

**MÉTHODE BASÉE SUR LA TEXTURE POUR  
L'ÉTIQUETAGE DES IMAGES**

par

**Maryse Turcotte**

mémoire présenté au Département de mathématiques et d'informatique  
en vue de l'obtention du grade de maître ès sciences (M.Sc.)

**FACULTÉ DES SCIENCES  
UNIVERSITÉ DE SHERBROOKE**

Sherbrooke, Québec, Canada, février 2000



**National Library  
of Canada**

**Acquisitions and  
Bibliographic Services**

**395 Wellington Street  
Ottawa ON K1A 0N4  
Canada**

**Bibliothèque nationale  
du Canada**

**Acquisitions et  
services bibliographiques**

**395, rue Wellington  
Ottawa ON K1A 0N4  
Canada**

*Your file Votre référence*

*Our file Notre référence*

**The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.**

**The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.**

**L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.**

**L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.**

0-612-67334-0

**Canada**

Le 09-03-2000 , le jury suivant a accepté ce mémoire dans sa version finale.  
date

**Président-rapporteur:** M. Djemel Ziou  
Département de mathématiques et d'informatique

**Membre:** M. François Dubeau  
Département de mathématiques et d'informatique

**Membre:** M. Jean Vaillancourt  
Département de mathématiques et d'informatique

**Membre:** M. Shengrui Wang  
Département de mathématiques et d'informatique

# SOMMAIRE

Dans ce travail, nous expérimentons 25 méthodes d'extraction de caractéristiques de texture et une méthode de sélection non-paramétrique séquentielle vers l'avant basée sur l'algorithme des  $K$  plus proches voisins sur des images à niveaux de gris de texture de Brodatz. Nous généralisons ensuite la méthode aux images couleur de scènes extérieures. Pour ce faire, nous développons une méthode simple mais efficace de segmentation des images couleur que nous appelons "*split-cluster-and-merge*", puis nous adaptons les méthodes d'extraction de caractéristiques de texture aux images couleur en effectuant un changement d'espace couleur. Nous affinons aussi la façon de sélectionner les caractéristiques en effectuant non seulement une sélection globale, mais aussi une sélection selon les classes. Nous validons ces résultats de la sélection à l'aide de trois types de classificateurs: l'algorithme des  $K$  plus proches voisins, le réseau de neurones RBF et le modèle de mixture de réseaux de neurones RBF.



# REMERCIEMENTS

Je tiens à remercier, en premier lieu, mon directeur de recherche, Monsieur Shengrui Wang, pour ses conseils pertinents, sa disponibilité et ses encouragements qui ont, tout au long de ma maîtrise, été fort appréciés.

Je tiens également à remercier mon codirecteur, Monsieur Jean Vaillancourt, pour sa précieuse collaboration.

Mes remerciements vont aussi à tous ceux qui ont contribué à faire de mes études supérieures une expérience agréable et enrichissante, plus particulièrement, merci à Jessica Lévesque et à Joey Champion.

Je remercie, finalement, le CRSNG pour son appui financier.

# TABLE DES MATIÈRES

<b>SOMMAIRE</b>	<b>ii</b>
<b>REMERCIEMENTS</b>	<b>iii</b>
<b>TABLE DES MATIÈRES</b>	<b>iv</b>
<b>LISTE DES TABLEAUX</b>	<b>v</b>
<b>LISTE DES FIGURES</b>	<b>vi</b>
<b>INTRODUCTION</b>	<b>1</b>
<b>CHAPITRE 1 — ÉTIQUETAGE DES IMAGES À NIVEAUX DE GRIS DE TEXTURE DE BRODATZ</b>	<b>5</b>
1.1 Découpage des images . . . . .	6
1.2 Extraction de caractéristiques de texture . . . . .	8
1.2.1 Statistiques de premier ordre sur les niveaux de gris . . . . .	10

1.2.2	Statistiques de second ordre sur la matrice de cooccurrence des niveaux de gris . . . . .	13
1.2.3	Statistiques sur les suites de niveaux de gris . . . . .	18
1.2.4	Statistiques sur l’histogramme des différences (en valeurs absolues) des niveaux de gris . . . . .	21
1.2.5	Modèle autorégressif . . . . .	22
1.3	Sélection des caractéristiques . . . . .	25
1.3.1	Motivation . . . . .	25
1.3.2	Méthode directe de sélection non-paramétrique des caractéristiques	29
1.4	Méthode de classification pour l’étiquetage des imagerie . . . . .	30
1.4.1	K plus proches voisins comme classificateur . . . . .	30
1.4.2	Algorithme des K plus proches voisins . . . . .	32
1.5	Conclusion . . . . .	33

**CHAPITRE 2 — ÉTIQUETAGE DES IMAGES COULEUR DE SCÈNES EXTÉRIEURES** **35**

2.1	Représentation dans l’espace couleur CIE $L^*u^*v^*$ . . . . .	37
2.1.1	Espace couleur CIE $L^*u^*v^*$ . . . . .	38
2.1.2	Transformation de l’espace RGB vers l’espace CIE $L^*u^*v^*$ . . . . .	39
2.2	Segmentation des images . . . . .	41
2.2.1	Étape de séparation ( <i>split</i> ) . . . . .	42

2.2.2	Étape de regroupement en agrégats ( <i>cluster</i> ) . . . . .	42
2.2.3	Étape de fusion des agrégats ( <i>merge</i> ) . . . . .	43
2.3	Extraction de caractéristiques de texture . . . . .	44
2.4	Sélection des caractéristiques . . . . .	45
2.5	Méthode de classification pour l'étiquetage des imagettes . . . . .	47
2.5.1	Réseau de neurones RBF (" <i>Radial Basis Functions</i> ") . . . . .	47
2.5.2	Mixture de réseaux de neurones RBF . . . . .	53
2.6	Conclusion . . . . .	54
 <b>CHAPITRE 3 — INTERFACE GRAPHIQUE</b>		<b>55</b>
3.1	Commandes interactives . . . . .	56
3.1.1	Segmentation . . . . .	56
3.1.2	Étiquetage manuel . . . . .	59
3.1.3	Calcul du vecteur de caractéristiques . . . . .	61
3.2	Commandes automatiques . . . . .	61
3.2.1	Sélection des caractéristiques . . . . .	62
3.2.2	Étiquetage . . . . .	64
3.3	Conclusion . . . . .	64
 <b>CHAPITRE 4 — RÉSULTATS EXPÉRIMENTAUX</b>		<b>65</b>
4.1	Résultats expérimentaux sur les images à niveaux de gris de texture de Brodatz . . . . .	65

4.1.1	Base de données d'images . . . . .	65
4.1.2	Sélection des caractéristiques de texture . . . . .	67
4.1.3	Étiquetage . . . . .	70
4.1.4	Sommaire . . . . .	72
4.2	Résultats expérimentaux sur les images couleur de scènes extérieures . .	74
4.2.1	Base de données d'images . . . . .	74
4.2.2	Sélection des caractéristiques de texture . . . . .	75
4.2.3	Étiquetage . . . . .	78
4.2.4	Sommaire . . . . .	83
	<b>CONCLUSION</b>	<b>86</b>
	<b>Bibliographie</b>	<b>88</b>

# LISTE DES TABLEAUX

1.1	Aperçu général de la méthode d'étiquetage pour les images à niveaux de gris de texture de Brodatz . . . . .	7
2.1	Aperçu général de la méthode d'étiquetage pour les images couleur de scènes extérieures . . . . .	37
2.2	Ensemble des caractéristiques décrites dans la section 1.2 . . . . .	46
4.1	Descriptions des tests effectués sur les images à niveaux de gris de texture de Brodatz . . . . .	67
4.2	Ordre de sélection (colonnes) des caractéristiques (lignes) pour les tests 1, 2, 3 et 4. Le ●, le ○, le * et le ★ indiquent des fréquences de 1, 2, 3 et 4 respectivement . . . . .	69
4.3	Liste des "meilleures" caractéristiques sélectionnées avec $d = 5$ . . . . .	70
4.4	Descriptions des tests effectués sur les images couleur de scènes extérieures	76
4.5	Liste des "meilleures" caractéristiques sélectionnées pour chaque classe et globalement . . . . .	79
4.6	Performances de bonne classification . . . . .	82

# LISTE DES FIGURES

1.1	Exemple d'un découpage d'image en un quadrillage de 64 imasettes. Les imasettes sur la diagonale sont celles qui servent à l'apprentissage. . . . .	7
2.1	Canaux de l'espace couleur CIE $L^*u^*v^*$ . . . . .	38
2.2	Exemple de séparation, regroupement et fusion . . . . .	42
2.3	Architecture de réseau RBF . . . . .	51
3.1	Menu <b>Texture</b> de l'application "Vision" pour la segmentation, l'étiquetage manuel et le calcul des caractéristiques . . . . .	57
3.2	Boîte de dialogue de la commande <b>Segmentation</b> . . . . .	58
3.3	Exemple d'images segmentées . . . . .	60
3.4	Boîte de dialogue de la commande <b>Étiquetage des imasettes</b> . . . . .	61
3.5	Fenêtre principale de l'application "Vision" avec le menu <b>Texture</b> . . . . .	62
3.6	Boîte de dialogue de la commande <b>Sélection des caractéristiques</b> . . . . .	63
3.7	Boîte de dialogue <b>Caractéristiques sélectionnées</b> . . . . .	63
3.8	Boîte de dialogue <b>Matrice de confusion</b> . . . . .	64

4.1	Exemples d'images à niveaux de gris de texture de Brodatz . . . . .	66
4.2	Pourcentages d'erreur de classification en fonction du nombre de "meilleures" caractéristiques sélectionnées pour les tests 1, 2, 3 et 4 . . . . .	68
4.3	Images correspondant aux colonnes 1,2,3,4,5,6 et 10 de la matrice de confusion et pour lesquelles la performance de classification est bonne . . . . .	73
4.4	Images correspondant aux colonnes 7,8 et 9 de la matrice de confusion et pour lesquelles la performance de classification est mauvaise . . . . .	73
4.5	Exemples d'images couleur de scènes extérieures . . . . .	75
4.6	Comparaison des pourcentages d'erreur de classification en fonction du nombre de caractéristiques sélectionnées pour tous les tests. La ligne pleine est pour la sélection globale et la ligne pointillée est pour la sélection selon les classes . . . . .	76
4.7	Un réseau de neurones RBF avec cinq neurones de sortie . . . . .	80
4.8	Cinq réseaux de neurones RBF avec un neurone de sortie pour chacun . . . . .	81
4.9	Cinq réseaux de neurones RBF avec cinq neurones de sortie pour chacun . . . . .	81
4.10	Échantillon des images étiquetées . . . . .	85



# INTRODUCTION

Au cours des dernières années, grâce à l'avancement technologique auquel nous assistons, la quantité d'information à notre disposition ne cesse de croître. On la retrouve sous forme de texte, de son, d'image fixe ou animée ou d'une combinaison de ces différents types de données. Ces informations sont stockées dans des banques de données de façon non structurée et, souvent, ne sont pas utilisées faute de méthodes efficaces d'indexation et de recherche. De ce phénomène est donc né un besoin évident de développer de bons outils qui nous permettront d'accéder à l'information désirée en un temps raisonnable.

Pour cette recherche, nous nous intéressons tout d'abord aux bases de données d'images. Plus spécifiquement, nous nous intéressons à l'indexation des bases de données en utilisant le contenu des images. Notre approche se distingue des autres approches à quelques égards. Entre autres, la plupart des approches existantes utilisent des caractéristiques de bas niveau telles la couleur, la texture, les régions, etc. pour faire l'indexation directement. Quoique ces méthodes sont polyvalentes, elles ne conviennent pas aux applications qui requièrent de l'information sémantique textuelle. Notre approche est complémentaire aux approches utilisant des caractéristiques de bas niveau et est basée sur l'étiquetage des images. L'étiquetage des images revêt une grande importance, car elle permet d'extraire une description textuelle des images, ce qui facilite l'indexation d'une grande base de don-

nées à l'aide de techniques conventionnelles qui ont déjà été développées. Ceci permettra d'améliorer grandement l'efficacité des bases de données d'images dans les applications reliées au multimédia, à l'Internet, à la médecine et à la défense.

L'étiquetage des images est un problème très difficile à cause du nombre infini de scènes, des conditions d'acquisition ainsi que de la grande variété de senseurs. L'étiquetage manuel des images est possible mais faisable seulement si la quantité d'images est petite. Quand le nombre d'images devient très grand, l'étiquetage automatique, sous une forme quelconque, devient indispensable. L'un des premiers systèmes à faire l'étiquetage automatique est décrit dans [22]. Dans ce système, les images sont systématiquement divisées en un quadrillage rectangulaire d'imagettes et quelques-unes des imagettes sont étiquetées manuellement. Ensuite, à l'aide d'un nombre prédéterminé de caractéristiques, des structures d'arbres sont construites et servent à l'étiquetage des imagettes. On retrouve aussi, dans [8], un système d'étiquetage inspiré du précédent. Ce dernier utilise les arbres de distance. Cette nouvelle méthode réduit le temps de calcul pour l'étiquetage. Il s'agit d'un critère non négligeable quand on travaille avec de grandes bases de données d'images. Ces travaux démontrent la faisabilité d'un système d'étiquetage automatique.

Dans ce travail, un nouveau système d'étiquetage d'images couleur de scènes extérieures est proposé. En effet, pour tirer avantage des techniques déjà existantes de manipulation de données alphanumériques, notre objectif est de développer un système automatique qui convertit l'information visuelle contenue dans les images en information textuelle. Tout d'abord, nous avons effectué une étude préliminaire par laquelle nous avons expérimenté, sur des images à niveaux de gris de texture de Brodatz, des méthodes d'extraction de caractéristiques de texture qui nous apparaissaient simples et descriptives [29]. Ensuite, nous avons étudié la contribution potentielle des caractéristiques de texture, i.e. la

caractérisation des imagerie par les caractéristiques de texture. L'attention a été portée sur l'utilisation collective des caractéristiques de texture et non sur leur utilisation individuelle. En effet, une caractéristique individuelle n'est normalement pas suffisante pour caractériser des imagerie associées à plusieurs étiquettes différentes. De même, trop de caractéristiques peuvent aussi avoir un effet négatif sur la caractérisation des imagerie. Pour ces raisons, nous avons utilisé un algorithme de sélection de caractéristiques sous-optimal afin de réduire la dimension du vecteur de caractéristiques de texture tout en maintenant une bonne capacité de discrimination pour l'étiquetage. Les résultats étant satisfaisants, nous avons entrepris de généraliser notre méthode pour les images couleur de scènes extérieures. Nous nous sommes penchés sur deux objectifs spécifiques. Premièrement, nous avons élaboré une méthode de segmentation simple, qui n'est plus arbitraire, mais plutôt basée sur le contenu des images et qui utilise la couleur et la variance seulement. Cette méthode permet de partitionner les images en plusieurs régions significatives telles que des étiquettes peuvent leur être associées. Deuxièmement, nous proposons une méthode simple pour adapter les méthodes d'extraction de caractéristiques de texture sur les images à niveaux de gris aux images multispectrales. Finalement, nous avons étudié l'utilisation des réseaux de neurones et de mixture de réseaux de neurones pour l'apprentissage et la généralisation de l'étiquetage des imagerie.

Avant d'aller plus loin, il est nécessaire de mentionner que le système développé ici n'a pas pour objectif de donner une interprétation détaillée des images. Selon les connaissances actuelles en vision par ordinateur et en intelligence artificielle, il s'agit encore d'un objectif lointain. Le but de notre système est d'étiqueter les régions qui sont grandes et relativement homogènes. Comme il a été dit précédemment, nous utilisons principalement les caractéristiques de texture dans ce travail. Plusieurs autres types de caractéristiques pourront être ajoutées dans ce système dans des travaux futurs. Comme nous le ver-

rons plus loin, le système proposé est très ouvert à d'autres types de caractéristiques. Nous avons choisi la texture en premier lieu, car c'est une caractéristique importante et omniprésente dans plusieurs types d'images. Elle joue aussi un rôle très important dans la perception visuelle chez l'humain [14]. La texture peut être le résultat des propriétés physiques des surfaces et aussi de différentes réflectances. Très souvent, les surfaces naturelles ont un aspect texturé.

Dans les chapitres qui suivent, nous présentons d'abord la méthode d'étiquetage des images à niveaux de gris de texture Brodatz, puis la méthode d'étiquetage des images couleur de scènes extérieures. Un chapitre est consacré à la présentation d'une interface graphique développée pour expérimenter les méthodes proposées pour l'étiquetage d'images couleur. Ensuite, nous présentons les résultats expérimentaux obtenus et effectuons une discussion. Nous terminons avec une conclusion et des perspectives.

# CHAPITRE 1

## ÉTIQUETAGE DES IMAGES À NIVEAUX DE GRIS DE TEXTURE DE BRODATZ

Dans ce chapitre, nous présentons une étude préliminaire effectuée sur des images à niveaux de gris de texture de Brodatz. Les objectifs spécifiques sont d'expérimenter différentes méthodes d'extraction de caractéristiques de texture et une méthode de sélection de caractéristiques pour réduire la dimension du vecteur de caractéristiques tout en maintenant une bonne capacité de discrimination. C'est pour mieux cibler nos objectifs que nous utilisons, pour l'instant, des images homogènes et texturées. En fait, nous posons l'hypothèse que la méthode d'étiquetage proposée dans ce chapitre fonctionne pour toutes les images qui contiennent une seule texture. Notre base de données contient 30 images tirées de "Textures: a Photographic Album for Artists and Designers" par Phil Brodatz (Dover Publications, NY 1966). La figure 4.1 montre un échantillon des images que nous avons utilisées.

Le tableau 1.1 donne un aperçu général de la méthode proposée. La première étape est de découper chaque image en une partition d'éléments de dimensions moindres appelés *imagettes*. Un sous-ensemble de ces imagettes dont nous connaissons *a priori* les classes associées servent pour l'apprentissage et le reste des imagettes servent pour la généralisation. Dans ce chapitre, chaque image correspond à une classe. On associe cette classe aux imagettes qui en sont issues par le découpage. Une fois la partition effectuée, il faut extraire, de chaque imagette, un ensemble de caractéristiques de texture qui leur sont associées et qui servent pour l'apprentissage et la classification. Ensuite, nous appliquons un algorithme de sélection sous-optimal et non-paramétrique qui permet de réduire la dimension du vecteur de caractéristiques extraites à l'étape précédente, tout en demeurant efficace. Cette étape permet, par là même, d'éliminer les caractéristiques inutiles ou redondantes. Finalement, chaque imagette est étiquetée en utilisant un algorithme de classification avec *apprentissage supervisé*, i.e. que les classes associées aux patrons d'apprentissage sont connues à l'avance. La méthode utilisée est la méthode des  $K$  plus proches voisins. Par conséquent, la méthode d'étiquetage décrite dans ce travail est dite *semi-automatique*.

## 1.1 Découpage des images

La première étape est de subdiviser les images en partitions d'éléments de dimensions moindres appelés imagettes. Une façon simple d'y arriver est de séparer les images avec un quadrillage carré comme à la Figure 1.1. Ce découpage est simple mais a comme désavantage de générer des imagettes qui peuvent contenir plusieurs textures différentes. Pour ce volet, nous utilisons des images de texture homogène. Par conséquent, ce type de découpage est adéquat pour l'instant.

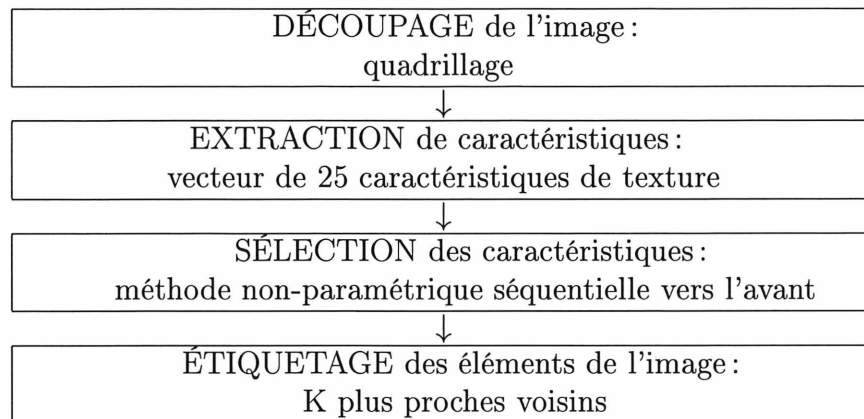


Tableau 1.1 – *Aperçu général de la méthode d'étiquetage pour les images à niveaux de gris de texture de Brodatz*

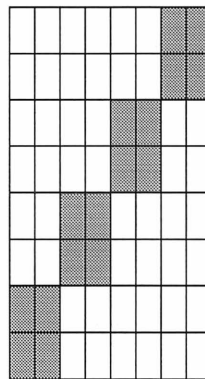


Figure 1.1 – *Exemple d'un découpage d'image en un quadrillage de 64 imagettes. Les imagettes sur la diagonale sont celles qui servent à l'apprentissage.*

De ce découpage, nous dégageons ensuite deux types d'imagettes. Les imagettes faisant partie de la diagonale ou situées près de la diagonale, ainsi que leurs classes associées, forment l'ensemble des patrons d'apprentissage. Les autres imagettes forment l'ensemble des patrons de généralisation, i.e. les imagettes auxquelles nous affecterons une classe. Pour ce travail, la cardinalité de l'ensemble d'apprentissage est telle qu'elle représente le quart du nombre total des imagettes. Ainsi, la diagonale peut avoir une allure telle que celle illustrée à la Figure 1.1.

## 1.2 Extraction de caractéristiques de texture

Comme nous l'avons déjà mentionné, nous nous intéressons, dans ce travail, aux caractéristiques de la texture. La texture est une caractéristique importante et omniprésente dans plusieurs types d'images. Elle joue également un grand rôle dans la perception visuelle chez l'humain [14]. Par conséquent, on essaie d'intégrer la texture dans les systèmes de vision artificielle. Généralement, on reconnaît une texture lorsqu'on la voit. Elle peut découler des propriétés physiques et photométriques de la surface. Les surfaces naturelles ont très souvent un aspect texturé. Toutefois, une définition précise de la texture n'existe pas. On peut cependant la voir comme un ensemble de patrons répétitifs tels que leurs éléments ou primitives sont arrangés suivant des règles de placement déterministes ou stochastiques.

L'analyse de la texture a pour objectif de faire ressortir les caractéristiques la décrivant, e.g. la granularité, le contraste, la régularité, etc. Pour ce faire, nous nous intéressons aux méthodes statistiques basées sur les niveaux de gris, la matrice de cooccurrence, la matrice des suites et l'histogramme des différences puis sur une méthode basée sur le modèle autorégressif. Nous détaillons chacune des caractéristiques extraites de ces méthodes dans la présente section.

### Normalisation des caractéristiques de texture

De ces méthodes, nous extrayons, au total, 25 caractéristiques de texture. Comme nous le verrons dans une section subséquente, la méthode de classification utilisée est la méthode des  $K$  plus proches voisins qui utilise, pour mesurer la proximité, la norme euclidienne. Puisque les différentes caractéristiques ont des ordres de grandeurs variables et qu'alors



leurs contributions à la norme euclidienne peuvent tantôt être trop grandes ou trop petites et ainsi biaiser les résultats, chaque caractéristique est normalisée à l'aide d'un coefficient de normalisation qui lui est propre et tel que la moyenne des normes de la caractéristique, pour l'ensemble des patrons d'apprentissage, est égale à 1, i.e. soit  $\mathbf{c}_{ij}$  le vecteur de la  $i$ -ième caractéristique du  $j$ -ème patron d'apprentissage pour une expérience faite avec  $n_{app}$  patrons d'apprentissage, alors le facteur de normalisation  $F_i$  est tel que :

$$\frac{1}{n_{app}} \sum_{j=1}^{n_{app}} \|F_i \mathbf{c}_{ij}\| = 1$$

où  $\|\mathbf{c}_{ij}\| = \sqrt{\sum_{k=1}^D c_{ijk}^2}$  et  $D$  est la dimension du vecteur de la  $i$ -ième caractéristique.

### Quantification des niveaux de gris

Pour certaines caractéristiques de texture, l'image est préalablement quantifiée pour passer de  $N_g = 256$  à  $N_g = 8$  niveaux de gris. Pour ce faire, nous utilisons la fonction suivante :

$$f(I(x,y)) = \begin{cases} 0 & \text{si } I(x,y) \in [0,31] \\ 1 & \text{si } I(x,y) \in [32,63] \\ \dots & \\ 7 & \text{si } I(x,y) \in [224,255] \end{cases}$$

où  $I(x,y)$  est le niveau de gris du pixel de l'image  $I$  en position  $(x,y)$  avant quantification et  $f(I(x,y))$  est le nouveau niveau de gris de l'image quantifiée.

Par cette transformation, nous pouvons réduire la dimension des matrices de cooccurrence et de suites ainsi que la dimension des histogrammes des différences, ce qui réduit le temps

de calcul pour les caractéristiques extraites de ces structures. De plus, le compromis sur  $N_g$  est intéressant puisque si  $N_g$  est trop grand, le nombre de paires de pixels  $P_d(i, j)$  de  $P_d$  sont petits et par conséquent, statistiquement moins significatifs. D'un autre côté, si  $N_g$  est trop petit, l'information de texture peut être perdue via la quantification.

### 1.2.1 Statistiques de premier ordre sur les niveaux de gris

Les méthodes d'extraction de caractéristiques présentées dans cette section sont des méthodes statistiques.

Les statistiques de premier ordre permettent de mesurer la vraisemblance d'observer une valeur de niveau de gris donnée à une position choisie aléatoirement dans l'image. Ces statistiques sont calculées à partir de l'histogramme de niveaux de gris de l'image et ne dépendent que des pixels pris individuellement et non de leur voisinage. Ainsi, ces statistiques n'apportent pas d'information relative à la position des pixels les uns par rapport aux autres [6]. L'histogramme des niveaux de gris d'une image se définit comme suit: soit  $z$  une variable aléatoire du niveau de gris d'une image  $I(x, y)$ ,  $x = 0, \dots, L_x - 1$  et  $y = 0, \dots, L_y - 1$ .  $L_x$  et  $L_y$  étant le nombre de lignes et de colonnes de l'image  $I$  respectivement. Alors  $h(z_i) = \#\{I(x, y) = i\}$ ,  $i = 0, 1, \dots, N_g - 1$ , est l'histogramme correspondant.  $\#$  est l'opérateur de cardinalité et  $N_g$  le nombre de niveaux de gris dans l'image.

#### Moyenne

La moyenne [26] se définit comme suit :

$$\mu_I = \frac{1}{N} \sum_{i=0}^{N_g-1} z_i h(z_i)$$

où

$$N = \sum_{i=0}^{N_g-1} h(z_i).$$

La moyenne est une mesure globale du degré de "clarté/obscurité" de l'image  $I$ . Elle tient compte des valeurs aberrantes.

### **Médiane**

La médiane [26] est la valeur de niveau de gris  $m_I$  telle qu'approximativement la moitié des pixels de l'image  $I$  ont une valeur de niveau de gris inférieure ou égale à  $m_I$ .

La médiane est aussi une mesure globale du degré de "clarté/obscurité" de l'image  $I$ . Elle tend à soustraire l'effet des valeurs aberrantes.

### **Mode**

Tout comme la moyenne et la médiane, le mode est une mesure de tendance centrale. Le mode est la valeur de niveau de gris  $mo_I$  telle que  $h(mo_I) \geq h(z_i)$ , pour tout  $i = 0, 1, \dots, N_g - 1$ . Autrement dit, il s'agit du niveau de gris avec la fréquence maximale. Si plus d'un niveau de gris ont cette fréquence maximale, il y a alors plusieurs modes. Il existe toujours au moins un mode, contrairement à ce qui est affirmé en [17], puisque  $\{h(z_i)\}$  est fini. Il s'agit d'un mode global.

## Variance et écart-type

La variance [26] se définit comme suit :

$$\sigma_I^2 = \frac{1}{N} \sum_{i=0}^{N_g-1} (z_i - \mu_I)^2 h(z_i)$$

La variance et l'écart-type ( $\sigma_I$ ) sont des mesures globales du contraste de l'image  $I$ . Par exemple, ils sont petits si les niveaux de gris sont proches de la moyenne et grands si les niveaux de gris sont très éparpillés.

## Écart interquantile

Soient  $m_{1I}$  le niveau de gris tel qu'approximativement 1/4 des pixels de l'image  $I$  ont une valeur de niveau de gris inférieure ou égale à  $m_{1I}$  et  $m_{3I}$  le niveau de gris tel qu'approximativement 3/4 des pixels de l'image  $I$  ont une valeur de niveau de gris inférieure ou égale à  $m_{3I}$ . Alors l'écart interquantile [26] est :

$$q_I = |m_{1I} - m_{3I}|.$$

C'est une mesure globale du contraste de l'image  $I$ .

## Moments centrés de l'histogramme de niveaux de gris

On s'intéresse aux troisième et quatrième moments centrés autour de la moyenne de l'histogramme des niveaux de gris [11]. Avant de calculer les moments, les niveaux de gris sont quantifiés, i.e. qu'on passe de  $N_g = 256$  à  $N_g = 8$  niveaux de gris. Cette transformation semble augmenter la performance de ces deux caractéristiques de texture. En effet,

si les niveaux de gris sont convenablement quantifiés, on obtient un histogramme plus significatif.

Le troisième moment centré, noté  $\mu_3$ , est une mesure de l'asymétrie de l'histogramme par rapport à la moyenne. Il se définit comme suit :

$$\mu_3 = \frac{1}{N} \sum_{i=0}^{N_g-1} (z_i - \mu_I)^3 h(z_i).$$

Le quatrième moment centré, noté  $\mu_4$ , est une mesure de l'aplatissement de l'histogramme. Il réfère à la forme de la queue de la distribution, i.e. qu'elle indique si une large proportion des niveaux de gris se trouvent aux extrémités de l'histogramme ou encore s'ils sont tous regroupés autour de la moyenne. Il se définit comme suit :

$$\mu_4 = \frac{1}{N} \sum_{i=0}^{N_g-1} (z_i - \mu_I)^4 h(z_i).$$

### **1.2.2 Statistiques de second ordre sur la matrice de cooccurrence des niveaux de gris**

Les méthodes d'extraction de caractéristiques présentées dans cette section sont aussi des méthodes statistiques.

Les statistiques de second ordre permettent de mesurer la vraisemblance d'observer une paire de niveaux de gris séparés par une distance et une orientation données,  $d$ , un vecteur de déplacement, en une position aléatoire dans l'image [6]. Elles sont calculées à partir

de la matrice de cooccurrence des niveaux de gris de l'image. La matrice de cooccurrence des niveaux de gris (GLCM) d'une image est très connue et très utilisée pour extraire les caractéristiques de texture [11] [16] [15] [31]. Elle fait ressortir certaines propriétés sur la distribution spatiale des niveaux de gris. Par exemple, si les éléments de la matrice sont concentrés autour de la diagonale, la texture a tendance à être grossière par rapport à la distance  $d$ .

La matrice de cooccurrence de niveaux de gris  $P_d$ , de dimension  $N_g \times N_g$ ,  $N_g$  le nombre de niveaux de gris dans l'image et  $d = (dx, dy)$  un vecteur de déplacement, se définit comme suit :

$$P_d(i, j) = \frac{\#\{((r, s), (t, v)) : I(r, s) = i, I(t, v) = j\}}{N_d}$$

avec  $(r, s) \in I$  et  $(t, v) = (r + dx, s + dy) \in I$ . Aussi,  $N_d = \#\{((r, s), (t, v)) : (t, v) = (r + dx, s + dy)\}$ .

Par conséquent,  $P_d(i, j)$  représente la probabilité conjointe d'avoir les niveaux de gris  $(i, j)$  avec une distance relative  $d$ .  $d$  a donc une influence sur les caractéristiques extraites de la GLCM. Il n'existe pas de méthode systématique pour le sélectionner. Dans ce travail, nous fixons  $d = (1, 0)$ . Aussi, afin de réduire la taille de la GLCM, les niveaux de gris sont quantifiés, i.e. qu'on passe de  $N_g = 256$  à  $N_g = 8$  niveaux de gris.

On remarquera qu'un grand nombre de caractéristiques de texture peuvent être calculées à partir de la GLCM. Ces caractéristiques peuvent être corrélées ou complémentaires. C'est une bonne raison pour considérer l'utilisation d'une méthode de sélection des caractéristiques.

téristiques. Il est à noter que les méthodes basées sur la GLCM peuvent mal fonctionner pour les textures avec des primitives de grande taille, i.e. les macrotextures.

### **Énergie ou uniformité ou second moment angulaire**

La caractéristique d'énergie [6] se définit comme suit :

$$ENER = \sum_i \sum_j P_d^2(i,j)$$

Sa valeur est minimale lorsque tous les  $P_d(i,j)$  sont égaux et tend à être grande lorsque les éléments de  $P_d$  sont concentrés, par exemple, autour de la diagonale. Il s'agit d'une mesure d'homogénéité des niveaux de gris de l'image. Sa valeur tend aussi à diminuer lorsque le nombre de niveaux de gris augmente.

### **Entropie**

La caractéristique d'entropie [6] se définit comme suit :

$$ENTR = - \sum_i \sum_j P_d(i,j) \log P_d(i,j)$$

L'entropie indique si les niveaux de gris sont distribués de façon aléatoire dans l'image. Sa valeur est maximale lorsque tous les  $P_d(i,j)$  sont égaux.

En théorie de l'information, l'entropie est une mesure d'incertitude et l'entropie maximale correspond à l'incertitude maximale. Sa valeur tend à augmenter avec la complexité de l'image et avec le nombre de niveaux de gris.

### **Différence de moment d'ordre $k$ ou contraste ou inertie**

La caractéristique de différence de moment d'ordre  $k$  [11] se définit comme suit :

$$DM(k,l) = \sum_i \sum_j (i - j)^k P_d(i,j)^l$$

La différence de moment d'ordre  $k$  tend à être faible lorsque la concentration des éléments autour de la diagonale est grande. Lorsque  $k = 2$ , la caractéristique est souvent appelée contraste et représente la quantité de variations locales présentes dans l'image. Dans ce travail, nous fixons  $k = 2$  et  $l = 1$  [6].

### **Différence de moment inverse d'ordre $k$ ou homogénéité locale**

La caractéristique de différence de moment inverse d'ordre  $k$  [11] se définit comme suit :

$$DMI(k,l) = \sum_i \sum_j \frac{P_d(i,j)^l}{1 + (i - j)^k}$$

La différence de moment inverse d'ordre  $k$  tend à être grande lorsque la concentration des éléments autour de la diagonale est grande. Dans ce travail, nous fixons  $k = 2$  et  $l = 1$  [26].

### **Homogénéité**

La caractéristique d'homogénéité [6] se définit comme suit :

$$HOMO = \sum_i \sum_j \frac{P_d(i,j)}{1 + |i - j|}$$



## Corrélation

La caractéristique de corrélation [6] se définit comme suit :

$$CORR = \sum_i \sum_j \frac{(ij)P_d(i,j) - \mu_x\mu_y}{\sigma_x\sigma_y}$$

où  $\mu_x, \mu_y$  sont les moyennes et  $\sigma_x, \sigma_y$ , les écart-types de  $P_d^x(\cdot)$  et  $P_d^y(\cdot)$  ( $P_d^x(\cdot) = \sum_j P_d(\cdot, j)$  et  $P_d^y(\cdot) = \sum_i P_d(i, \cdot)$ ). Il s'agit d'une mesure de la dépendance linéaire des niveaux de gris dans l'image.

## Probabilité maximum

La caractéristique de probabilité maximum [11] se définit comme suit :

$$PM = \max_{i,j} P_d(i,j)$$

La probabilité maximum représente la plus forte réponse à un vecteur de déplacement  $d$ .

## Probabilité d'une suite de longueur n pour un niveau de gris i

La caractéristique de probabilité d'une suite de longueur  $n$  [16] pour un niveau de gris  $i$  se définit comme suit :

$$PS(n) = \sum_i \frac{(P_d(i) - P_d(i,i))^2 (P_d(i,i))^{n-1}}{P_d(i)^n}$$

où  $P_d(i) = \sum_j P_d(i,j)$ .

Dans ce travail, nous fixons  $n = 2$ .

### **Tendance de regroupement**

La caractéristique de tendance de regroupement [16] se définit comme suit :

$$TR = \sum_i \sum_j (i + j - \mu_i - \mu_j)^k P_d(i, j)$$

où  $\mu_i = \sum_i i \sum_j P_d(i, j)$  et  $\mu_j = \sum_j j \sum_i P_d(i, j)$ .

Dans ce travail, nous fixons  $k = 2$ .

### **1.2.3 Statistiques sur les suites de niveaux de gris**

Une suite de niveaux de gris [13] est un ensemble connexe de pixels colinéaires ayant tous le même niveau de gris. Une suite est caractérisée par son niveau de gris, sa longueur et sa direction. La matrice des suites  $P$ , de dimension  $N_g \times N_r$ , où  $N_g$  est le nombre total de niveaux de gris et  $N_r$  le nombre de longueurs de suites dans l'image, est telle que  $P(i, j)$  est le nombre de fois qu'une suite de longueur  $j$ , de niveau de gris  $i$  ou appartenant au  $i$ -ème intervalle de niveau de gris et de direction  $d$ , apparaît dans l'image.

Afin de réduire la taille de  $P$ , les niveaux de gris sont quantifiés et on passe de  $N_g = 256$  à  $N_g = 8$  niveaux de gris. Aussi, puisque nous travaillons avec des images de dimension  $512 \times 512$ , on a, dans ce travail,  $N_r = 512$ . Le nombre d'opérations nécessaires au calcul de telles matrices est proportionnel au nombre de pixels dans l'image.

Cette méthode peut s'avérer sensible au bruit et elle n'apporte pas d'information sur la distribution spatiale des niveaux de gris entre eux. Les statistiques extraites de la matrice  $P(i,j)$  donnent plutôt de l'information sur la "grossièreté" de la texture. En effet, les textures grossières contiennent de longues suites de pixels à niveau de gris presque constant.

Finalement, on s'attend à ce que les longueurs des suites soient dépendantes de la direction  $d$  lorsque la texture est caractérisée par une direction. Pour ce travail, nous fixons  $d = (1,0)$ .

### **"Short Run Emphasis Moments"**

La caractéristique "Short Run Emphasis Moments" [13] est définie comme suit :

$$\frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \frac{P(i,j)}{j^2}}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} P(i,j)}$$

où le dénominateur est le nombre de suites dans l'image et agit comme facteur de normalisation. Elle a tendance à être grande lorsque l'image contient de courtes suites de pixels, i.e. si la texture est fine.

### **"Long Run Emphasis Moments"**

La caractéristique "Long Run Emphasis Moments" [13] est définie comme suit :

$$\frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} j^2 P(i,j)}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} P(i,j)}$$

où le dénominateur est le nombre de suites dans l'image et agit comme facteur de normalisation. Elle a tendance à être grande lorsque l'image contient de longues suites de pixels, i.e. si la texture est grossière.

### **Non-uniformité des niveaux de gris**

La caractéristique de non-uniformité des niveaux de gris [13] est définie comme suit :

$$\frac{\sum_{i=1}^{N_g} (\sum_{j=1}^{N_r} P(i,j))^2}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} P(i,j)}$$

où le dénominateur est le nombre de suites dans l'image et agit comme facteur de normalisation. Elle dépend de la distribution des niveaux de gris et a tendance à être petite lorsque les niveaux de gris des suites sont uniformément distribués. Ce sont les longues suites qui contribuent le plus à cette statistique.

### **Non-uniformité des suites de niveaux de gris**

La caractéristique de non-uniformité des suites de niveaux de gris [13] est définie comme suit :

$$\frac{\sum_{i=1}^{N_r} (\sum_{j=1}^{N_g} P(i,j))^2}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} P(i,j)}$$

où le dénominateur est le nombre de suites dans l'image et agit comme facteur de normalisation. Elle dépend de la distribution des suites de niveaux de gris et a tendance à être petite lorsque les suites sont uniformément distribuées dans l'image. Ce sont les grandes longueurs de suites qui contribuent le plus à cette statistique.

## "Run percentage"

La caractéristique "Run percentage" [13] est définie comme suit :

$$\frac{\sum_{i=1}^{N_r} \sum_{j=1}^{N_g} P(i,j)}{N^2}$$

où  $N$  est le nombre de pixels dans l'image. Cette statistique est le rapport entre le nombre total de suites et le nombre total de suites possibles si elles étaient toutes de longueur 1. Elle a tendance à être grande lorsque les suites sont toutes petites. Elle a tendance à être petite lorsque l'image a une structure linéaire.

### 1.2.4 Statistiques sur l'histogramme des différences (en valeurs absolues) des niveaux de gris

Soit  $h_d(z_l)$ ,  $l = 0, 1, \dots, N_g - 1$ , l'histogramme des différences (en valeurs absolues) des niveaux de gris, où  $h_d(z_l)$  est la somme des entrées de la matrice de cooccurrence  $P_d(i, j)$  (telle que décrite dans la section 1.2.2) sur les lignes parallèles à la diagonale, i.e. telles que  $|i - j| = z_l$ . Avec cet histogramme, on peut calculer plusieurs caractéristiques analogues à celles de la matrice de cooccurrence des niveaux de gris.

#### Moyenne de différence

La caractéristique moyenne de différence [31] se définit comme suit :

$$MD = \sum_{l=0}^{N_g-1} z_l h_d(z_l).$$

## Second moment

La caractéristique second moment [31] se définit comme suit :

$$SM = \sum_{l=0}^{N_g-1} z_l^2 h_d(z_l).$$

C'est l'analogie de la caractéristique de contraste sur la matrice de cooccurrence des niveaux de gris.

## Entropie de différence

La caractéristique entropie de différence [31] se définit comme suit :

$$ED = \sum_{l=0}^{N_g-1} h_d(z_l) \log h_d(z_l).$$

### 1.2.5 Modèle autorégressif

Le modèle autorégressif [1] utilise des estimés linéaires du niveau de gris d'un pixel, étant donné les niveaux de gris de ses voisins, pour caractériser la texture. Il permet de modéliser les interactions spatiales entre les pixels d'une image. Pour une texture grossière, les coefficients sont tous similaires, alors que, pour une texture fine, les coefficients sont très variables.

Soit  $X(k) = x$  une variable aléatoire d'un processus stochastique où  $k = (i, j)$  est la position du pixel dans l'image et  $x$  son niveau de gris correspondant. Le modèle autorégressif suppose que  $X(k)$  soit stationnaire, i.e.  $X(k)$  et  $X(k + c)$ ,  $c$  une constante, ont

les mêmes statistiques dans leur voisinage ( $\mu_k = \mu$  et  $\sigma_k = \sigma$ , pour tout  $k$ ) et ergodique, i.e. que toutes ses caractéristiques peuvent être déterminées avec un nombre de voisins fini et suffisant. Le modèle autorégressif permet de bien représenter les microtextures et les macrotextures. En général, il est d'ordre 2,3 ou 4, tout dépendant de la nature de la texture. Une analyse plus détaillée sur la représentation de la texture par le modèle autorégressif est fournie dans [1]. Dans ce travail, nous utilisons un modèle autorégressif d'ordre 3.

Nous définissons le modèle autorégressif par l'équation suivante :

$$X_s = a_0 W_s + \sum_{r \in \Omega^+} a_r X_{s+r}$$

où  $X_s$  représente le niveau de gris à la position  $s$ ,  $s$  la position  $(i,j)$  sur les lignes et les colonnes,  $\Omega^+ = \Omega \setminus \{s\}$  le voisinage sur les lignes et les colonnes de  $X_s$  (n'incluant pas  $X_s$ ) et  $\{a_0, a_r, r \in \Omega^+\}$  les paramètres à estimer du modèle.

On définit  $\Omega^+$ , un voisinage causal, par  $\Omega^+ = \{(-1,0), (-1,1), (0,1)\}$ .

$W_s$  est un bruit gaussien tel que :

$$E[W_s] = 0,$$

$$E[W_s W_{s+r}] = \begin{cases} 1 & \text{si } r = (0,0) \\ 0 & \text{sinon,} \end{cases}$$

$$E[W_s X_{s+r}] = \begin{cases} a_0 & \text{si } r = (0,0) \\ 0 & \text{sinon.} \end{cases}$$

Pour l'estimation des paramètres  $\{a_0, a_r, r \in \Omega^+\}$ , nous allons utiliser la méthode des moindres carrés. Comme le bruit est présumé gaussien, cette méthode coïncide ici avec celle du maximum de vraisemblance - ce n'est pas le cas en général.

L'erreur d'estimation (qui correspond au bruit  $W_s$ ), en posant  $b_0 = 1$  et  $b_r = -a_r$ , pour tout  $r \neq 0$ , est donnée par :

$$E_s = \frac{1}{a_0} \sum_{r \in \Omega^+} b_r X_{s+r}.$$

Pour toute l'image, l'erreur quadratique est :

$$E = \sum_{s \in I} \left( \frac{1}{a_0} \sum_{r \in \Omega^+} b_r X_{s+r} \right)^2.$$

La méthode des moindres carrés consiste à déterminer les coefficients  $\{a_0, a_r, r \in \Omega^+\}$  tels que les dérivées partielles de  $E$  s'annulent, i.e.

$$\frac{\partial E}{\partial b_r} = 0, r \in \Omega^+$$

et

$$\frac{\partial E}{\partial a_0} = 0.$$



Les solutions de ce système d'équations sont :

$$\left(\sum_s Z(s)Z^T(s)\right)A = \sum_s Z(s)X_s$$

et

$$a_0^2 = \frac{1}{N^2} \sum_s (X_s - A^T Z(s))^2$$

où  $Z(s)$  est le vecteur  $[X_{s+r}, r \in \Omega^+]$  et  $A$  est le vecteur des paramètres  $[b_r, r \in \Omega^+]$ .

Puisque nous désirons connaître les valeurs des paramètres  $\{a_r, r \in \Omega^+\}$ , il suffit de résoudre le système linéaire  $(\sum_s Z(s)Z^T(s))A = \sum_s Z(s)X_s$ .

On peut estimer le coefficient du bruit  $a_0$  grâce à la deuxième relation, i.e.

$$a_0 = \frac{1}{N} \sqrt{\sum_s (X_s - A^T Z(s))^2}.$$

## 1.3 Sélection des caractéristiques

### 1.3.1 Motivation

Maintenant que nous sommes en mesure d'obtenir un nombre intéressant de caractéristiques, nous nous intéressons, dans cette section, à la dimension de cet ensemble de

caractéristiques.

Dans les applications, il est souvent nécessaire de réduire la dimension de l'espace des caractéristiques en sélectionnant les "meilleures" caractéristiques afin que le problème soit plus facile à manipuler et que le modèle soit moins complexe tout en maintenant un niveau de performance satisfaisant. Les caractéristiques extraites doivent alors contenir l'information discriminante nécessaire pour résoudre le problème et ne doivent pas être redondantes.

En fait, il a été démontré [5], théoriquement et expérimentalement, que la performance de classification a tendance à se détériorer avec l'augmentation de la dimension de l'ensemble des caractéristiques. Souvent, cela est dû au fait que le modèle résultant a une variance élevée ("*over-fitting*") au lieu d'être lisse, ce qui diminue les capacités de généralisation du modèle et, par conséquent, la performance de la classification. Ce phénomène s'appelle la "malédiction des grandes dimensions". Ce phénomène peut aussi s'expliquer par un petit nombre de données d'apprentissage qui résulte en un modèle moins représentatif et plus dépendant des données disponibles.

Supposons que l'on veut sélectionner un sous-ensemble de  $d$  caractéristiques parmi  $D$  au total. Le nombre total de sous-ensembles de dimension  $d$  à considérer est

$$\binom{D}{d} = \frac{D!}{(D-d)!d!}.$$

Vu le grand nombre de sous-ensembles que cela amène à considérer, une approche de sélection *sous-optimale* s'avère intéressante. Les méthodes de sélection les plus connues

[10] sont les algorithmes exponentiels (ex : "*branch and bound*"), les algorithmes stochastiques (ex : méthodes génétiques) et les algorithmes séquentiels (ex : sélection séquentielle vers l'avant, sélection séquentielle vers l'arrière).

La méthode de sélection séquentielle vers l'avant [2] est une méthode sous-optimale. Le sous-ensemble de caractéristiques de départ est vide. La première étape consiste à sélectionner la caractéristique qui, à elle seule, donne la meilleure performance de classification. Les étapes subséquentes consistent à sélectionner, une à une, les nouvelles caractéristiques qui, combinées aux caractéristiques déjà sélectionnées, donnent les meilleures performances de classification et ce, jusqu'à l'obtention du vecteur de caractéristiques de dimension  $d$ . Cette méthode donne une solution sous-optimale car elle n'examine pas tous les sous-ensembles de caractéristiques possibles. Cependant, elle présente l'avantage que le nombre de sous-ensembles de caractéristiques à considérer est maintenant réduit à

$$\sum_{i=1}^d (D - i + 1) = \frac{d(2D - (d - 1))}{2}.$$

De plus, lors des premières étapes, la procédure d'évaluation devrait avoir une bonne précision puisque le ratio entre le nombre de patrons d'apprentissage et la dimension du sous-ensemble de caractéristiques à évaluer est élevé. Aussi, puisque la méthode sélectionne la caractéristique qui performe la mieux avec les caractéristiques déjà sélectionnées, elle tient compte des dépendances entre les caractéristiques pour la classification. Finalement, soit  $E$  l'ensemble des caractéristiques déjà sélectionnées, alors tous les sous-ensembles de caractéristiques  $E \cup \{c_i\}, c_i \notin E$  sont évalués.

Il existe aussi diverses façons d'évaluer l'habileté de discrimination des sous-ensembles

de caractéristiques [9] lorsque nous utilisons un algorithme de sélection séquentielle. Les fonctions d'évaluation les plus connues sont la mesure de distance ou séparabilité ou discrimination, la mesure de l'information, la mesure de dépendance ou corrélation, la mesure de la probabilité d'erreur de classification, etc.

Pour ce travail, le premier critère d'évaluation est la mesure de la probabilité d'erreur de classification. La fonction d'évaluation est issue du classificateur lui-même. Il a été conjecturé que ce type de méthodes donne de meilleurs résultats [18]. Cependant, cette méthode est plutôt coûteuse en temps de calcul. De plus, il est à noter que la probabilité d'erreur de classification comme critère d'évaluation présente les désavantages suivants :

1. Pour que le niveau de confiance sur les résultats soit élevé, il faut beaucoup de données;
2. Les résultats issus de la méthode de sélection sur un ensemble d'images peuvent ne pas être pertinents, i.e. qu'ils peuvent ne pas représenter la capacité de discrimination réelle des caractéristiques de texture;
3. La probabilité d'erreur de classification ne donne pas d'indication à savoir pourquoi une caractéristique performe bien et une autre, non. Elle ne fait que mesurer l'efficacité de la sélection de caractéristiques de texture.

Le deuxième critère d'évaluation est le pourcentage de rejet de classification. En effet, nous ne voudrions pas privilégier une caractéristique avec laquelle plusieurs patrons ne sont pas classifiés.

### 1.3.2 Méthode directe de sélection non-paramétrique des caractéristiques

L'objectif de la présente méthode est de sélectionner le "meilleur" sous-ensemble de  $d$  caractéristiques à partir d'un ensemble de  $D$  caractéristiques. L'algorithme de sélection des caractéristiques est séquentiel vers l'avant et se déroule comme suit :

1. Initialisation de l'ensemble des  $d$  "meilleures" caractéristiques:  $E := \emptyset$
2. Initialisation du pourcentage de bonne classification:  $MinErreur := 100$  et du pourcentage de rejet  $MinRejet := 100$
3. Pour chaque caractéristique  $c_i \notin E, i = 1, \dots, D$ :
  - $\tilde{E} := E \cup \{c_i\}$
  - Classifier toutes les imagettes appartenant à l'ensemble de généralisation à l'aide des caractéristiques de  $\tilde{E}$
  - Calculer le pourcentage d'erreur de classification  $PctErreur_i$  et le pourcentage de rejet  $PctRejet_i$
  - Si ( $PctErreur_i < MinErreur$ )  
alors  $MinErreur := PctErreur_i$ ,  $MinRejet := PctRejet_i$  et  $meilleur := c_i$
  
  - Sinon, si ( $PctErreur_i == MinErreur$ ) ET ( $PctRejet_i < MinRejet$ )  
alors  $MinRejet := PctRejet_i$  et  $meilleur := c_i$
4.  $E := E \cup \{ meilleur \}$

5. Si  $\text{card}(E) == d$ , alors la sélection est terminée, sinon, aller à 2.

Le premier critère d'évaluation est la probabilité d'erreur de classification. Elle est estimée en calculant le rapport entre le nombre d'imagettes mal classifiées (incluant les imagettes non classifiées) et le nombre total d'imagettes à classifier. Le deuxième critère d'évaluation est le pourcentage de rejet de classification. Il correspond au rapport entre le nombre d'imagettes non classifiées et le nombre total d'imagettes à classifier. Les caractéristiques sélectionnées sont celles qui minimisent d'abord la probabilité d'erreur de classification puis ensuite le pourcentage de rejet de classification. En cas d'égalité, on choisit la première caractéristique qui minimise le premier critère, puis ensuite, le second.

## 1.4 Méthode de classification pour l'étiquetage des imagettes

Une fois la sélection des caractéristiques effectuée, nous pouvons procéder à l'étiquetage ou classification des imagettes à l'aide des caractéristiques sélectionnées. La méthode de classification utilisée pour l'étiquetage des imagettes est la méthode des  $K$  plus proches voisins [5]. En fait, cet algorithme est utilisé comme critère d'évaluation pour la sélection des caractéristiques de texture ainsi que pour la validation de la sélection.

### 1.4.1 $K$ plus proches voisins comme classificateur

Il s'agit d'une méthode de classification *non-paramétrique*, i.e. qui ne présuppose pas de distribution statistique particulière. L'idée de cette méthode est de construire un classi-

ficateur en s'inspirant du théorème de Bayes. Il faut alors estimer, pour chaque classe, la densité *a priori* et la densité conditionnelle de classe qui nous permettent d'obtenir la densité *a posteriori* que l'on utilise directement pour bâtir la règle de décision.

Supposons que notre ensemble d'images d'apprentissage contienne  $N$  images au total et que  $N_k$  images appartiennent à la classe  $C_k$ , i.e.  $\sum_k N_k = N$ . Il faut alors construire, autour du vecteur de caractéristiques de l'image à classifier, une hypersphère contenant  $K$  patrons d'apprentissage et ce, indépendamment des classes auxquelles ils appartiennent. Supposons que cette hypersphère, de volume  $V$ , contient  $K_k$  images appartenant à la classe  $C_k$ . Alors les approximations des densités de probabilité conditionnelle de classe, de probabilité non conditionnelle et de probabilité *a priori* sont données par :

$$p(x|C_k) = \frac{K_k}{N_k V},$$

$$p(x) = \frac{K}{NV}$$

et

$$p(C_k) = \frac{N_k}{N}$$

respectivement.

$\frac{K_k}{N_k}$  est la proportion des données de la classe  $C_k$  se trouvant dans l'hypersphère de volume  $V$ . En supposant que la densité des patrons d'apprentissage est relativement uniforme

dans cette région, on obtient une estimation de  $p(x|C_k)$  par  $\frac{K_k}{N_k V}$ . Un raisonnement similaire s'applique pour l'estimation de la probabilité non conditionnelle. Par conséquent, le choix de  $K$  est important puisque  $V$  est le volume de l'hypersphère qui contient  $K$  patrons d'apprentissage. En effet, comme nous le redirons à la section 1.4.2, la valeur de  $K$  ne doit être ni trop grande, ni trop petite pour que les estimations des probabilités ci-haut soient satisfaisantes.

En utilisant le théorème de Bayes, on obtient les densités de probabilité *a posteriori*:

$$p(C_k|x) = \frac{p(x|C_k)p(C_k)}{p(x)} = \frac{K_k}{K}.$$

Afin de minimiser la probabilité d'erreur de classification d'une imagerie à classifier, il faut lui affecter la classe dont le rapport  $\frac{K_k}{K}$  est maximal. Cette règle est connue sous le nom de *règle de décision des K plus proches voisins*.

## 1.4.2 Algorithme des K plus proches voisins

Cette méthode consiste à trouver, autour du vecteur de caractéristiques de l'imagerie à classifier, une hypersphère contenant  $K$  patrons d'apprentissage et ce, indépendamment des classes auxquelles ils appartiennent et d'affecter à l'imagerie à classifier la classe "la plus représentée" parmi les classes correspondant aux patrons d'apprentissage compris à l'intérieur de l'hypersphère.

Notre méthode de classification diffère légèrement de la méthode originale. En effet, dans notre cas, nous définissons qu'une classe est "la plus représentée" si, parmi les  $K$  plus



proches voisins, on la retrouve au minimum  $K/2$  fois. Dans le cas où aucune classe ne se qualifie comme "la plus représentée", il y a rejet et on ne fait aucune classification. Dans l'algorithme original, on affecte un patron à la classe pour laquelle  $K_k$  est le plus grand. Cette légère modification rend donc l'algorithme plus restrictif.

Un désavantage de cette méthode est qu'il faut calculer la distance avec tous les patrons d'apprentissage à chaque fois que l'on veut classifier un vecteur.

Il existe une valeur optimale pour  $K$ , i.e. un nombre de voisins qui permet de bien représenter la classe du patron à classifier. Toutefois, il n'existe pas de méthode précise pour la calculer puisque le choix judicieux de  $K$  requiert la connaissance *a priori* de l'espace des caractéristiques et des classes associées. Si  $K$  est trop grand, il peut y avoir des régions de l'espace telles que les approximations ci-haut sont trop "lissées" et par conséquent, moins représentatives. L'opposé peut également se produire. Si  $K$  est trop petit, il peut y avoir des régions de faible densité où les estimés sont alors très sensibles aux données disponibles. Dans ce travail, il semble raisonnable de fixer  $K = 3$ .

## 1.5 Conclusion

Dans cette section, nous avons présenté une méthode d'étiquetage pour les images de texture de Brodatz. Nous avons d'abord découpé les images à l'aide d'un quadrillage rectangulaire. Les images de texture de Brodatz représentant une seule texture, elles ne nécessitent pas une méthode de segmentation sophistiquée. Ensuite, nous avons extrait, pour chaque image obtenue du découpage, un vecteur de 25 caractéristiques de texture. Afin de diminuer la dimension de l'espace de caractéristiques, nous avons appliqué

un algorithme de sélection séquentielle vers l'avant. Dans le chapitre 4, nous présentons quelques résultats expérimentaux. Ceux-ci démontrent que la sélection permet de contrer la "malédiction des grandes dimensions". Aussi, on peut déduire, par la consistance des résultats obtenus, que certaines caractéristiques ont une capacité discriminative supérieure. Finalement, nous effectuons une classification ou étiquetage des imagettes. Les résultats expérimentaux du chapitre 4 montrent les pourcentages de bonne classification. Ces résultats sont très satisfaisants et nous ont encouragés à tenter de généraliser, ce que l'on décrit dans le chapitre qui suit, notre méthode aux images couleur de scènes extérieures.

## CHAPITRE 2

# ÉTIQUETAGE DES IMAGES COULEUR DE SCÈNES EXTÉRIEURES

Le travail décrit dans ce chapitre est une extension à la méthode proposée dans le chapitre précédent aux images couleur de scènes extérieures. Cette modification amène deux nouveaux problèmes. Le premier est la segmentation des images en régions distinctes de textures homogènes. En effet, il est préférable, pour ce travail, de considérer des imagerie qui contiennent, le plus possible, une seule texture. On retrouve, dans la littérature, une quantité importante de méthodes pour la segmentation des images. Dans ce travail, nous nous inspirons de la méthode de *"split-and-merge"* [11] pour développer une méthode que nous appelons *"split-cluster-and-merge"*. Le second est la modification des méthodes d'extraction de caractéristiques de texture. Avant, les caractéristiques de texture étaient extraites d'imagerie monochromes de formes et de dimensions fixes. Avec les images de scènes extérieures, nous obtenons des imagerie couleur de formes et de dimensions

variables. Par conséquent, il faut voir comment adapter les méthodes d'extraction de caractéristiques aux images multispectrales. Il est à noter qu'à cause des formes et dimensions variées des imagettes issues de la segmentation, certaines méthodes d'extraction de caractéristiques de texture perdent de leur invariabilité. Cependant, sur ce point, aucune mesure correctrice n'est apportée. Il sera intéressant de voir comment ce problème se reflètera dans les résultats issus de l'algorithme de sélection des caractéristiques.

Le tableau 2.1 donne un aperçu général de la méthode proposée. Initialement, nous convertissons les images à un espace couleur tel que la distance euclidienne représente la similarité dans la perception de la couleur. Une fois cette étape effectuée, nous pouvons segmenter chaque image en une partition d'images de dimensions moindres et aussi homogènes que possible en texture appelées *imagettes*. Un sous-ensemble de ces imagettes dont nous connaissons *a priori* les classes associées servent pour l'apprentissage et un autre sous-ensemble d'imagettes servent pour la généralisation. Dans ce chapitre, nous avons choisi d'associer aux imagettes issues du découpage les 5 classes suivantes: "Ciel", "Eau", "Arbre", "Mur" et "Herbe". Une fois la segmentation effectuée, il faut extraire, de chaque imagette, un ensemble de caractéristiques de texture qui leur sont associées et qui servent pour l'apprentissage et la classification. Ensuite, nous appliquons un algorithme de sélection sous-optimal et non-paramétrique qui permet de réduire la dimension du vecteur de caractéristiques extraites à l'étape précédente, tout en demeurant efficace. Cette étape permet, par là même, d'éliminer les caractéristiques inutiles ou redondantes. La sélection est effectuée d'abord globalement et, par la suite, selon la classe. Finalement, chaque imagette est étiquetée en utilisant différents algorithmes de classification avec *apprentissage supervisé*, i.e. que les classes associées aux patrons d'apprentissage sont connues à l'avance. Les méthodes de classification utilisées sont les  $K$  plus proches voisins, un réseau de neurones RBF et deux modèles de mixture de réseaux

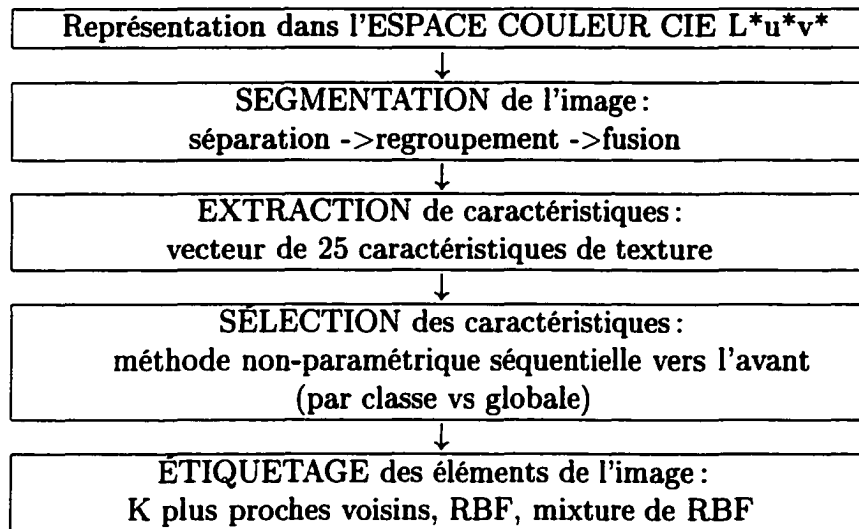


Tableau 2.1 – Aperçu général de la méthode d'étiquetage pour les images couleur de scènes extérieures

de neurones RBF. Par conséquent, la méthode d'étiquetage décrite dans ce travail est dite *semi-automatique*.

## 2.1 Représentation dans l'espace couleur CIE L\*u\*v

Lorsque l'on veut se servir de la couleur comme critère de similarité entre deux régions, deux problèmes principaux se posent :

- le choix de l'espace couleur;
- le choix de la métrique dans l'espace couleur.

Dans l'espace couleur RGB, il n'existe pas de relation d'ordre définie sur les couleurs. En fait, la nature intrinsèquement non euclidienne de l'espace des couleurs induit la variabilité de la valeur de l'élément métrique d'un point à l'autre de cet espace. Alors qu'il est possible de paramétrer une couleur et donc, théoriquement, de l'identifier de manière

absolue, il est impossible d'obtenir, à partir de cet espace d'origine, un système parfaitement uniforme [28].

### 2.1.1 Espace couleur CIE $L^*u^*v^*$

Pour ce travail, nous effectuons une transformation vers l'espace couleur CIE (Commission Internationale de l'Éclairage)  $L^*u^*v^*$  qui a été développé exprès pour que la distance euclidienne quantifie la similarité entre deux couleurs, i.e. que deux couleurs proches dans cet espace sont également perceptuellement similaires. Il s'agit de l'un des meilleurs modèles de la couleur. Il est dit *uniforme* [20]. La Figure 2.1 montre les trois canaux de l'espace couleur CIE  $L^*u^*v^*$ .

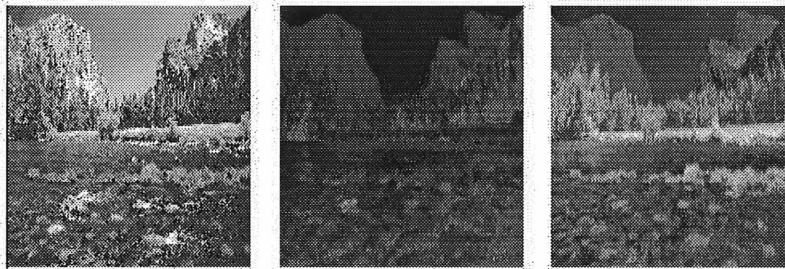


Figure 2.1 – *Canaux de l'espace couleur CIE  $L^*u^*v^*$*

Le canal  $L^*$  représente la clarté psychométrique ou luminance. Sa valeur servira pour l'extraction des caractéristiques de texture. Ceci nous évitera d'avoir à adapter les méthodes d'extraction pour des images multispectrales. Les valeurs de  $u^*$  et  $v^*$  sont reliées à la teinte et à la saturation comme suit :

$$\text{teinte} = \arctan\left(\frac{v^*}{u^*}\right)$$

et

$$\text{saturation} = \sqrt{u^{*2} - v^{*2}}.$$

### 2.1.2 Transformation de l'espace RGB vers l'espace CIE L\*u\*v\*

L'espace couleur CIE L\*u\*v\* est défini à l'aide de l'espace couleur 1931 CIE XYZ formé de trois couleurs primaires définies artificiellement pour décrire toutes les couleurs visibles. Les valeurs X, Y et Z sont le résultat d'associations de couleurs par un observateur standard soumis à des conditions d'illumination et de visionnement données [25].

Les équations qui suivent permettent de calculer les valeurs de L\*, u\* et v\* en fonction des trois canaux R, G et B [28].

**Calcul de L\* :**

$$L^* = \begin{cases} 116(Y/Y_B)^{1/3} - 16 & \text{si } (Y/Y_B) > 0,008856 \\ 903.3(Y/Y_B) & \text{sinon} \end{cases}$$

**Calcul de u\* :**

$$u^* = 13L^*(u' - u'_B)$$

avec

$$u' = \begin{cases} \frac{4X}{X+15Y+3Z} & \text{si } X \neq 0 \text{ ou } Y \neq 0 \text{ ou } Z \neq 0 \\ 0 & \text{sinon} \end{cases}$$

et

$$u'_B = \frac{4X_B}{X_B + 15Y_B + 3Z_B}$$

**Calcul de  $v^*$  :**

$$v^* = 13L^*(v' - v'_B)$$

avec

$$v' = \begin{cases} \frac{9Y}{X+15Y+3Z} & \text{si } X \neq 0 \text{ ou } Y \neq 0 \text{ ou } Z \neq 0 \\ 0 & \text{sinon} \end{cases}$$

et

$$v'_B = \frac{9Y_B}{X_B + 15Y_B + 3Z_B}$$

$X, X_B, Y, Y_B, Z$  et  $Z_B$  se calculent de la façon suivante [23] :

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} 0.412453 & 0.357580 & 0.180423 \\ 0.212671 & 0.715160 & 0.072169 \\ 0.019334 & 0.119193 & 0.950227 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix}$$



et

$$\begin{pmatrix} X_B \\ Y_B \\ Z_B \end{pmatrix} = \begin{pmatrix} 0.412453 & 0.357580 & 0.180423 \\ 0.212671 & 0.715160 & 0.072169 \\ 0.019334 & 0.119193 & 0.950227 \end{pmatrix} \begin{pmatrix} 255 \\ 255 \\ 255 \end{pmatrix}$$

## 2.2 Segmentation des images

Dans cette partie du travail, nous tentons de développer une technique de segmentation simple qui permettra d'obtenir une partition d'éléments de dimensions moindres et de textures homogènes appelés imagettes. Pour les images de scènes extérieures, il est parfois difficile de faire une bonne segmentation car les régions adjacentes de textures différentes ne sont pas toujours séparées par une frontière bien définie. De plus, les différences d'échelle pour de mêmes textures rendent la tâche plus difficile.

La méthode de segmentation est inspirée de la méthode de "*split-and-merge*". Nous l'appelons "*split-cluster-and-merge*". C'est une méthode de segmentation non-supervisée. La méthode proposée se subdivise en trois étapes :

- séparation systématique en un quadrillage régulier de dimension fixe (*split*);
- regroupement en agrégats selon un critère de regroupement (*cluster*);
- fusion des agrégats selon un critère de proximité (*merge*).

Il est à noter que la performance de la segmentation dépend de la qualité des caractéristiques dans l'image, i.e. du bruit à l'intérieur des régions homogènes et du contraste entre elles.

### 2.2.1 Étape de séparation (*split*)

Comme mentionné ci-haut, l'étape de séparation est systématique. On fixe les dimensions d'un quadrillage dont sont issues les imagettes qui deviennent les unités de base pour le résultat final de la segmentation. La dimension d'un rectangle doit être suffisamment petite afin d'être à la fois efficace à l'intérieur des petites régions et de fournir un faible taux d'erreur près des frontières de texture. La Figure 2.2a. donne un exemple de l'étape de séparation avec un quadrillage en imagettes  $8 \times 8$  pixels.

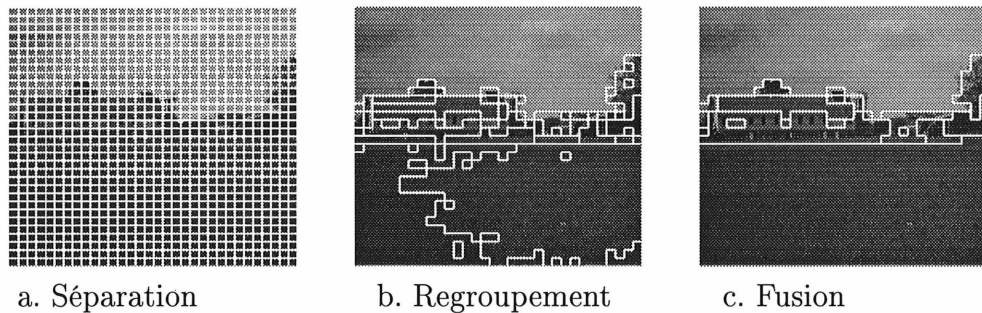


Figure 2.2 – Exemple de séparation, regroupement et fusion

### 2.2.2 Étape de regroupement en agrégats (*cluster*)

L'étape de regroupement est réalisée à l'aide de la version *batch* de l'algorithme de K-moyenne. Pour cet algorithme, il faut d'abord fixer  $K$ , le nombre d'agrégats. Il existe une valeur optimale pour  $K$ . Toutefois, on ne connaît pas de méthode précise pour la calculer. Des méthodes d'estimation de  $K$  ont toutefois été développées [3] [4] [24] [27] [32]. Pour ce travail, nous fixons  $K$  manuellement.

En guise d'initialisation des agrégats, on affecte, de façon systématique, les imagettes obtenues à l'étape précédente aux  $K$  agrégats. Soit  $N$ , le nombre d'imagettes obtenues,

on numérote, pour ce travail, les imasettes de gauche à droite et de haut en bas et on affecte l'agrégat  $i$  modulo  $K, i = 1, 2, \dots, N$  à la  $i$ -ième imasette.

On calcule ensuite, pour chaque imasette et, globalement, pour chaque agrégat, les caractéristiques qui serviront comme critères de regroupement. Pour ce travail, nous avons choisi, les valeurs CIE  $L^*u^*v^*$  moyennes, i.e.  $[\overline{L^*} \ \overline{u^*} \ \overline{v^*}]^T$  et la variance sur la bande  $L^*$ , i.e. sur la clarté psychométrique. Ces caractéristiques sont simples et permettent de distinguer les agrégats de façon satisfaisante. De plus, on remarque que si certaines régions sont moins texturées, la couleur est une information importante. Par contre, s'il y a davantage de texture, la moyenne et la variance sont des statistiques de premier ordre tout de même significatives. Les vecteurs obtenus sont appelés les *centres* des agrégats.

L'étape subséquente consiste à réaffecter chaque imasette à un nouvel agrégat tel que la distance entre l'imasette et le centre du nouvel agrégat est minimale. Par la suite, les centres des  $K$  nouveaux agrégats sont recalculés et la procédure de réaffectation est exécutée jusqu'à ce qu'il n'y ait plus de changement dans les regroupements d'imasettes. La convergence est rapide, i.e. environ 20 à 30 itérations pour 1024 imasettes de  $8 \times 8$  pixels chacunes, par exemple. La figure 2.2b. donne un exemple de l'étape de regroupement en agrégats avec  $K = 7$ .

### 2.2.3 Étape de fusion des agrégats (*merge*)

Cette étape est essentielle pour réduire la sur-segmentation résultant de l'étape précédente. Nous essaierons de fusionner les imasettes similaires et adjacentes et ce, jusqu'à la satisfaction d'un critère d'arrêt. À une étape donnée, nous fusionnons les paires d'imasettes adjacentes qui ont la plus petite valeur de fusion. Pour ce faire, nous affectons à

chaque frontière entre deux imagettes adjacentes une valeur de fusion qui est le produit du nombre de pixels dans la plus petite des deux imagettes et la distance euclidienne entre les vecteurs  $[\overline{L^*} \ \overline{u^*} \ \overline{v^*}]^T$  des deux imagettes. Voici la méthode proposée :

1. Pour chaque région, on calcule le vecteur moyen  $[\overline{L^*} \ \overline{u^*} \ \overline{v^*}]^T$ . En effet, l'oeil humain a tendance à faire la moyenne des fragments d'image presque uniformes. Puis on affecte, à chaque frontière entre deux régions distinctes, le produit du nombre de pixels dans la plus petite des deux imagettes et la distance euclidienne entre les vecteurs  $[\overline{L^*} \ \overline{u^*} \ \overline{v^*}]^T$  des deux imagettes. C'est la valeur de fusion.
2. Les frontières sont triées en ordre croissant de valeur de fusion.
3. Les deux régions les plus proches, i.e. de plus petite valeur de fusion sont fusionnées. Dans ce travail, pour accélérer la procédure, nous fusionnons, à cette étape, les 20 paires d'imagettes les plus proches.
4. On répète les étapes 1 à 3 jusqu'à l'obtention d'un critère d'arrêt.

Pour ce travail, le critère d'arrêt est le nombre final de régions. La figure 2.2c. donne un exemple de l'étape de fusion des agrégats jusqu'à l'obtention de 20 régions.

## 2.3 Extraction de caractéristiques de texture

Dans le chapitre précédent, les caractéristiques de texture étaient extraites d'imagettes monochromes de formes et de dimensions fixes. Avec les images de scènes extérieures, nous obtenons des imagettes couleur de dimensions et de formes variables.

D'abord, en ce qui concerne l'adaptation des méthodes d'extraction de caractéristiques aux images multispectrales, nous extrayons simplement les caractéristiques de texture du canal  $L^*$  de l'espace couleur CIE  $L^*u^*v^*$ . Ce canal représente la clarté psychométrique ou luminance. Il génère une image semblable à une image à niveaux de gris et il apparaît que l'information de texture y est plutôt bien préservée.

Ensuite, puisque les imageries issues de la segmentation sont de formes et de dimensions variables, il se peut que certaines statistiques extraites de la matrice des suites de niveaux de gris varient de beaucoup pour des textures similaires. Aucune mesure correctrice n'est apportée. Il sera intéressant de voir comment ce problème se reflètera dans les résultats issus de l'algorithme de sélection des caractéristiques.

Les méthodes d'extraction de caractéristiques appliquées au canal  $L^*$  de l'espace couleur CIE  $L^*u^*v^*$  utilisées sont les mêmes que celles décrites à la section 1.2. Le tableau 2.2 donne un résumé de ces caractéristiques.

## 2.4 Sélection des caractéristiques

La méthode de sélection des caractéristiques est décrite à la section 1.3. Nous utilisons aussi cette méthode dans ce volet pour les images couleur de scènes extérieures. Dans ce volet, cependant, nous avons affiné notre analyse de la sélection en l'appliquant à deux niveaux différents. En effet, nous pensons que les caractéristiques discriminantes ne sont pas nécessairement les mêmes pour toutes les classes. Premièrement, nous effectuons une sélection globale, i.e. que le critère d'évaluation de l'algorithme de sélection est appliqué à toutes les classes. Par conséquent, on obtient une unique liste ordonnée

Numéro	Caractéristique	Calculées à partir de(s) ...
1	Moyenne	Niveaux de gris
2	Médiane	Niveaux de gris
3	Mode	Niveaux de gris
4	Variance	Niveaux de gris
5	Écart interquantile	Niveaux de gris
6	Asymétrie ( <i>Skewness</i> )	Niveaux de gris
7	Aplatissement ( <i>Kurtosis</i> )	Niveaux de gris
8	Énergie	Matrice de cooccurrence
9	Entropie	Matrice de cooccurrence
10	Inertie	Matrice de cooccurrence
11	Différence de moment inverse d'ordre 2	Matrice de cooccurrence
12	Homogénéité	Matrice de cooccurrence
13	Corrélation	Matrice de cooccurrence
14	Probabilité maximum	Matrice de cooccurrence
15	Probabilité d'une suite de longueur n=2	Matrice de cooccurrence
16	Tendance de regroupement	Matrice de cooccurrence
17	"Short Run Emphasis Moments"	Matrice des suites
18	"Long Run Emphasis Moments"	Matrice des suites
19	Non-uniformité des niveaux de gris	Matrice des suites
20	Non-uniformité des suites de niveaux de gris	Matrice des suites
21	"Run Percentage"	Matrice des suites
22	Moyenne de différence	Histogramme des différences
23	Second moment	Histogramme des différences
24	Entropie de différence	Histogramme des différences
25	Modèle autorégressif	Niveaux de gris

Tableau 2.2 – Ensemble des caractéristiques décrites dans la section 1.2

des "meilleures" caractéristiques de texture pour toutes les classes. Deuxièmement, nous effectuons une sélection selon la classe, i.e. que le critère d'évaluation de l'algorithme de sélection est appliqué à chacune des classes. Par conséquent, on obtient une liste ordonnée des "meilleures" caractéristiques de texture pour chaque classe. Nous verrons, au chapitre 4, que les résultats semblent être meilleurs lorsqu'on utilise une méthode qui peut tenir compte des caractéristiques sélectionnées selon les classes.

## 2.5 Méthode de classification pour l'étiquetage des im- gettes

Les méthodes de classification utilisées pour l'étiquetage des imgettes sont la méthode des  $K$  plus proches voisins [5], un réseau de neurones RBF ("*Radial Basis Functions*") [5] et deux modèles de mixture de réseaux de neurones RBF [5]. L'algorithme des  $K$  plus proches voisins est utilisé comme critère d'évaluation pour la sélection des caractéristiques de texture et comme méthode de classification pour la validation de la sélection ou étiquetage. Les autres méthodes proposées sont utilisées seulement comme méthode de classification pour la validation de la sélection ou étiquetage en comparaison avec l'algorithme des  $K$  plus proches voisins.

La méthode des  $K$  plus proches voisins a été décrite à la section 1.4. Nous présenterons, dans les sous-sections qui suivent, le réseau de neurones RBF et le modèle de mixture de réseaux de neurones RBF.

### 2.5.1 Réseau de neurones RBF ("*Radial Basis Functions*")

#### Réseau de neurones RBF comme classificateur

Le réseau de neurones RBF a été conçu pour résoudre les problèmes de classification ou d'approximation. Dans cette section, l'idée est de construire un classificateur en s'inspirant du théorème de Bayes [5]. L'objectif est de modéliser la probabilité *a posteriori*  $P(C_k|\mathbf{x})$  pour chaque classe. Nous avons

$$\begin{aligned}
P(C_k|\mathbf{x}) &= \frac{p(\mathbf{x}|C_k)P(C_k)}{p(\mathbf{x})} \\
&= \frac{p(\mathbf{x}|C_k)P(C_k)}{\sum_{k'} p(\mathbf{x}|C_{k'})P(C_{k'})}
\end{aligned} \tag{2.1}$$

avec  $P(C_k)$  les probabilités *a priori*.

Pour modéliser chacune des distributions conditionnelles de classes  $p(\mathbf{x}|C_k)$ , on utilise un ensemble de  $M$  fonctions de bases qui tiendront compte des modes dans la représentation du problème de classification à résoudre. On suppose :

$$p(\mathbf{x}|C_k) \sim \sum_{j=1}^M p(\mathbf{x}|j)P(j|C_k). \tag{2.2}$$

Il est à noter que (2.2) est une approximation. La formule de décomposition des probabilités s'énonce comme suit :

$$p(\mathbf{x}|C_k) = \sum_{j=1}^M p(\mathbf{x}|j, C_k)P(j|C_k) \tag{2.3}$$

et (2.2) est égale à (2.3) si la connaissance des caractéristiques  $j$  permet de toujours trouver  $C_k$ .

Dans ce travail, nous recherchons une fonction  $g(\text{para}_k, x)$  dont les paramètres  $\text{para}_k$  permettent d'estimer  $p(\mathbf{x}|C_k)$ . Pour que  $g$  soit considérée comme une distribution de probabilité, nous la posons sous la forme particulière



$$\sum_{j=1}^M p(\mathbf{x}|j)P(j|C_k).$$

Si nous choisissons de représenter  $p(\mathbf{x}|j)$  par une gaussienne et si nous n'avons pas de limite sur le nombre  $M$ , il est alors possible d'approcher n'importe quelle distribution expérimentale, et ce jusqu'au niveau de précision désiré. Quand on choisit un  $M$  plutôt petit, on obtient une approximation grossière mais qui n'est pas nécessairement mauvaise en raison de la malédiction des grandes dimensions. Finalement, les tests présentés au chapitre 4 permettent de motiver l'utilisation de cette approximation.

En sommant sur toutes les classes, on obtient la probabilité

$$\begin{aligned} p(\mathbf{x}) &\sim \sum_k p(\mathbf{x}|C_k)P(C_k) \\ &= \sum_{j=1}^M p(\mathbf{x}|j)P(j) \end{aligned} \quad (2.4)$$

où  $P(j) = \sum_k P(j|C_k)P(C_k)$ .

L'équation (2.4) est une représentation de *mixture*. Les  $M$  fonctions de base  $p(\mathbf{x}|j)$  jouent le rôle des modes dans la représentation du problème de classification à résoudre et les  $P(j)$  jouent le rôle des coefficients de la mixture.

Pour retrouver la probabilité *a posteriori*, il suffit de substituer (2.2) et (2.4) dans le théorème de Bayes (2.1). On obtient :

$$\begin{aligned}
P(C_k|x) &\sim \frac{\sum_{j=1}^M p(\mathbf{x}|j)P(j|C_k)P(C_k) P(j)}{\sum_{j'=1}^M p(\mathbf{x}|j')P(j')} \frac{P(j)}{P(j)} \\
&= \sum_{j=1}^M \frac{p(\mathbf{x}|j)P(j)}{\sum_{j'=1}^M p(\mathbf{x}|j')P(j')} \frac{P(j|C_k)P(C_k)}{P(j)} \\
&= \sum_{j=1}^M w_{kj} \phi_j(x). \tag{2.5}
\end{aligned}$$

L'expression (2.5) représente un réseau de neurones RBF dans lequel les fonctions de bases normalisées de la première couche sont données par

$$\phi_j(x) = \frac{p(\mathbf{x}|j)P(j)}{\sum_{j'=1}^M p(\mathbf{x}|j')P(j')} = P(j|x)$$

et les poids de la seconde couche par

$$w_{kj} = \frac{P(j|C_k)P(C_k)}{P(j)} = P(C_k|j).$$

La première couche peut être interprétée comme la probabilité *a posteriori* ( $P(j|x)$ ) de l'appartenance au mode  $j$  étant donné  $\mathbf{x}$  et les poids de la seconde couche comme les probabilités *a posteriori* d'appartenance à la classe  $k$  ( $P(C_k|j)$ ) étant donné l'appartenance au mode  $j$ . La figure 2.3 donne l'architecture de réseau RBF correspondante.

### Optimisation du réseau de neurones RBF

L'architecture du réseau de neurones RBF est montrée à la Figure 2.3 avec

$$y_k = \sum_{j=1}^M w_{kj} \phi_j(x) + w_{k0}$$

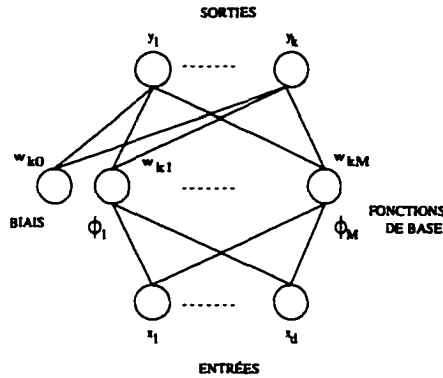


Figure 2.3 – Architecture de réseau RBF

et

$$\phi_k(x) = e^{-\frac{(\|x - \mu_j\|)^2}{2\sigma_j^2}}$$

Il existe plusieurs possibilités quant au choix de la fonction radiale  $\phi(\cdot)$ . La fonction gaussienne est celle qui est la plus communément utilisée. Les fonctions radiales  $\phi_j$  agissent comme neurones cachés. Les lignes connectant ces neurones cachés aux entrées sont les composantes  $\mu_{ji}$  où  $i = 1, \dots, d$  et  $j = 1, \dots, M$  des centres  $\mu_j$  des fonctions radiales  $\phi_j$ . À chaque  $\phi_j$  correspond un  $\sigma_j$  associé. Les poids  $w_{kj}$  où  $k = 1, \dots, k$  et  $j = 1, \dots, M$  relient les neurones cachés aux neurones de sortie. Le biais sert à compenser pour la différence entre la valeur moyenne des entrées et la valeur moyenne des sorties.

Pour l'apprentissage de la première couche, nous utilisons la version *on-line* de l'algorithme de K-moyenne. Cet apprentissage est dit *non-supervisé* car il utilise seulement le vecteur d'entrée  $x$ . L'algorithme de K-moyenne utilisé est le suivant :

1. Fixer  $K$ , le nombre de centres

2. Initialiser les  $\mu_j$ . Pour chaque centre :

- Sélectionner un patron d'apprentissage  $\mathbf{x}$  au hasard parmi les pixels d'apprentissage
- $\mu_j$  est initialisé avec ce patron d'apprentissage

3. Mise-à-jour des  $\mu_j$  de façon stochastique. Pour un nombre prédéterminé d'itérations :

- Sélectionner un patron d'apprentissage  $\mathbf{x}$  au hasard dans l'image
- Trouver le centre le plus près de ce pixel, i.e.  $\mu_{j_0}$  tel que  $\|\mathbf{x} - \mu_{j_0}\|^2 = \min\{\|\mathbf{x} - \mu_j\|^2 | j = 1, \dots, L\}$
- Ajuster le centre le plus près de ce pixel, i.e.  $\Delta\mu_{j_0} = \gamma(\mathbf{x} - \mu_{j_0})$

4. Déterminer les largeurs  $\sigma_j$ , i.e  $\sigma_j = \psi \min\{\|\mu_j - \mu_l\| | l = 1, \dots, K \text{ et } j \neq l\}$

Cette étape permet de déterminer les paramètres  $\mu_j$  et  $\sigma_j$  pour les fonctions radiales  $\phi_j$ .

Pour l'apprentissage de la seconde couche, nous utilisons une adaptation de l'algorithme de rétropropagation du gradient (RPG). Cet apprentissage est dit *supervisé* car il utilise les vecteurs d'entrée  $\mathbf{x}$  et les sorties désirées  $d_i$  qui sont les classes associées aux patrons d'apprentissage. Cette étape permet de déterminer les poids de la couche cachée par la correction d'erreur. L'algorithme de RPG trouve les erreurs commises par les neurones de sorties et les neurones cachés. Il est très populaire quoique pas très efficace car il est difficile de choisir un pas de correction optimal et de plus, le gradient ne pointe pas nécessairement sur le minimum. Toutefois, il est simple et a la capacité de résoudre des problèmes complexes.

## 2.5.2 Mixture de réseaux de neurones RBF

Le modèle de mixture des réseaux de neurones fournit un mécanisme permettant de subdiviser la solution d'un problème entre différents réseaux de neurones. Cela est possible en utilisant un ensemble de réseaux dont chacun (expert) est associé à une source particulière en entrée ou, comme c'est le cas dans ce travail, à un sous-ensemble spécifique de caractéristiques en entrée. Il existe deux grandes catégories de méthodes pour composer l'architecture d'un modèle de mixture. L'une est une approche globale appelée mixture hiérarchique. Elle est décrite dans [19]. L'autre est plus heuristique et est utilisée dans [12] et [30]. Même si l'approche globale permet d'entraîner tous les paramètres en même temps, l'approche heuristique "locale" permet d'obtenir de meilleurs résultats. Cela s'explique par le fait que la connaissance du problème de classification peut y être plus aisément intégrée. Dans ce travail, nous adoptons la seconde approche. Cela signifie que nous devons entraîner chaque réseau de neurones RBF (expert) individuellement et ensuite, déterminer les stratégies pour combiner les sorties des différents experts.

En fait, dans un problème comme le nôtre, nous essayons d'apprendre une fonction qui peut être différente pour différentes régions de l'espace d'entrée. C'est une raison pour laquelle nous avons appliqué l'algorithme de sélection de caractéristiques d'abord globalement et, par la suite, selon la classe. De cette façon, nous pouvons expérimenter avec des mixtures dont les experts qui les composent peuvent prendre soit le même vecteur en entrée pour toutes les classes, soit un vecteur d'entrée spécifique pour chaque classe. Dans ce travail, il y a un expert pour chaque classe. Les stratégies pour combiner les différents experts sont des fonctions simples comme le maximum et la majorité. Les figures 4.8 et 4.9 montrent de telles architectures. L'apprentissage des réseaux et les stratégies de combinaison entre les différentes sorties sont décrits plus loin dans ce travail.

## 2.6 Conclusion

Dans cette section, nous avons présenté une méthode d'étiquetage pour les images couleur de scènes extérieures. Il s'agit d'une extension à la méthode développée dans le chapitre précédent. Premièrement, les images sont converties de l'espace couleur RGB à l'espace couleur  $L^*u^*v^*$ . Dans cet espace, deux couleurs proches sont également perceptuellement similaires. Les images couleur de scènes extérieures comprennent souvent plusieurs textures différentes. Pour cette raison, elles nécessitent une méthode de segmentation qui permet de subdiviser les images en partitions d'images de dimensions moindres et aussi homogènes que possible en texture. Ensuite, nous avons extrait, pour chaque imagerie obtenue de la segmentation, un vecteur de 25 caractéristiques de texture. Afin de diminuer la dimension de l'espace de caractéristiques, nous avons appliqué un algorithme de sélection séquentielle vers l'avant. La sélection est effectuée d'abord globalement et, par la suite, selon la classe. Dans le chapitre 4, nous présentons quelques résultats expérimentaux. Ceux-ci démontrent que la sélection permet de contrer la "malédiction des grandes dimensions". Aussi, on peut déduire, par la consistance des résultats obtenus, que certaines caractéristiques ont une capacité discriminative supérieure. Finalement, il s'avère qu'entre la sélection globale et la sélection selon la classe, c'est cette dernière qui semble être la plus avantageuse. Il semble donc que ce ne sont pas les mêmes caractéristiques qui décrivent bien chaque type de texture. Finalement, nous effectuons une classification ou étiquetage des imageries. Les résultats expérimentaux du chapitre 4 montrent les pourcentages de bonne classification. Parmi les trois types de méthodes utilisés, le modèle de mixture de réseaux de neurones RBF est celui qui donne les meilleurs résultats.

# CHAPITRE 3

## INTERFACE GRAPHIQUE

La présente section a pour objectif de décrire les menus **Texture** ajoutés à l'interface graphique de l'application "Vision" en cours de développement par des étudiants du groupe de Moivre (MOdélisation en Imagerie, Vision et REseaux de neurones) du département de mathématiques et d'informatique de l'Université de Sherbrooke. Les fonctions de base utilisées appartiennent à la librairie de fonctions NVLib 1.0 [21]. Ce volet de l'application a été développé afin d'expérimenter avec la méthode proposée dans mon projet de maîtrise, i.e. la segmentation et l'étiquetage manuel des images ainsi que l'extraction et la sélection des caractéristiques de texture pour l'étiquetage des images couleur de scènes extérieures.

Beaucoup a été écrit à propos des interfaces graphiques et sur les attraits qu'elles présentent. La présente interface est développée avec Microsoft Visual C++ 5. Le rôle d'une interface graphique est de réagir à des actions de l'utilisateur et aussi à des instructions données par le programme. En effet, l'élément central de tout programme Windows, c'est la boucle de message. Windows est en permanence à l'écoute de toutes les instructions

provenant de l'utilisateur, il les envoie sous forme de messages aux applications en cours d'exécution qui renvoient ces messages aux fenêtres concernées.

L'interface graphique de l'application "Vision" permet d'expérimenter la segmentation et l'étiquetage manuel des images ainsi que l'extraction et la sélection des caractéristiques de texture pour l'étiquetage des images couleur de scènes extérieures. Elle offre à la fois convivialité, facilité d'utilisation, flexibilité pour tester divers paramètres et donne un aspect professionnel à l'application. Les sections qui suivent vont nous montrer comment utiliser les items des menus **Texture** de l'application "Vision".

### 3.1 Commandes interactives

Dès qu'une image est ouverte, un ensemble de menus contenant des commandes interactives est proposé. Le menu **Texture** contient toutes les commandes servant à la préparation des imageries pour la sélection de caractéristiques de texture et l'étiquetage des imageries. On y retrouve les commandes permettant d'effectuer la segmentation des images, l'affectation manuelle des étiquettes aux imageries résultant de la segmentation et le calcul des vecteurs de caractéristiques de texture associés aux imageries. La Figure 3.1 montre le menu **Texture**. Pour l'instant, seule la commande **Segmentation** est accessible car elle est préalable aux autres commandes du menu.





Figure 3.1 – Menu **Texture** de l'application "Vision" pour la segmentation, l'étiquetage manuel et le calcul des caractéristiques

### 3.1.1 Segmentation

La méthode proposée dans mon projet de maîtrise pour la segmentation peut se résumer comme suit :

1. Initialisation de la segmentation avec des imagettes de taille  $m \times n$
2. Regroupement ou "*clustering*" de ces imagettes selon des caractéristiques de texture globales sur les imagettes avec un algorithme de K-moyenne

3. Fusion des regroupements selon les critères de distance définie sur  $L^*u^*v^*$  et/ou de taille jusqu'à l'obtention du nombre de régions désiré

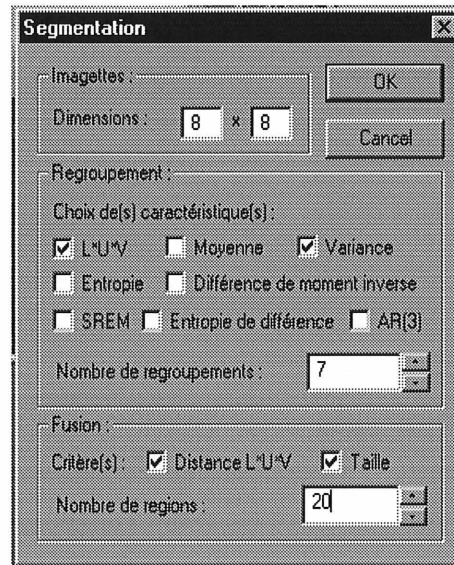


Figure 3.2 – Boîte de dialogue de la commande **Segmentation**

La boîte de dialogue **Segmentation** (Figure 3.2) contient tous les contrôles disponibles pour la segmentation. En voici une brève description :

#### **Imagettes :**

- Dimensions: permet de spécifier les dimensions en pixels des imagettes pour l'initialisation de la segmentation.

#### **Regroupement :**

- Choix de(s) caractéristique(s) : permet de spécifier le(s) caractéristique(s) globale(s) sur les imagettes qui serviront pour calculer les centres de l'algorithme de K-

moyenne. Les caractéristiques globales proposées sont  $L^*u^*v^*$  (représentation de la couleur), la moyenne et la variance (statistiques de premier ordre), l'entropie et la différence de moment inverse (statistiques de second ordre), le SREM ou *Short Run Emphasis Moments* (statistique extraite de la matrice des suites de niveaux de gris), l'entropie de différence (statistique extraite de l'histogramme des différences en valeurs absolues) et AR(3) (modèle autorégressif d'ordre 3). Toutes ces caractéristiques, sauf la première, sont extraites du canal  $L^*$  de l'image.

- Nombre de regroupements : permet de spécifier la valeur de la constante K de l'algorithme de K-moyenne.

#### **Fusion :**

- Critère(s) : permet de spécifier le(s) critère(s) de fusion des regroupements d'images. Les critères proposés sont la distance définie sur  $L^*u^*v^*$  si l'on veut fusionner les régions adjacentes "proches" en distance définie sur  $L^*u^*v^*$  et la taille des images si l'on veut fusionner les images de petites tailles. Dans le cas où les deux options sont sélectionnées, le critère de fusion devient la distance définie sur  $L^*u^*v^*$  multipliée par la taille.
- Nombre de régions : permet de spécifier le nombre final de régions désiré pour la segmentation.

La Figure 3.3 montre un exemple d'images segmentées. On remarque que lorsque les images sont segmentées, les commandes **Étiquetage manuel** et **Calcul des caractéristiques de texture** sont accessibles.

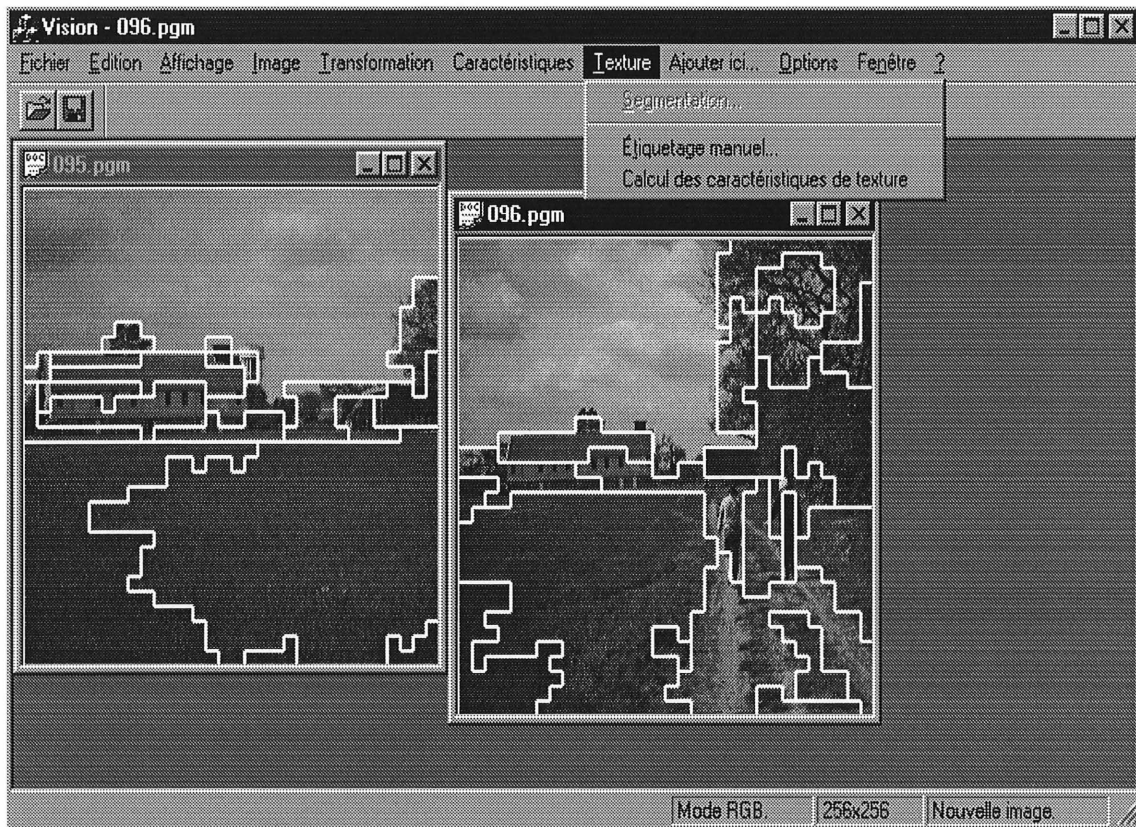


Figure 3.3 – Exemple d'images segmentées

### 3.1.2 Étiquetage manuel

La boîte de dialogue **Étiquetage manuel** (Figure 3.4) contient une copie de l'image segmentée active ainsi que les champs **Remplacer l'étiquette** et **Par l'étiquette**. En voici une brève description :

- Remplacer l'étiquette : affiche l'étiquette déjà affectée à l'imagette se trouvant sous la souris. Ce champ est mis-à-jour au fur et à mesure que l'on se déplace sur la copie de l'image segmentée avec la souris.

- Par l'étiquette: permet de saisir le nom de l'étiquette à affecter à l'imagette sur laquelle on vient de cliquer (bouton de gauche) dans la boîte de dialogue.

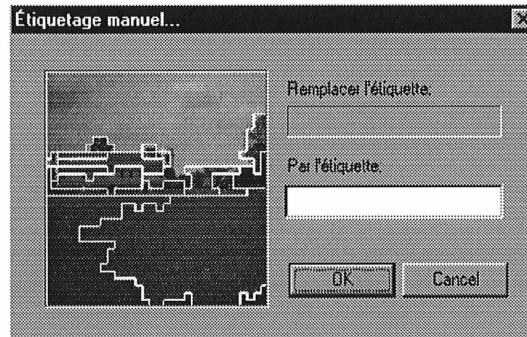


Figure 3.4 – Boîte de dialogue de la commande **Étiquetage des imagettes**

### 3.1.3 Calcul du vecteur de caractéristiques

Cette commande enclenche le calcul des vecteurs des 25 caractéristiques de texture décrits dans ce document.

## 3.2 Commandes automatiques

La Figure 3.5 montre la fenêtre principale de l'application avec le menu **Texture**. Les commandes qui s'y trouvent sont exécutées automatiquement sur notre base de données d'images de scènes extérieures.

Dans le menu **Texture**, on retrouve la commande **Sélection des caractéristiques** qui permet la sélection des "meilleures" caractéristiques de texture pour l'étiquetage. Cette

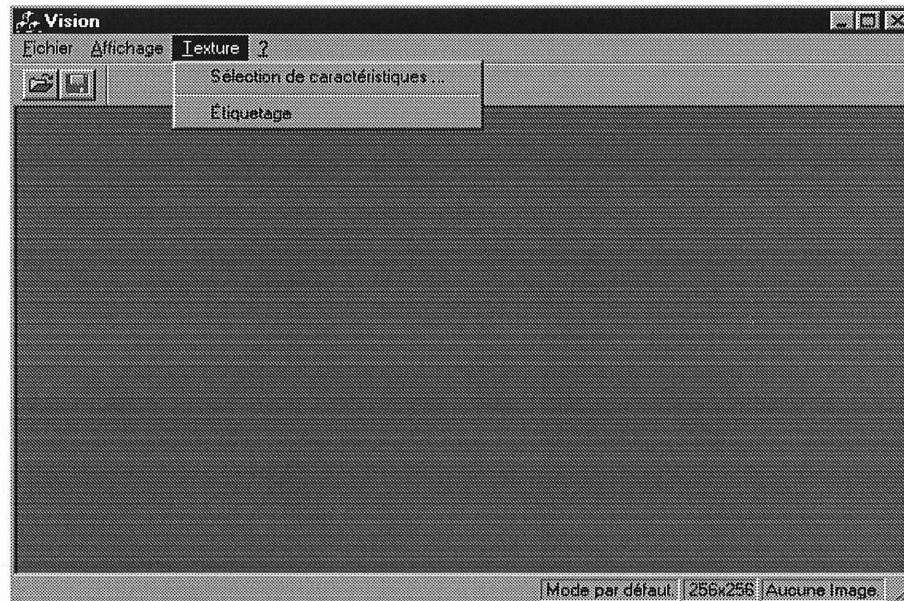


Figure 3.5 – Fenêtre principale de l'application "Vision" avec le menu **Texture**

commande permet d'aller chercher directement dans un fichier les images sur lesquelles la sélection de caractéristiques est appliquée et ne nécessite pas l'ouverture des images. On retrouve aussi dans ce menu la commande **Étiquetage** qui permet d'étiqueter les images en allant également chercher directement dans un fichier les images sur lesquelles les algorithmes de classification sont appliqués.

### 3.2.1 Sélection des caractéristiques

La boîte de dialogue **Sélection des caractéristiques** (Figure 3.6) contient le champ nombre de caractéristiques à sélectionner. En voici une brève description :

- Nombre de caractéristiques à sélectionner : permet de spécifier le nombre de "meilleures" caractéristiques de texture pour l'étiquetage d'images couleur de scènes extérieures

que l'on désire sélectionner.

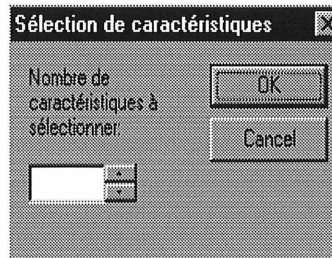


Figure 3.6 – Boîte de dialogue de la commande **Sélection des caractéristiques**

La boîte de dialogue **Caractéristiques sélectionnées** (Figure 3.7) contient les résultats de la commande **Sélection des caractéristiques** que l'on vient de décrire.

### 3.2.2 Étiquetage

Lorsque l'on exécute la commande **Étiquetage**, on obtient les résultats de l'étiquetage dans la boîte de dialogue **Matrice de confusion** (Figure 3.8). Les lignes représentent les classes obtenues par classification. Les colonnes représentent les classes réelles.

## 3.3 Conclusion

Cette interface a été très utile, comme nous l'avons dit, pour effectuer la segmentation, l'étiquetage manuel des images, l'extraction des caractéristiques de texture, la sélection de caractéristiques de texture et l'étiquetage. C'est avec convivialité, facilité d'utilisation et flexibilité qu'elle a permis de tester divers paramètres et ce, tout en offrant un aspect professionnel à l'application.



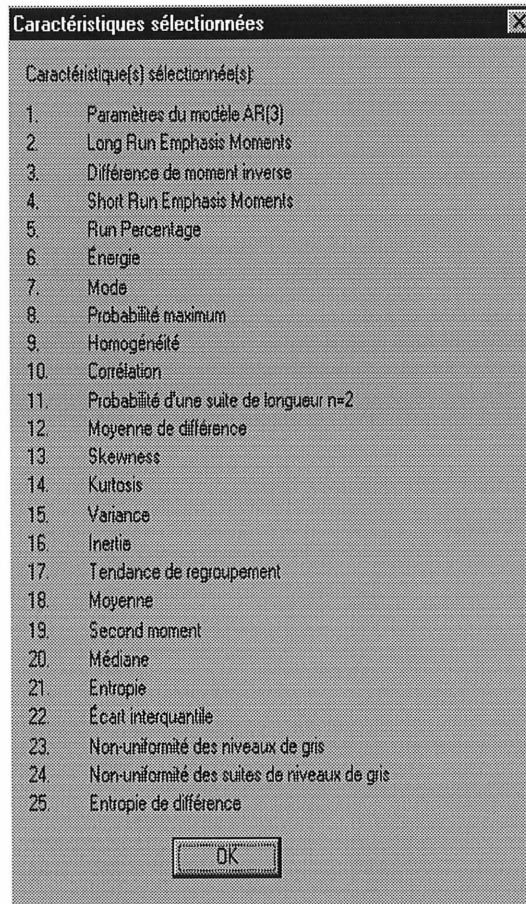


Figure 3.7 – Boîte de dialogue *Caractéristiques sélectionnées*



The image shows a dialog box titled "Matrice de confusion" with a close button in the top right corner. The dialog contains a 5x5 grid of numbers. At the bottom center, there is an "OK" button.

66	19	8	11	9
4	21	3	16	5
7	17	47	6	17
2	1	12	41	6
1	3	0	0	1

Figure 3.8 – *Boîte de dialogue Matrice de confusion*

# CHAPITRE 4

## RÉSULTATS EXPÉRIMENTAUX

### 4.1 Résultats expérimentaux sur les images à niveaux de gris de texture de Brodatz

#### 4.1.1 Base de données d'images

Ce volet concerne les images à niveaux de gris de texture de Brodatz. Toutes les images que nous avons utilisées contiennent une seule texture. Les textures de Brodatz sont une collection d'images de textures monochromes, plutôt homogènes et photographiées sous un éclairage de studio à angle parallèle au plan du film. Elles sont tirées de "Textures: a Photographic Album for Artists and Designers" par Phil Brodatz (Dover Publications, NY 1966). Notre base de données contient 30 images. La figure 4.1 montre un échantillon des images que nous avons utilisées.

Chaque image de texture est subdivisée tantôt en 16 imageries de  $64 \times 64$  pixels, tantôt

en 64 imagettes de  $32 \times 32$  pixels. Toutes les imagettes ont été étiquetées dans les 30 classes qui correspondent aux 30 images de texture de Brodatz. Afin de pallier le nombre limité d'images à notre disposition, nous avons divisé notre base de données d'images en trois groupes,  $T_1, T_2$  et  $T_3$ , mutuellement disjoints. Chaque groupe contient dix images. Le nombre d'imagettes dans chaque  $T_i$  dépend de la façon dont les images sont subdivisées.

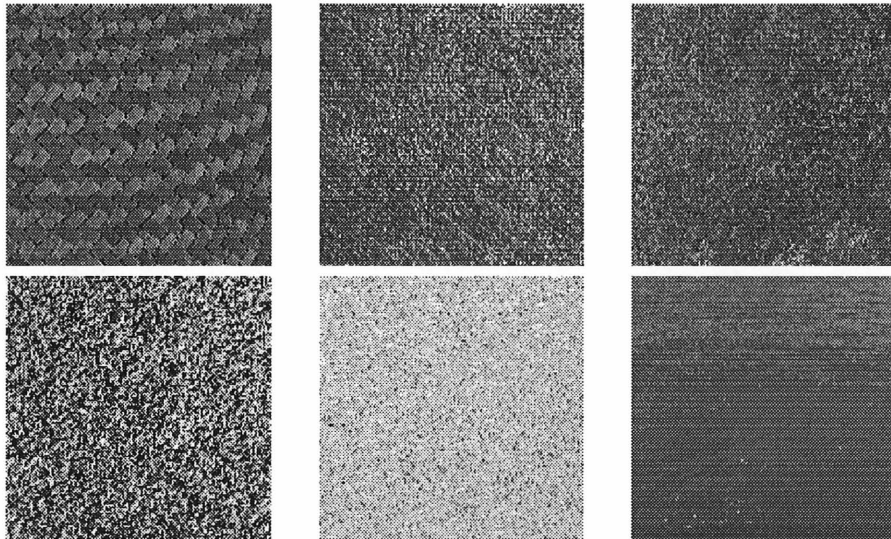


Figure 4.1 – *Exemples d'images à niveaux de gris de texture de Brodatz*

Pour chaque test de sélection impliquant un ensemble d'images  $T_i$ , les données d'apprentissage et les données de généralisation sont obtenues comme suit. Les imagettes faisant partie de la diagonale du découpage ou situées près de la diagonale (voir section 1.1) ainsi que leurs classes associées formeront les sous-ensembles d'apprentissage. Le reste des imagettes pour chaque ensemble d'imagettes formeront les sous-ensembles de généralisation, i.e les sous-ensembles d'imagettes auxquelles nous affecterons une classe ou étiquette. Les cardinalités des sous-ensembles d'apprentissage correspondent au quart du nombre total des imagettes pour chaque ensemble d'imagettes  $T_i, i = 1, 2, 3$ . Nous avons essayé deux résolutions différentes pour le découpage des imagettes, soit un découpage en 16 imagettes et un découpage en 64 imagettes. Le tableau 4.1 donne la description

des tests effectués.

Test	Ensemble d'images	Découpage (imagettes/images)
1	$T_1$	16
2	$T_2$	16
3	$T_1$	64
4	$T_3$	64

Tableau 4.1 - *Descriptions des tests effectués sur les images à niveaux de gris de texture de Brodatz*

#### 4.1.2 Sélection des caractéristiques de texture

##### Performance de classification pour la sélection

La Figure 4.2 présente les pourcentages d'erreur de classification étant donné la sélection des  $d$  "meilleures" caractéristiques de texture pour les tests 1, 2, 3 et 4.

Tout d'abord, on remarque que les courbes de pourcentages d'erreur de classification sont un peu plus prononcées lorsque le découpage est en seize imagettes. Toutefois, les allures générales des courbes sont les mêmes pour les quatre tests. La meilleure performance a été enregistrée pour le test 2. Par conséquent, nous utiliserons le découpage en seize imagettes pour l'étiquetage comme moyen de validation des résultats. Aussi, il est intéressant d'observer la "malédiction des grandes dimensions" manifestée dans les résultats, i.e. la tendance des résultats à ne pas s'améliorer ou même à se détériorer lorsque le nombre de caractéristiques augmente.

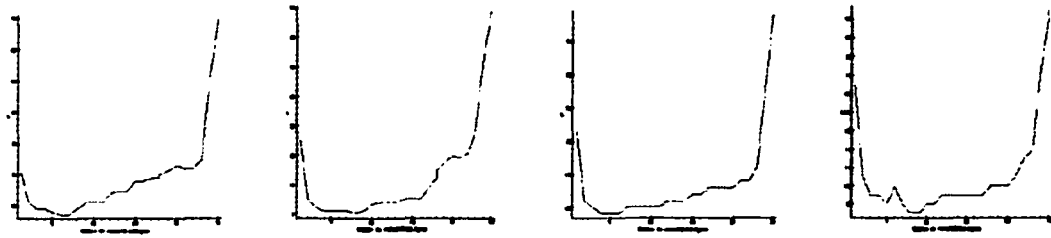


Figure 4.2 – *Pourcentages d'erreur de classification en fonction du nombre de "meilleures" caractéristiques sélectionnées pour les tests 1, 2, 3 et 4*

### **Ordre de sélection des caractéristiques**

Dans le Tableau 4.2, on trouve l'ordre (colonne) dans lequel les caractéristiques (lignes) ont été sélectionnées. Le ●, le ○, le \* et le ★ indiquent des fréquences de 1, 2, 3 et 4 respectivement. Par exemple, la première ligne du tableau signifie que la première caractéristique (moyenne) a été sélectionnée sixième, dixième, quinzième et dix-septième lors des tests. La ligne 25 indique que la vingt-cinquième caractéristique (paramètres du modèle autorégressif) a été sélectionnée première pour les quatre tests.

Le modèle autorégressif est très efficace. Dans les tests préliminaires, nous n'avions pas inclu cette caractéristique. Or, il s'était avéré que les caractéristiques qui performaient bien étaient les caractéristiques de la matrice de cooccurrence, de la matrice des suites et de l'histogramme des différences, i.e. les caractéristiques telles que l'information spatiale entre les pixels est considérée. C'est cette constatation qui nous a incités à inclure le modèle autogéressif qui utilise des estimés linéaires du niveau de gris d'un pixel, étant donné les niveaux de gris de ses voisins, pour caractériser la texture.

Caract. / Rang	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1 Moyenne					•					•					•		•								
2 Médiane													◦			•		•							
3 Mode															•			•			•	•			
4 Variance														◦			•		•						
5 Écart interquantile			•								•	•					•								
6 Asymétrie																								*	•
7 Aplatissement																•	•	•	•						
8 Énergie																•					◦		•		
9 Entropie				•	◦	•																			
10 Inertie									•		•	•										•			
11 Diff. mom. inv.			•		•	•		•																	
12 Homogénéité							•	◦		•															
13 Corrélacion																								•	*
14 Prob. maximum																							•	*	
15 Prob. suite n=2			•					•	•													•			
16 Tend. de regroup.																•		•	◦						
17 "SREM"				◦	•	•																			
18 "LREM"															•						*				
19 Non-unif. des NG		•		•						•			•												
20 Non-unif. SNG							•			•	◦														
21 "Run Percentage"	◦							•			•														
22 Moy. de diff.								•					•	◦											
23 Second moment												•			•							◦			
24 Entropie de diff.	◦	•						•																	
25 Modèle AR	*																								

Tableau 4.2 – *Ordre de sélection (colonnes) des caractéristiques (lignes) pour les tests 1, 2, 3 et 4. Le •, le ◦, le \* et le \* indiquent des fréquences de 1, 2, 3 et 4 respectivement*

On remarque aussi que les caractéristiques sont sélectionnées dans un ordre qui est assez consistant d'un test à l'autre. Par exemple, la caractéristique 25 est toujours sélectionnée la première, les caractéristiques 9, 11, 12, 17 et 24 sont toujours sélectionnées parmi les dix premières et les caractéristiques 6, 7, 8, 13, 14 et 16 sont toujours sélectionnées parmi les dix dernières. Par conséquent, on peut dire que la méthode de sélection utilisée donne

Numéro	Caractéristique sélectionnée	Particularité
25	Modèle autorégressif	Tient compte des niveaux de gris de ses voisins
24	Entropie de différence	Tient compte de la distribution des niveaux de gris
17	"Short Run Emphasis Moments"	Donne de l'information sur la "grossièreté" de la texture
9	Entropie	Tient compte de la distribution des niveaux de gris
11	Diff. de moment inv. d'ordre $n = 2$	Tient compte de la concentration des éléments autour de la diagonale dans la matrice de cooccurrence

Tableau 4.3 – *Liste des "meilleures" caractéristiques sélectionnées avec  $d = 5$*

des résultats cohérents pour les tests effectués.

D'après les résultats obtenus (voir Figure 4.2), il paraît raisonnable de choisir  $d = 5$ . En effet, ce seuil semble être un bon compromis entre la taille du sous-ensemble de caractéristiques sélectionnées et la performance de bonne classification. On remarque que cinq caractéristiques sont toujours sélectionnées parmi les neuf premières. Le Tableau 4.3 montre ces caractéristiques. On remarque qu'elles tiennent toutes compte des relations spatiales entre les pixels.

### 4.1.3 Étiquetage

Dans cette section, on effectue un étiquetage comme test de validation sur un nouveau sous-ensemble de dix images de texture de Brodatz.

Les images sont découpées en seize imagettes, le vecteur de caractéristiques est de dimension  $d = 5$  et les caractéristiques utilisées sont les paramètres du modèle autorégressif, l'entropie de différence (histogramme des différences), "Short Run Emphasis Moments" (matrice des suites), l'entropie (matrice de cooccurrence) et la différence de moment inverse d'ordre  $k = 2$  (matrice de cooccurrence). La méthode de classification est l'algorithme des  $K$  plus proches voisins avec  $K = 3$ .

Voici la matrice de confusion obtenue pour l'étiquetage :

$$\begin{pmatrix} 11 & 0 & 0 & 0 & 0 & 0 & 3 & 3 & 0 & 0 \\ 0 & 12 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 12 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 12 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 12 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 12 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & 2 & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 12 \end{pmatrix}$$

Les lignes représentent les classes obtenues par classification. Les colonnes représentent les classes réelles. Par exemple, pour les classes 2,3,4,5,6 et 10, les 12 imagettes à classer sont correctement classifiées dans leurs classes réelles. Par contre, pour la classe 1, 11 des imagettes à classer sont correctement classifiées dans leurs classes réelles et 1 des imagettes est incorrectement classifiée dans la classe 3. La Figure 4.3 montre les images qui correspondent à ces classes dont la performance de classification est très bonne.



Pour trois classes qui restent, la classification est mauvaise. Par exemple, pour la classe 7, seulement 6 des 12 imajettes à classifier sont classifiées, les 6 autres imajettes sont rejetées pour la classification. Des 6 imajettes classifiées, 3 imajettes sont incorrectement classifiées dans la classe 1, 1 imajette est incorrectement classifiée dans la classe 8 et 2 imajettes sont incorrectement classifiées dans la classe 9. La classification n'est pas très bonne non plus pour les classes 8 et 9. La Figure 4.4 montre les images qui correspondent à ces classes dont la performance de classification est mauvaise. Les caractéristiques sélectionnées ne sont pas efficaces avec ces images. Cela peut découler du fait que les ensembles de tests n'incluaient pas d'image avec des textures similaires.

Pour toutes les classes, le pourcentage de bonne classification est de 73% et le pourcentage de rejet est de 16%.

#### **4.1.4 Sommaire**

Dans ce volet du travail, nous avons expérimenté différentes méthodes d'extraction de caractéristiques de texture et leur avons appliqué une méthode de sélection de caractéristiques pour réduire la dimension du vecteur de caractéristiques tout en maintenant une bonne capacité de discrimination. Nous avons obtenu des résultats de classification intéressants.

Notre prochain objectif est d'étendre notre méthode aux images couleur de scènes extérieures contenant de multiples textures. Cette extension amène les problèmes suivants. D'abord, il faut segmenter les images en régions distinctes de textures homogènes. Ensuite, il faut modifier les méthodes d'extraction de caractéristiques de texture pour les

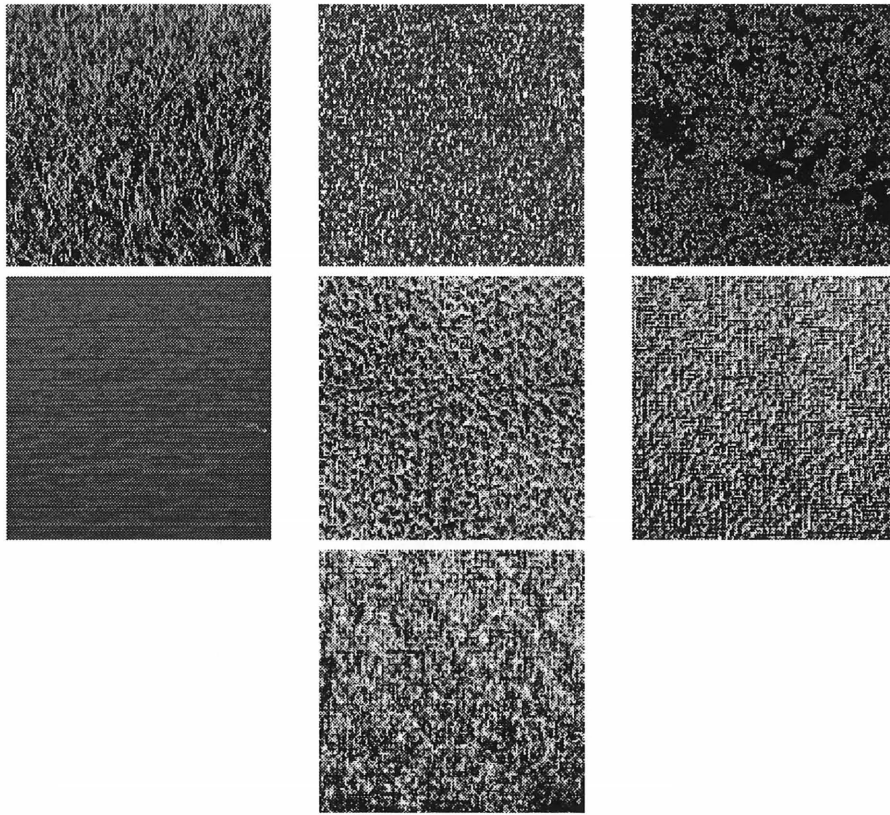


Figure 4.3 – Images correspondant aux colonnes 1,2,3,4,5,6 et 10 de la matrice de confusion et pour lesquelles la performance de classification est bonne

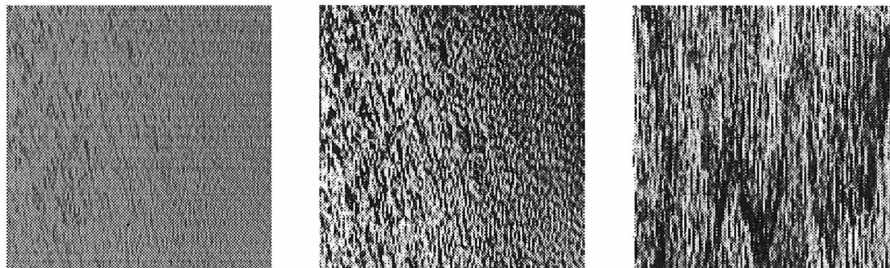


Figure 4.4 – Images correspondant aux colonnes 7,8 et 9 de la matrice de confusion et pour lesquelles la performance de classification est mauvaise

images couleur. Enfin, pour obtenir de meilleurs résultats d'étiquetage, nous introduisons aussi de nouvelles méthodes de classification dans notre système.

## 4.2 Résultats expérimentaux sur les images couleur de scènes extérieures

### 4.2.1 Base de données d'images

Ce volet concerne les images couleur de scènes extérieures. La plupart des images que nous avons utilisées contiennent plusieurs textures. Elles sont choisies de différentes sources incluant VisTex (développé par Rosalind Picard, Chris Graczyk, Steve Mann, Josh Wachman, Len Picard et Lee Campbell du *Media Laboratory* au *MIT, Massachusetts*) et la base de données utilisée dans [7]. Notre base de données contient 92 images. La figure 4.5 montre un échantillon des images que nous avons utilisées.

Chaque image de texture est segmentée en 20 imasettes de dimensions et de formes variables. Des 1840 imasettes obtenues, 456 imasettes ayant une sémantique significative ont été étiquetées, i.e. qu'elles ont été affectées à l'une des cinq classes suivantes : "Ciel", "Eau", "Arbre", "Mur" et "Herbe". Les 1384 autres imasettes n'ont pas été étiquetées soit parce qu'elles n'appartenaient clairement à aucune des cinq classes mentionnées ci-haut, soit parce qu'elles étaient trop petites ou trop foncées ou qu'elles n'appartenaient pas clairement à l'une des cinq classes : par exemple, une imasette qui contient un mélange uniforme de feuillage et de ciel. Afin de pallier le nombre limité d'imasettes étiquetées à notre disposition, nous avons subdivisé notre ensemble de 456 imasettes étiquetées en quatre sous-ensembles de même cardinalité et mutuellement disjoints :  $A_1$ ,  $A_2$ ,  $A_3$  et  $A_4$  que nous utiliserons par la suite pour les tests.



Figure 4.5 – Exemples d'images couleur de scènes extérieures

#### 4.2.2 Sélection des caractéristiques de texture

Soit  $S$  l'ensemble des 456 imageries étiquetées disponibles pour tester l'algorithme de sélection et soit  $A_i, i = 1, 2$  deux des quatre sous-ensembles mutuellement disjoints présentés à la section 4.2.1. Un ensemble de tests sont effectués à partir de  $A_1$  pour l'apprentissage et  $S \cap \overline{A_1}$  pour la généralisation. Un autre ensemble de tests sont effectués à partir de  $A_2$  pour l'apprentissage et  $S \cap \overline{A_2}$  pour la généralisation. Le classificateur est la méthode des  $K$  plus proches voisins avec  $K = 1$  et  $K = 3$ . Finalement, nous avons appliqué l'algorithme de sélection à deux niveaux. Premièrement, nous effectuons une sélection globale, i.e. que le critère d'évaluation de l'algorithme de sélection est appliqué à toutes les classes. Par conséquent, on obtient une unique liste ordonnée des "meilleures" caractéristiques de texture pour toutes les classes. Deuxièmement, nous effectuons une sélection selon la classe, i.e. que le critère d'évaluation de l'algorithme de sélection est appliqué à chacune des classes. Par conséquent, on obtient une liste ordonnée des "meilleures" caractéristiques de texture pour chaque classe. Le tableau 4.4 donne la description des

tests effectués.

Test	<i>KNN</i> K=...	Sous-ensemble d'apprentissage	Niveau de sélection
1	1	$A_1$	Globale
2	1	$A_1$	Selon la classe
3	1	$A_2$	Globale
4	1	$A_2$	Selon la classe
5	3	$A_1$	Globale
6	3	$A_1$	Selon la classe
7	3	$A_2$	Globale
8	3	$A_2$	Selon la classe

Tableau 4.4 – *Descriptions des tests effectués sur les images couleur de scènes extérieures*

### Performance de classification pour la sélection

La figure 4.6 montre l'erreur de classification en fonction du nombre de caractéristiques sélectionnées pour chaque test. Les lignes pleines représentent les résultats pour la sélection globale et les lignes pointillées, les résultats pour la sélection selon les classes.

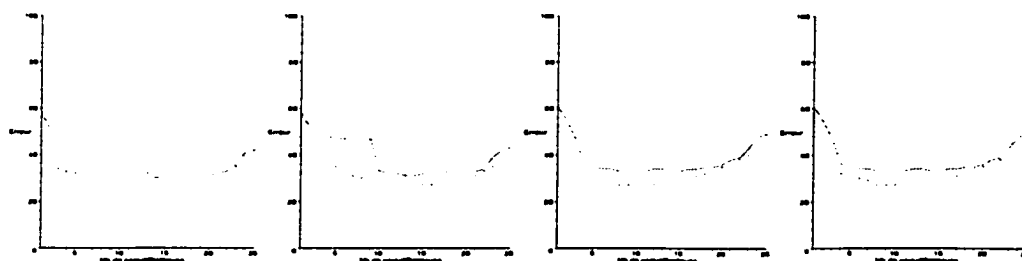


Figure 4.6 – *Comparaison des pourcentages d'erreur de classification en fonction du nombre de caractéristiques sélectionnées pour tous les tests. La ligne pleine est pour la sélection globale et la ligne pointillée est pour la sélection selon les classes*

Nous avons appliqué l'algorithme de sélection à l'ensemble des caractéristiques à deux

niveaux différents, i.e. globalement et selon les classes car nous pensons que les caractéristiques discriminantes ne sont pas nécessairement les mêmes pour toutes les classes. Nous verrons dans les résultats subséquents que les caractéristiques sélectionnées selon les classes semblent améliorer la performance de classification, par exemple, lorsque nous utilisons une mixture de réseaux de neurones.

Dans la figure 4.6, la ligne pointillée représente la somme du nombre d'images correctement classifiées pour chaque classe (avec les caractéristiques sélectionnées pour ces classes respectivement) divisé par le nombre total d'images à classifier. Comme nous avons un ensemble de caractéristiques sélectionnées pour chacune des classes, le nombre de caractéristiques nécessaire pour obtenir ces résultats peuvent varier entre  $d$ , la valeur sur l'axe des  $x$ , et un total possible de  $d \times 5$  caractéristiques puisqu'il y a 5 classes. Pour cela, il est important de se rappeler que les résultats doivent être vus comme une comparaison parce que le nombre total de caractéristiques dont on a besoin pour obtenir ces résultats n'est pas en accord avec le but initial de la procédure de sélection qui est de réduire le plus possible le nombre de caractéristiques. Très souvent, nous remarquons que l'erreur de classification (le premier critère d'évaluation de l'algorithme de sélection) de la sélection selon les classes est plus petite que l'erreur de classification de la sélection globale.

La "malédiction des grandes dimensions" est également visible dans les résultats présentés à la figure 4.6. L'erreur de classification tend soit à se maintenir ou à se détériorer lorsque le nombre de caractéristiques sélectionnées augmente au lieu de diminuer comme notre intuition pourrait nous laisser croire.

## Ordre de sélection des caractéristiques

L'ordre de sélection est, de façon similaire à ce que nous avons obtenu à la section 4.2, assez consistant, i.e. que chaque caractéristique tend à être sélectionnée approximativement aux alentours de la même position pour chaque test. Les caractéristiques qui tiennent compte des relations spatiales entre les pixels continuent de donner de bons résultats. Le tableau 4.5 montre la liste des caractéristiques de texture qui ont, pour chaque classe et globalement, été sélectionnées parmi les 13 "meilleures" caractéristiques. Il paraît raisonnable de choisir le seuil égal à 13 pour obtenir environ  $d = 5$  caractéristiques par classe comme à la section 4.2 et en accord avec les résultats du tableau 4.6. On note qu'il n'y a que 2 caractéristiques de texture, parmi les 25 que nous avons choisies, qui se démarquent pour la caractérisation des imagerie appartenant à la classe "Herbe". Nous avons aussi répété les résultats du tableau 4.3 de la section 4.2 pour démontrer la cohérence des résultats. Finalement, on remarque que certaines caractéristiques sont souvent sélectionnées. Ce sont probablement les caractéristiques de texture les plus discriminantes.

### 4.2.3 Étiquetage

Soit  $S$  l'ensemble des 456 imagerie étiquetées disponibles pour tester la performance des caractéristiques sélectionnées et soit  $A_i, i = 1, \dots, 4$  les quatre sous-ensembles mutuellement disjoints présentés à la section 4.2.1. Quatre tests sont effectués avec  $A_i, i = 1, \dots, 4$ , pour l'apprentissage, et  $S \cap \bar{A}_i, i = 1, \dots, 4$ , respectivement, pour la généralisation. Les classificateurs utilisés sont l'algorithme des  $K$  plus proches voisins avec  $K = 1$  et  $K = 3$ , un réseau de neurones RBF et deux modèles de mixture de réseaux de neurones RBF. Les figures 4.7, 4.8 et 4.9 présentent ces trois derniers classificateurs.

La figure 4.7 montre la première architecture de réseau de neurones : un réseau de neu-

Caractéristiques sélectionnées	Ciel	Eau	Arbre	Mur	Herbe	Tous	Brodatz
3. Mode				•			
4. Variance	•						
5. Écart interquantile		•					
7. Kurtosis	•						
8. Énergie							•
9. Entropie			•	•			
10. Contraste	•					•	
11. Homogénéité locale	•	•	•	•	•	•	•
12. Homogénéité	•	•			•		
17. "SREM"		•	•	•		•	•
18. "LREM"		•	•				
22. Moyenne de différence			•				
24. Entropie de différence						•	•
25. Modèle autorégressif				•		•	•

Tableau 4.5 – Liste des "meilleures" caractéristiques sélectionnées pour chaque classe et globalement

rones RBF avec cinq neurones de sortie. Dans ce cas, un patron d'apprentissage est un vecteur composé de caractéristiques de texture sélectionnées globalement ainsi que la classe à laquelle le patron appartient. Chaque neurone de sortie représente une classe. Lors de la phase d'apprentissage supervisée des poids de la seconde couche, les valeurs désirées  $d_i$  du  $i$ -ième neurone de sortie est 1 si la classe du patron d'apprentissage appartient à la classe représentée par le neurone de sortie et 0, sinon. Pour la classification, la classe auquel correspond le neurone de sortie avec l'activation maximale est affectée au patron à classifier. Par conséquent, le pourcentage de rejet est toujours nul.

La figure 4.8 montre la deuxième architecture de réseaux de neurones : cinq réseaux de neurones RBF avec un neurone de sortie chacun. Dans ce cas, chaque réseau de neurone représente une classe. Un patron d'apprentissage est un vecteur composé de caractéris-



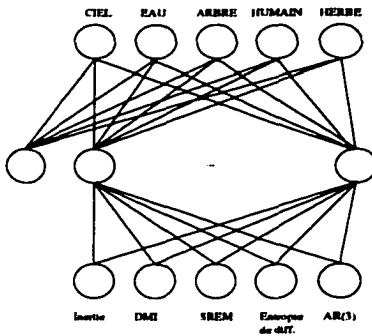


Figure 4.7 – Un réseau de neurones *RBF* avec cinq neurones de sortie

tiques de texture sélectionnées selon la classe représentée par le réseau de neurones ainsi que la classe à laquelle le patron appartient. Chaque neurone de sortie correspond à la classe représentée par le réseau de neurones. Lors de la phase d'apprentissage supervisée des poids de la seconde couche, les valeurs désirées  $d_i$  du  $i$ -ième neurone de sortie est 1 si la classe du patron d'apprentissage appartient à la classe représentée par le réseau de neurones et 0, sinon. Pour la classification, la classe auquel correspond le réseau de neurone auquel appartient le neurone de sortie avec l'activation maximale est affectée au patron à classifier. Par conséquent, le pourcentage de rejet est toujours nul.

La figure 4.9 montre la troisième architecture de réseaux de neurones: cinq réseaux de neurones *RBF* avec cinq neurones de sortie chacun. Dans ce cas, chaque réseau de neurones représente une classe. Un patron d'apprentissage est un vecteur composé de caractéristiques de texture sélectionnées selon la classe représentée par le réseau de neurones ainsi que la classe à laquelle le patron appartient. Pour chaque réseau de neurones, il y a un neurone de sortie qui correspond à chacune des classes. Lors de la phase d'apprentissage supervisée des poids de la seconde couche, les valeurs désirées  $d_i$  du  $i$ -ième neurone de sortie, pour chaque réseau, est 1 si la classe du patron d'apprentissage appartient à

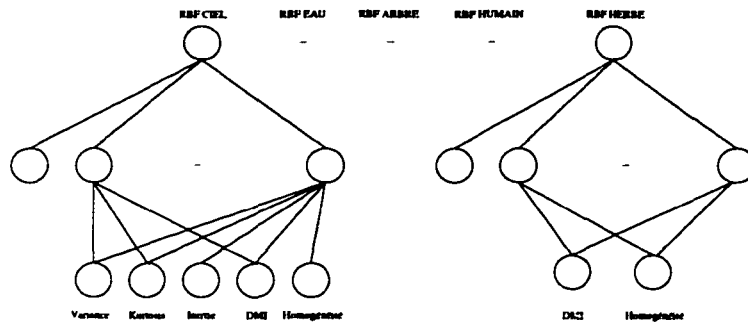


Figure 4.8 – Cinq réseaux de neurones *RBF* avec un neurone de sortie pour chacun

la classe à laquelle correspond le neurone de sortie et 0, sinon. Pour la classification, nous retenons, pour chaque réseau de neurones, la classe qui correspond au neurone de sortie avec l'activation maximale. Si une classe est retenue par la majorité des réseaux de neurones, i.e. trois fois ou plus sur cinq, alors cette classe est affectée au patron à classifier. Par conséquent, comme il est possible qu'il n'y ait pas de classe remplissant cette condition, il est possible d'avoir un pourcentage de rejet non nul.

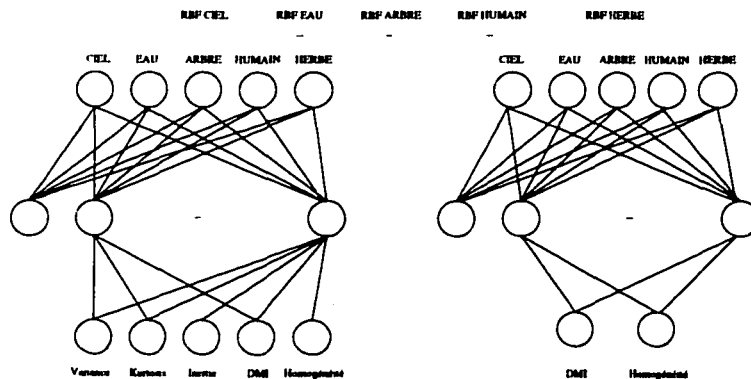


Figure 4.9 – Cinq réseaux de neurones *RBF* avec cinq neurones de sortie pour chacun

Tous les réseaux de neurones ont 20 neurones cachés.

Le tableau 4.6 montre la performance de bonne classification selon la méthode de classification. Les résultats obtenus en utilisant des réseaux de neurones RBF individuellement ou avec un modèle de mixture sont meilleurs que ceux obtenus en utilisant l'algorithme des  $K$  plus proches voisins. Dans le cas des  $K$  plus proches voisins, les résultats sont meilleurs lorsque  $K = 1$  puisqu'alors il existe toujours un voisin et par conséquent, il n'y a jamais de rejet pour la classification. Cependant, lorsque  $K = 3$ , il n'existe pas toujours une majorité de voisins appartenant à la même classe et le pourcentage de rejet est souvent non nul. Le même commentaire s'applique pour la cinquième méthode, i.e. le modèle de mixture de cinq réseaux de neurones RBF. Avec ce modèle, puisqu'il faut aussi obtenir une majorité pour pouvoir prendre une décision, le pourcentage de rejet est souvent non nul. Voilà pourquoi, en partie, les résultats de ces méthodes sont un peu moins bons.

Sous-ens. d'apprentissage	K=1 plus proche voisin	K=3 plus proches voisins	Un RBF, cinq neurones de sortie	Cinq RBF, un neurone de sortie chacun	Cinq RBF, cinq neurones de sortie chacun
$A_1$	48.5%	45.0%	50.3%	<b>56.7%</b>	50.3%
$A_2$	49.4%	43.0%	48.2%	<b>52.0%</b>	49.4%
$A_3$	50.9%	40.4%	38.0%	<b>57.0%</b>	53.2%
$A_4$	52.0%	44.4%	46.2%	<b>53.8%</b>	51.5%

Tableau 4.6 – *Performances de bonne classification*

La quatrième méthode, i.e. le modèle de mixture de réseaux de neurones avec un neurone de sortie pour chacun donnent les meilleurs résultats. Ainsi, il semble que chaque réseau de neurones représente bien sa classe. Les pourcentages de bonne classification varient entre 52.0% et 57.0%. C'est très encourageant, spécialement lorsqu'on considère que chaque patron peut être classifié dans l'une des cinq classes possibles. Dans la plupart des cas, le modèle de mixture de réseaux de neurones performe mieux que les classificateurs

individuels.

La figure 4.10 montre un échantillon des images étiquetées extraites du troisième test et de la quatrième méthode. Comme on peut le voir dans cet exemple, la texture permet de bien caractériser les régions des classes "Ciel", "Eau", "Arbre" et "Mur". Cependant, les régions de la classe "Herbe" sont mal représentées. Bien-sûr, cela peut s'expliquer par le fait que l'herbe a souvent une texture similaire aux arbres et aussi à l'eau et au ciel. Il est certain que si les caractéristiques de couleur et de relations spatiales entre les objets sont intégrées au système, la performance sera grandement améliorée.

#### **4.2.4 Sommaire**

Dans ce volet du travail, nous avons expérimenté l'étiquetage d'images couleur avec les caractéristiques de texture. Pour pouvoir appliquer les méthodes d'extraction de caractéristiques de texture aux différentes régions des images, nous avons développé un algorithme simple mais efficace pour segmenter une image en utilisant l'information de couleur et la variance. Il est à noter que la couleur n'est pas utilisée dans les étapes subséquentes. De cette façon, nous pouvons concentrer notre étude sur la contribution des caractéristiques de texture.

Pour sélectionner un ensemble approprié de caractéristiques de texture pour les cinq étiquettes ou classes, nous avons utilisé un algorithme de sélection basé sur les  $K$  plus proches voisins. Nous avons aussi sélectionné un ensemble de caractéristiques de texture commun à toutes les classes et aussi des ensembles de caractéristiques de texture spécifiques pour chacune des classes. Pour utiliser ces caractéristiques, trois types de classificateurs ont été construits pour étiqueter les régions. Ces classificateurs sont basés

sur l'algorithme des  $K$  plus proches voisins, le réseau de neurones RBF et le modèle de mixture de réseaux de neurones RBF. Les classificateurs basés sur un modèle de mixture donne de meilleurs résultats que les classificateurs individuels. Le modèle de mixture de réseaux de neurones RBF avec un seul neurone de sortie chacun donne les meilleurs résultats. De plus, les résultats sont consistants avec la théorie de la "malédiction des grandes dimensions".

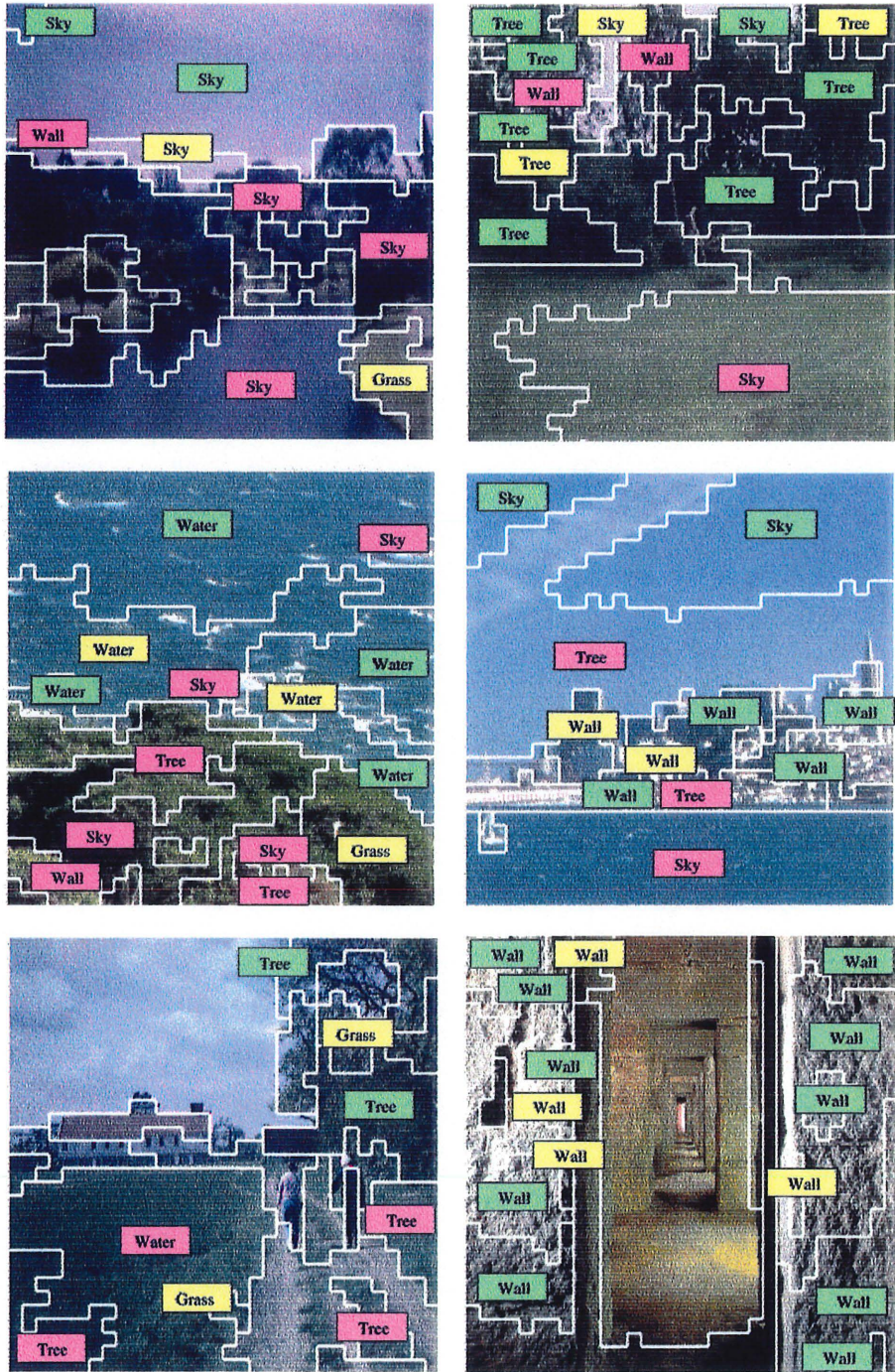


Figure 4.10 – Échantillon des images étiquetées

# CONCLUSION

Dans ce projet, les objectifs étaient, d'abord, d'expérimenter des méthodes d'extraction de caractéristiques de texture et une méthode de sélection non-paramétrique séquentielle vers l'avant sur des images à niveaux de gris de texture de Brodatz. Nous avons pu conclure, à cause d'observations à l'égard du phénomène de la "malédiction des grandes dimensions", que la sélection des caractéristiques est de mise. Aussi, les caractéristiques de texture sélectionnées permettent d'effectuer un étiquetage fort satisfaisant.

Par conséquent, nous avons tenté d'élargir cette méthode aux images couleur de scènes extérieures. Cela nous a conduit au développement d'une méthode simple mais efficace de segmentation des images couleur, ainsi qu'à une adaptation des méthodes d'extraction de caractéristiques de texture aux images couleur basée sur le changement de l'espace couleur. Nous avons également affiné la façon de sélectionner les caractéristiques en effectuant non seulement une sélection globale, mais aussi une sélection selon les classes. Ces résultats de la sélection ont finalement été validés à l'aide de trois types de classificateurs : l'algorithme des  $K$  plus proches voisins, le réseau de neurones RBF et le modèle de mixture de réseaux de neurones RBF. Le modèle de mixture de réseaux de neurones RBF permet d'obtenir des résultats très encourageants, surtout si l'on tient compte du fait que nous n'avons utilisé que la texture dans notre système.

Nous pouvons conclure que la méthode développée pour les images à niveaux de gris de texture de Brodatz est effectivement généralisable aux images couleur. Nos résultats démontrent aussi que la texture est une information importante en vision.

Dans des travaux ultérieurs, il serait bien entendu très intéressant d'incorporer à notre système des caractéristiques d'images autres que la texture, comme par exemple la couleur et la position relative des objets dans l'image. Il va sans dire que la qualité des résultats en serait certainement accrue.



# Bibliographie

- [1] N. Abbadenni, D. Ziou, and S. Wang. Recherche d'images basée sur le contenu. Représentation de la texture par le modèle autorégressif. Rapport no. 216, 1998.
- [2] D.W. Aha and R.L. Bankert. *Artificial Intelligence and Statistics*. Springer-Verlag, 1996.
- [3] E.M. Aitnouri, S. Wang, D. Ziou, J. Vaillancourt, and L. Gagnon. An Algorithm for Determination of the Number of Modes for *pdf* Estimation of Multi-Modal Histograms. *12th Conference on Vision Interface Proceedings*, pages 368–374, 1999.
- [4] H. Akaike. Information and an extension of the maximum likelihood principle. *2nd International Symposium of Information Theory*, pages 267–281, 1973.
- [5] C. Bishop. *Neural Networks for Pattern Recognition*. Oxford University Press Inc, 1995.
- [6] C. Chen, L. Pau, and P. Wang. *Handbook of Pattern Recognition and Computer Vision*. World Scientific, 1993.
- [7] G. Daigle. Système pour l'étiquetage automatique de l'image. Master Thesis, Département de mathématiques et d'informatique, Université de Sherbrooke, 1999.
- [8] G. Daigle, S. Wang, , D. Ziou, and B. Ayeb. A System for Picture Labeling. *IEEE Proceedings of International Conf. on Systems, Man, and Cybernetics*, pages 4453–4458, 1998.

- [9] M. Dash and H. Liu. Feature Selection for Classification. *Intelligent Data Analysis*, Vol. 1, no. 3, <http://www.elsevier.com/locate/ida>, 1997.
- [10] J. Doak. An Evaluation of feature selection methods and their application to computer security. Technical Report CSE-92-18, University of California, Department of Computer Science, Davis CA, December 1992.
- [11] R.C. Gonzalez et R.E. Woods. *Digital Image Processing*. Addison-Wesley Publishing Company, 1992.
- [12] L. Gagnon and R. Klepko. Hierarchical Classifier Design for Airborne SAR Images of Ships. *Automatic Target Recognition VIII, SPIE Aerosense '98*, 3371:38–49, April 1998.
- [13] M. Galloway. Texture Analysis Using Gray Level Run Lengths. *Computer Graphics and Image Processing*, pages 172–179, 1975.
- [14] S. Mori H. Tamura and T. Yamawaki. Textural Features Corresponding to Visual Perception. *IEEE Transactions on Systems, Man and Cybernetics*, pages 460–473, 1978.
- [15] R. Haralick, K. Shanmugam, and I. Dinstein. Textural Features for Image Classification. *Systems, Man and Cybernetics*, pages 610–621, 1973.
- [16] R. Haralick and L. Shapiro. *Computer and Robot Vision*. Addison-Wesley Publishing Company, 1992.
- [17] J. Gerbrands I. Young and L. Van Vliet. Image Processing Fundamentals. <http://www.ph.tn.tudelft.nl/Courses/FIP/noframes/fip-Statisti.html>.
- [18] G. John, R. Kohavi, and K. Pfleger. Irrelevant features and the subset selection problem. *Proceedings of the Eleventh International Machine Learning Conference*, pages 121–129, 1994.
- [19] M.I. Jordan and R.A. Jacobs. Hierarchical Mixtures of Experts and the EM Algorithm. *Neural Computation*, 6(2):181–214, 1994.

- [20] G. Joy and Z. Xiang. Color Image Quantization by Agglomerative Clustering. *IEEE Computer Graphics and Applications*, pages 14(3):44–48, 1994.
- [21] M.F. Auclair Fortier and F. Deschenes and S. Lawrence. NVLib 1.0: Librairie de fonctions C++ pour la programmation en vision et en traitement d'images. Rapport technique, Département de mathématiques et d'informatique, Université de Sherbrooke, Canada, 2000.
- [22] R.W. Picard and T.P. Minka. Vision Texture for Annotation. Technical Report 302, MIT Media Laboratory, 1995.
- [23] Charles A. Poynton. Frequently Asked Questions about Color. <http://www.inforamp.net/poynton/ColorFAQ.html>, 1995.
- [24] J. Mac Queen. Some Methods for Classification and Analysis of Multivariate Observation. *Proc. of the 5th Berkeley Symp. on Math. Stat. and Prob.*, pages 281–297, 1967.
- [25] P.K. Robertson and J.F. O'Collaghan. The Generation of Color Sequences for Univariate and Bivariate Mapping. *IEEE Computer Graphics and Applications*, 6(2):24–32, 1986.
- [26] A. Rosenfeld and A. Kak. *Digital Picture Processing*. Academic Press, 1982.
- [27] M. Stone. Cross-Validatory Choice and Assessment of Statistical Predictions. *Journal of the Royal Statistical Society B*, 36(1):131–147, 1974.
- [28] A. Trouvé. *La mesure de la couleur. Principes, technique et produits du marché*. Afnor, 1991.
- [29] M. Turcotte, S. Wang, and J. Vaillancourt. Extraction et sélection de caractéristiques de texture pour la classification d'images. Rapport technique no. 231, Département de mathématiques et d'informatique, Université de Sherbrooke, Sherbrooke, Canada, 1999.

- [30] A.M. Ukrainec and S. Haykin. A Modular Neural Network for Enhancement of Cross-polar Radar Targets. *Neural Networks*, 9(1):143–168, 1996.
- [31] J. Weszka, C. Dyer, and A. Rosenfeld. A Comparative Study of Texture Measures for Terrain Classification. *Systems, Man and Cybernetics*, pages 269–285, 1976.
- [32] J. Zhang and J.W. Modestino. A Model-Fitting Approach to Cluster Validation with Application to Stochastic Model-Based Image Segmentation. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 12(10):1009–1017, 1990.