en 1991, il rejoint en 1992 le Conservatoire National des Arts et Métiers qu'il intègre en tant que Maître de Conférences. Ses domaines d'intérêt au sein du L.E.R.I.S.S. concernent la segmentation automatique des images par des approches orientées vers la détection de contour/région, ainsi que leurs applications à l'imagerie médicale.

Jacques LEMOINE, ingénieur SUPELEC et Docteur-Ingénieur, est Maître de Conférences à l'Université Paris XII-Val de Marne depuis 1986. Ses activités de recherche au L.E.R.I.S.S. concernent le Traitement d'Images et la reconnaissance de Formes appliquée aux problèmes d'interface Homme/Machine. Il dirige une formation universitaire de troisième cycle dans le domaine du Traitement de Signal et de ses applications aux signaux physiologiques.

Eric PETIT est né à Toulouse en 1957. Il développe depuis une dizaine d'années une activité de recherche dans le domaine de l'analyse et la segmentation des images. Récemment nommé Professeur à l'Université Paris XII-Val de Marne. Il anime depuis 1987 la division de Traitement d'Images du L.E.R.I.S.S. et dirige l'option Bio-Ingénierie de l'image du D.E.A. de Génie Biologique et Médical d'Ile de France.

En raison de la mauvaise visibilité une fois réduite, nous reproduisons ici la figure 15 :

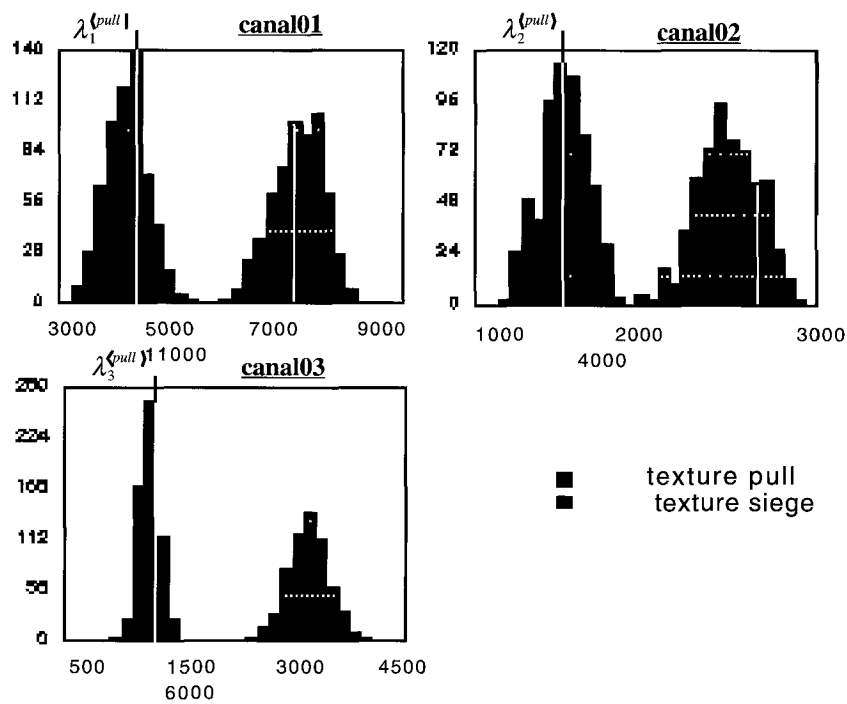


Figure 15 : Attributs des textures "pull" et "siège" filtrés par le banc de filtres  $\Phi^{(pull)}$ .