

# RÔLE DE LA CLASSIFICATION STATISTIQUE DANS LA COMPRESSION DU SIGNAL D'IMAGE : PANORAMA ET UNE ÉTUDE SPÉCIFIQUE DE CAS <sup>1</sup>.

Nadia Ghazzali\*, Alain Léger\*\*, Israël César Lerman\*\*\*

\* Université de Laval, Québec, Canada. \*\* CCETT, 4 rue du clos Courtel, F-35512 Cesson Sévigné. \*\*\* IRISA, campus universitaire de Beaulieu, F-35042 Rennes Cédex.

## RÉSUMÉ

La compression des images à base de quantification vectorielle ("QV") constitue toujours un champ actif de recherches. Même si cette technique employée seule ne fournit pas, aujourd'hui, des taux de compression suffisants eu égard aux besoins des applications et à "ses soeurs concurrentes" (méthodes par transformées, méthodes prédictives, méthodes affines, etc. ...), elle n'en reste pas moins qu'associée à d'autres techniques, elle apporte des compléments de compression non négligeables - démontrés théoriquement -, et de plus, présente des qualités de simplicité - au décodeur - fort appréciées.

Nous commençons par une présentation générale du domaine de la compression image en y montrant comment la QV s'y inscrit; et, la situation de cette approche vis à vis d'autres, plus traditionnelles (transformées en "cosinus"), la QV peut directement être obtenue par une méthode de classification automatique dont il y a lieu d'adapter les paramètres. On entre ainsi dans le corps des méthodes de l'analyse classificatoire des données et nous rapportons une analyse expérimentale originale, présentant divers aspects méthodologiques - et couvrant deux thèses - sur l'impact comparé de diverses méthodes de classification; où la méthode de l'Analyse de la Vraisemblance du Lien (AVL) est considérée avec un accent particulier.

## SOMMAIRE

### 1 - INTRODUCTION

### 2 - ASPECTS GÉNÉRAUX

#### 2.1 - L'information et son codage

##### 2.1.1 - Introduction

##### 2.1.2 - Résultats de Shannon

##### 2.1.3 - Codage à réduction de débit de l'information

#### 2.2 - La compression du signal d'image

##### 2.2.1 - le signal d'image et sa représentation

##### 2.2.2 - le codage à réduction de débit du signal d'image

##### 2.2.3 - la restitution: Décompression et Visualisation

#### 2.3 - Normes et Standards : la norme "JPEG" de compression des images fixes

### 3 - LA QUANTIFICATION VECTORIELLE DES IMAGES FIXES

#### 3.1 - Motivations

#### 3.2 - Exposé des principes

#### 3.3 - Image en niveaux de gris; Traitements et résultats

##### 3.3.1 - Quantification scalaire

##### 3.3.2 - Quantification vectorielle d'une image en niveau de gris

##### 3.3.3 - Résultats

#### 3.4 - Image en couleur; Traitements et résultats

##### 3.4.1 - La donnée de base

##### 3.4.2 - Méthodes et résultats

### 4 - CONCLUSIONS

### 5 - RÉFÉRENCES

---

<sup>1</sup> Ce travail a donné lieu à un exposé aux "Journées MODULAD, Applications Industrielles de l'analyse des données, Lannion 17-18 juin 1993.

## 1- INTRODUCTION<sup>2</sup>

La compression des images s'inscrit dans le large domaine des techniques de traitement de l'information et plus spécifiquement du "signal" . Elle répond à un besoin sûrement très ancien de transmission d'images (tableaux de souverains ou de notables que l'on faisait porter par coursier parfois sur de longues distances!), mais ici, nous le situons dans le prolongement d'événements marquant indiscutablement cette fin du XIX ième -photographie (1826), télégraphe (1837), téléphone (1876) et T.S.F. (1896) - avec l'invention du Bélinographe qui, au tout début du XXIème (1905) révolutionne les communications de presse: 12 minutes d'attente par image entre Paris et Berlin, au lieu de 23 heures!

Aujourd'hui, la disponibilité d'outils théoriques et la maturité de la technologie numérique, d'une part; l'essor de la micro-informatique proposant des puissances de traitement tout à fait étonnantes (des dizaines de millions d'opérations par seconde), l'accroissement des ressources de stockage magnétique et optique (600 Moctets ou plus) et l'offre de réseaux publiques de télécommunication à débits élevés (de 64 Kbit/s à 2 Mbit/s, voire 150 Mbit/s) de transfert d'information, d'autre part, rendent réaliste le développement au profit d'un large public, de services de télécommunication ou d'applications informatiques "temps réel", où l'image joue un rôle central. On peut citer pour exemples, des applications de bibliothèques ou de musées, où la vidéo se substitue, en partie, au papier, (médiathèque du centre Beaubourg, bibliothèque du congrès aux États Unis, musées européens (projet européen NARCISSE), ...), des applications d'enseignement à distance (projet européen DELTA), des services de Presse pour lesquels, sauf rares exceptions, l'image reste un élément essentiel, des applications d'imagerie satellitaire pour l'analyse des couvertures nuageuses ou des ressources géologiques, les exemples sont très variés, mais ont tous au moins un élément commun, la compression de l'information d'image en vue de son traitement, mais surtout de son transport, ou de son stockage.

Les organismes de normalisation ISO (Organisme International de Normalisation) et CCITT (Comité Consultatif International du Télégraphe et du Téléphone) pour la période 1984-1992, ainsi qu'un projet collaboratif Européen ESPRIT (N°563 ou "Photographic Image Coding Algorithm" (1985-1988), ont fédéré les études de codage à réduction de débit de l'image en vue d'arrêter un schéma unique reconnu partout, tous services confondus. Ceci a donné naissance à l'algorithme "JPEG" ("Joint Photographic Expert

---

<sup>2</sup> Une partie de cet article est emprunté à l'article d'Alain Léger, "La compression des données images" paru dans la collection "La science au présent", Encyclopédie Universalis, 1993.

Group" de l'ISO JTC1/SC18 "Traitement de l'information"). Enfin, en rapport avec l'exposé qui suit ces structures ont fourni un cadre très dynamique aux études comparatives d'algorithmes de codage, nommément ici, la technique à base de transformée en cosinus et la technique par quantification vectorielle.

Un panorama restreint de l'état de l'art, tant théorique que technique, d'une discipline en pleine évolution, et non "achevée" à ce jour, est proposé ci-après en vue de poser des éléments marquants du domaine, pour ensuite exposer, "en contrechamp", le résumé du travail de recherche de deux thèses [14], [9] menées par le CCETT et l'IRISA, où la première s'inscrivait dans le projet ESPRIT "PICA" (Photographic Image Compression Algorithm").

## 2 - ASPECTS GÉNÉRAUX

### 2.1 - L'information et son codage

#### 2.1.1 - Introduction

La théorie de l'information et du codage est associée à un processus de communication dont l'objet est l'acheminement de messages d'un expéditeur vers un destinataire, au travers d'un canal de communication (fig.1).

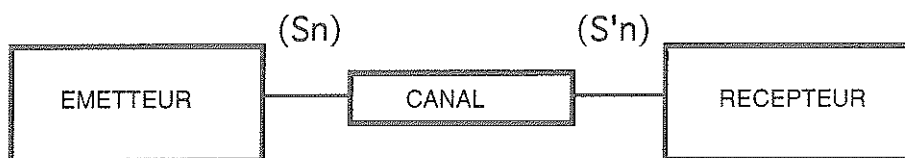


Figure 1: modèle général de transmission de messages

Ce modèle est bien sûr très général, et peut porter aussi bien sur l'envoi de courrier postal - transport humain de lettres, télégrammes ou colis - que sur l'envoi de messages radioélectriques - signaux télévisuels ou fac-similé - Le contenu véhiculé par le message, est perçu du récepteur, comme un événement incertain. C'est ainsi que la théorie de l'information est une théorie statistique. Elle introduit la notion fondamentale de "quantité d'information" portée par le message du récepteur vers l'émetteur. Ces éléments donneront corps au premier théorème de Shannon dit "théorème du codage de source".

De plus, l'information devant être très naturellement portée à distance, on est amené à considérer "la capacité" et "la fiabilité" du canal de transmission, c'est-à-dire son aptitude à acheminer plus ou moins rapidement (ex. débit d'un réseau de télécommunication) et

dans des conditions plus ou moins sûres (ex. taux d'erreurs d'un réseau de télécommunication). D'où la constatation, à l'arrivée, d'une dégradation du message expédié. Le souci de maintenir une intelligibilité suffisante du message transmis se trouve formalisé dans le deuxième théorème fondamental de Shannon, dit "codage de canal bruité", qui assure qu'il est possible de transmettre un message à une cadence finie sur un canal perturbé en maintenant une protection du message aussi grande que l'on veut.

Ces notions sont fondamentales pour l'étude de la représentation codée de l'information ou *codage*, on distingue deux type de codage:

- *le codage de source* dont l'objectif est la réduction de débit de la source, par élimination de toute sorte de redondance, au prix, éventuellement, d'une perte admissible d'information. On parlera, alors, de codage sans perte et de codage avec perte d'information sous contraintes de fidélité. (rem: le cryptage de l'information est une technique de codage de source dont l'objectif est d'assurer le secret du message transmis)

- *le codage de canal* dont l'objectif est de sécuriser l'information transmise sur un canal qui peut-être bruité ; il réintroduit, de façon apparemment contradictoire, une redondance maîtrisée, qui selon les cas permet de détecter ou de corriger les erreurs de transmission.

### 2.1.2 - Résultats de Shannon

La notion d'entropie prend son origine dans le besoin général de caractériser le degré de désordre (ex. thermodynamique) ou d'incertitude dans la réalisation d'une variable aléatoire. Dans un schéma de communication, cette notion va être utile à caractériser la *quantité moyenne d'information* fournie par une source de symboles qui vu du récepteur présente un caractère aléatoire. A la suite de nombreux considérants, elle a été définie par l'expression générale:

$$H = - \sum_i p_i \log_2 p_i \quad (\text{l'unité est le shannon (Sh) ou bit d'information, par symbole})$$

La borne minimale théorique de l'entropie d'une source d'information en vue de sa représentation codée - sans perte d'information - est formulée par le premier théorème de Shannon:

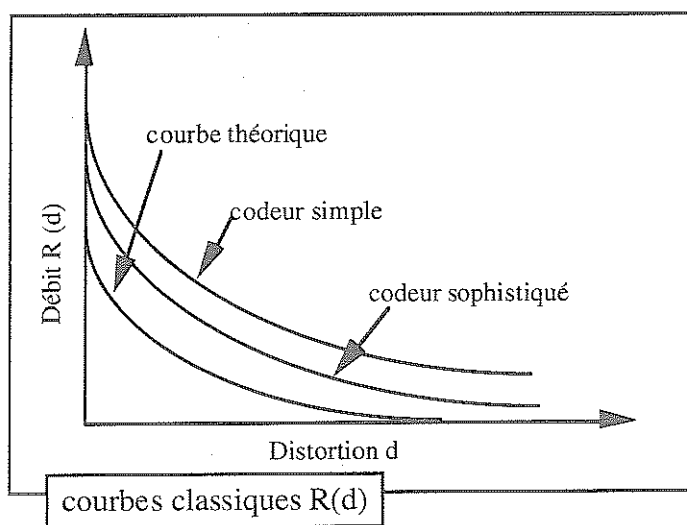
*Si H est l'entropie d'une source discrète d'information, la suite des signes générés par cette source ne peut-être représentée exactement par une suite de symboles binaires*

utilisant en moyenne moins de  $H$  bits par signe. Par contre, on peut théoriquement se rapprocher de la valeur  $H$  autant qu'on le veut.

L'aptitude d'un canal bruité à transmettre des messages dans des limites de qualité bornée ou l'aptitude d'un schéma de codage à réduction de débit avec perte d'information, sous contrainte de qualité bornée, s'exprime dans le second théorème de Shannon:

*Si l'on transmet sur un canal bruité de capacité  $C$  une information à une cadence  $R$ , il existe en général une possibilité de codage permettant de réduire la probabilité d'erreur de transmission à une valeur arbitrairement faible, tant que la cadence  $R$  ne dépasse pas la capacité  $C$  du canal.*

Ce théorème a conduit à l'élaboration d'une théorie classique ("Rate-Distortion theory") [cf. Références aux ouvrages] qui fournit une courbe utile à l'expression d'une limite théorique ( courbe " $R(d)$ " ) qui pour une distorsion donnée " $d$ "- aussi petite soit-elle - fixe le débit minimum de source (" $R$ " pour "Rate" en anglais) (ex. nombre de bits par symbole de source, pour un codage de source) qu' on peut atteindre tout en garantissant la distorsion " $d$ ".



### 2.1.3 - Codage à réduction de débit de l'information

- Compression sans perte d'information

Shannon, dans son premier théorème, définit une limite théorique à l'élimination de redondance statistique d'une source d'information, mais aucune méthode pour l'atteindre n'est décrite. C'est ainsi que des algorithmes astucieux de codage - on parlera souvent de

"codage entropique" car cherchant à approcher cette valeur d'entropie limite - ont vu le jour. Les plus connus sont le codage de Huffman [13], le codage arithmétique [ voir Références aux revues] et le codage Lempel-Ziv. [30] Ces méthodes sont aussi appelées "réversibles", car elles assurent après décodage une restitution identique à l'original de l'information de source.

• Compression sous contraintes de fidélité

Cette compression a pour objectif de réduire le débit d'information de la source dans des proportions en général très larges, et donc de dépasser la limite de Shannon (sans perte). De très nombreuses techniques existent, car contrairement à la compression sans perte, ces méthodes mettent en scène des représentations de signaux (audio, vidéo, ) et des représentations de caractérisation de la fidélité, qui sauf domaine très spécifique (robotique, radar, ...) , résistent totalement - ou presque -aujourd'hui, à toute modélisation. Il est toutefois vrai que le 2<sup>ième</sup> théorème de Shannon a engendré de nombreuses études théoriques qui fournissent des réponses dans des cas bien limités (e.g. situation de source Gaussienne et contrainte de fidélité selon le critère de l'erreur quadratique moyenne), mais malheureusement la réalité en compression du signal d'image "naturelle" est bien souvent différente! Notre situation se schématise donc ainsi (fig.2):

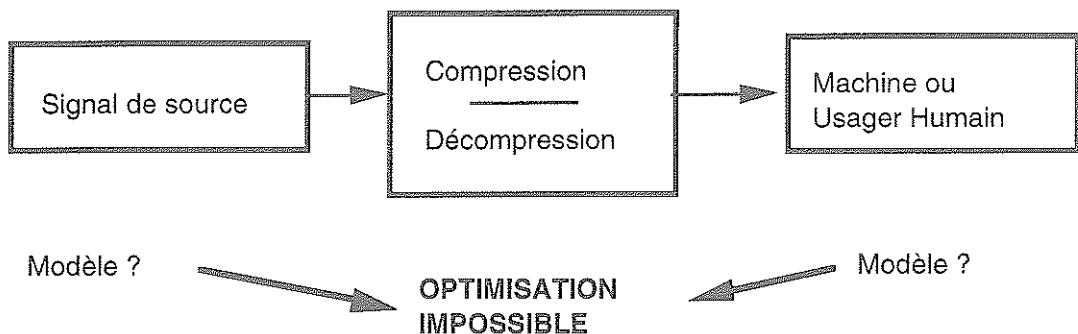


Figure 2: problématique de la compression sous contraintes de fidélité

Cette limitation de nos connaissances engendre depuis de nombreuses années une abondante littérature scientifique sur le thème de la compression, et reflète bien une situation où malgré l'usage de fondements théoriques fiables - l'intuition joue encore un rôle évident. Toutefois, pour comprimer l'information de source au-delà de la limite de Shannon, nous sommes amenés à mettre en oeuvre une procédure non linéaire (d'où la non réversibilité de la compression) dite "de quantification" assujettie à une contrainte "de qualité" de l'information reconstituée. L'art va consister à éliminer cette fois, non plus la redondance statistique (Shannon premier théorème), mais la redondance de non-

pertinence liée à la sémantique du message transmis et à son contexte applicatif, tous deux fort dépendants de l'expertise disponible sur le domaine, et conjuguant des approches à la fois pragmatique et théorique, ex. non-pertinence psychoacoustique pour la compression d'un son, non-pertinence psychovisuelle pour la compression d'une image, etc.....(voir encadrés "Quantification" et "critères de qualité").

## 2.2 - La compression du signal d'image

### 2.2.1 - le signal d'image et sa représentation

En traitement du signal d'image, le point de départ est toujours l'observation et l'acquisition de mesures délivrées par un ou plusieurs capteurs (caméra vidéo, télescope, scanner médical, ...). Dans le contexte qui nous intéresse, ces signaux sont présentés sous forme numérique ( suite de nombres représentant la grandeur mesurée) avant tout traitement, et constituent "l'information brute". Le soin que l'on apporte à cette étape est capital pour la suite, puisque toute information perdue à ce niveau ne pourra jamais être restituée, par exemple, "la finesse d'analyse" (ou échantillonnage), le bruit de quantification et le bruit du capteur sont des éléments à définir avec rigueur.

### 2.2.2 - le codage à réduction de débit du signal d'image

#### • Modèle de codeur

Il est d'usage de modéliser un codeur sous la forme de trois étapes , conformément à la figure 3, où les variables sont exprimées sous forme vectorielle.

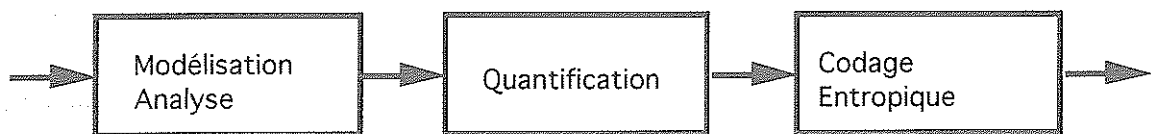


Figure 3: modèle général pour le codage à réduction de débit

#### • Modélisation et analyse

La première étape dite de "modélisation et d'analyse" est très générale, et est utile à bien d'autres domaines que celui de l'image (le son par exemple). C'est l'étape d'analyse ou

de modélisation au cours de laquelle les caractéristiques du signal d'image, *pertinentes* à l'utilisateur sont mises en évidence. La représentation sera soit non paramétrique: analyse spectrale, analyses temps-fréquence et temps échelle, ..., soit paramétrique via la définition d'un modèle qui prendra en compte l'information disponible sur le signal (le signal émis par un radar ou un sonar ou un échographe est bien sûr connu à priori, et donc modélisé, avant le traitement du signal d'écho, en revanche les signaux d'image de type photographie naturelle résistent toujours aux nombreuses tentatives de modélisation (la synthèse d'image naturelle par géométrie fractale en propose toutefois une approximation). Enfin il est important de noter que l'étape d'analyse n'apporte *aucune* compression mais simplement *une autre représentation de cette information*.

Les méthodes classiques d'analyse pour l'image sont: Fourier, Karhunen-Loewe, Cosinus, ou méthodes Différentielles, ou analyse par plages ( utilisée pour les services de télécopie).

- La Compression: Quantification et Codage entropique

Ces deux étapes constituent les étapes de compression.

1- La *Quantification* - étape centrale - apporte une compression avec perte maîtrisée de l'information, au moyen d'un critère de qualité. Ce dernier - de nature quantitative pour le calcul - ne capte que très imparfaitement les aspects psychophysiques de la vision humaine. De plus, tous les défauts visibles ou non, sur l'image restituée ne proviennent généralement que de ce traitement. C'est aussi ce traitement qui par le choix du niveau de qualité assure le réglage du débit de sortie du codeur - en limite, par un choix judicieux du pas de quantification, cette étape est neutre et la qualité de l'image restituée est parfaite. Enfin, pour l'image fixe, ce réglage permet de calibrer la taille des fichiers d'images numériques comprimées; pour l'image animée, il permet d'ajuster dynamiquement le débit du codeur en fonction de contraintes de débit du canal de transmission (réseau de transmission, débit de sortie d'un lecteur de Disque Compact,...). Voir encadré "Quantification et critère de qualité".

2- Le *codage entropique* apporte une compression sans perte d'information. Les méthodes les plus classiques sont le codage de Huffman [13], le codage arithmétique [voir Références aux revues] et le codage de "Lempel-Ziv" [30].



### 2.2.3 - La restitution: Décompression et Visualisation

La décompression est l'opération "inverse" (ce qui n'est généralement pas le cas au sens mathématique) de l'opération de compression. Pour des applications de communication dite "de face à face"(e.g. visiophonie) une répartition équilibrée de la complexité de réalisation doit-être trouvée entre le compresseur et le décompresseur. Mais ceci n'est pas nécessairement le cas pour les systèmes de consultation d'information (e.g. videotex), où la vitesse de retrait et de visualisation des images, ainsi que le coût des terminaux sont essentiels. On trouvera donc des situations où les codeurs sont beaucoup plus complexes que les décodeurs (ex. technique par fractales, technique de Karhunen-Loewe, ...). En général c'est l'étape "Analyse et Modélisation" qui apporte cette dissymétrie, car plus délicate et nécessitant des temps de traitement plus longs.

La visualisation est l'étape de présentation de l'information décomprimée. Elle utilise un support vidéo ou un support papier pour les situations humaines usuelles. Mais l'utilisateur final peut très bien être une machine - en robotique par exemple -.

### 2.3 - Normes et Standards: la norme "JPEG" de compression des images fixes

On distingue généralement les normes (dite "de jure") issues de travaux collaboratifs d'instances internationales de normalisation comme l'ISO (Organisation International de Normalisation), le CCITT (Comité Consultatif International duTélégraphe et du Téléphone), ou le CCIR (Comité Consultatif International des Radiocommunications), des standards (ou normes "de facto") élaborés par des associations d'industriels ou d'usagers , comme l'ECMA (association européenne des fabricants d'ordinateurs). A titre d'illustration des techniques de compression des données, quelques techniques normalisées et largement utilisées, sont présentées ci-après.

La norme JPEG (IS 10918) est le fruit d'une activité de normalisation "anticipée". La compression des images a pris un réel essor dès l'apparition des techniques numériques, c'est à dire vers les années 60. Toutefois, il faut attendre le début des années 80, pour qu'apparaisse un souci de travaux collaboratifs sur ce thème. C'est le Vidéotex qui le premier fédère les études de compression des images fixes (1978), en vue d'offrir aux

usagers le complément de l'image à la présentation textuelle ou graphique ( système "CAPTAIN" Japonais et Photovideotex Européen). C'est sur cette base, qu'un peu plus tard, est lancé le projet ESPRIT-PICA (Photographic Image Compression Algorithm) en 1985, rejoint par l'ISO/IEC JTC1/SC29/WG8 (codage des images et du son). En 1986, vue l'importance du sujet pour les applications de télécommunications, le CCITT Commission VIII (protocoles et terminaux pour l'audiographie, la télécopie et le vidéotex), se joint au groupe ISO. Ainsi, dès 1985, différents algorithmes à réduction de débit allaient s'affronter, en vue de l'élaboration d'une norme de compression pour les images fixes.

Cette norme de compression des images fixes à niveaux de gris ou en couleur a été définie dans la perspective d'applications les plus diversifiées, en tenant cependant compte de contraintes réalistes par rapport aux applications les plus visibles. Ceci a conduit à imposer les fonctionnalités suivantes:

- 1) Compression sans perte et compression avec pertes.
- 2) Affichage en pleine résolution ou à résolution progressive des images décompressées.
- 3) Mise en oeuvre en logiciel et sur silicium.
- 4) Codage et Décodage en "temps réel".
- 5) Contrôle souple de la contrainte de qualité des images comprimées.

Ce qui correspond, pour un cas typique d'imagerie (image télévisuelle de qualité "studio numérique"), aux couples suivants de valeurs (compression, qualité):

- i) compression de 200: "image reconnaissable"
- ii) compression de 64: "image de qualité moyenne"
- iii) compression de 20: "image de qualité excellente"
- iv) compression de 8: "image visuellement identique à l'originale"
- v) compression de 2 à 3: "image identique à l'originale" (compression sans pertes).

La technique se schématise selon les blocs fonctionnels décrits sur la figure 4:

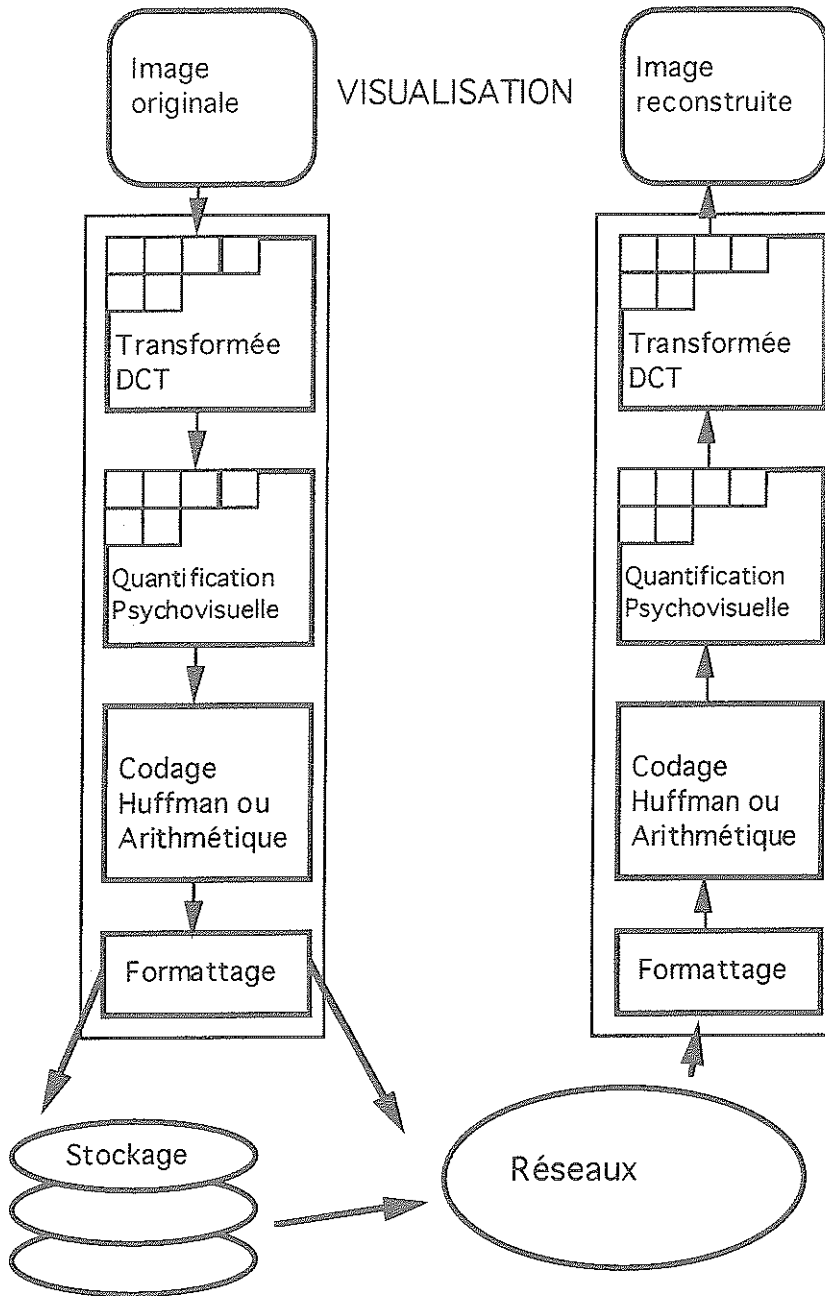


Figure 4: Schéma de codage et de décodage de la norme "JPEG"

Les trois composantes communes à tous les modes de codage basé DCT, sont largement illustrées dans la norme (International Standard 10918), toutefois les traitements de base appliqués aux échantillons d'un bloc 8x8 d'une composante d'image, sont ici rappelés, en termes simples.

•Transformation

Une Transformée en Cosinus Discrète de format 8x8 ("DCT 8x8") a été choisie comme base d'analyse des blocs d'images contigus de 8x8 pixels. Ce choix a tenu compte d'une

part, des très bonnes propriétés de condensation de l'information image à niveaux de gris et de couleurs et d'autre part de la complexité raisonnable de mise en oeuvre.

- Quantification

Chacun des 64 coefficients DCT - résultats de l'opération de transformation linéaire "DCT" du bloc 8x8 de pixels d'une composante de l'image d'origine - sont ensuite quantifiés selon une loi linéaire, et selon un pas de quantification dit "psychovisuel". *C'est là le point-clé de l'algorithme de compression.* En effet, les défauts apparaissant, éventuellement, lors de la reconstruction de l'image comprimée, ont pour origine ce "bruit de quantification". C'est ainsi, qu'un soin tout particulier doit-être apporté à la définition de ces pas élémentaires de quantification. La méthode retenue, bien qu'imparfaite, met en oeuvre des expérimentations d'analyse des seuils de perception de la vision humaine, dans des conditions d'expérience bien définie, pour déterminer la valeur de ces seuils. Quelques exemples sont donnés dans le standard, mais c'est à chaque application de définir le jeu de seuils adaptés à son contexte d'usage.

- Codage Entropique

Deux techniques sont proposées dans le standard: la méthode classique de Huffman (imposée) et le codage arithmétique (optionel). Ces deux techniques ont chacune leur mérite, l'une apporte la simplicité, l'autre apporte une adaptation plus fine au contexte locale de l'image et de ce fait un petit gain en compression (max 10%), au prix d'une complexité accrue.

### 3 - LA QUANTIFICATION VECTORIELLE DES IMAGES FIXES

#### 3 -1 Motivations

On les trouve dans l'expression des deux théorèmes de Shannon qui pour le premier appliqué à une source de symboles multi-dimensionnels (vs mono-dimensionnel) - à composantes corrélées ou non - nous montre que l'entropie de la source diminue et que donc le débit moyen par composante sera plus faible, et qui pour le deuxième nous donne une courbe  $R(d)$  plus "avantageuse", pour d'une part les raisons évoquées pour le 1<sup>er</sup> théorème, mais aussi par le fait que le critère de distorsion "d" est aussi multi-dimensionnel et reflète donc mieux "la nature des choses" qui semble être plus multidimensionnelle que monodimensionnelle.

Dans la situation qui nous intéresse, ici, le critère "d" qui est une mesure de la qualité (ou de la dégradation) de l'image décomprimée, on comprendra aisément que cette

appréciation subjective est "très globale". Ceci ne veut pas dire pour autant, que le critère trouve, du coup, une réponse analytique définitive! On fait juste un pas en avant.

### 3-2 Exposé des principes

La donnée d'une image numérisée est représentée par un tableau  $T$  de nombres suivants:

$$T = \{ ((i, j), t_{ij}) / 0 \leq i \leq N_c - 1, 0 \leq j \leq N_l - 1 \} \quad (1)$$

$N_c$  (resp.  $N_l$ ) est le nombre de colonnes (lignes) de  $T$

$(i, j)$  est un point courant de ce tableau, que l'on appelle communément "pixel" ou "pel" (de l'anglais "picture element" et du français "point image").  $T$  représente une application  $t$  de l'ensemble - rectangle des points images dans une échelle uni ou multidimensionnelle. Le cas multidimensionnel correspond au produit direct d'échelles unidimensionnelles. Ces dernières ont a priori un caractère totalement ordinal. Mais, on s'empresse pour des raisons de commodité - de les plonger dans des échelles numériques. Ainsi,  $R_+$  désignant l'ensemble des nombres réels positifs,  $t$  est une application de la forme :

$$t: \{ (i, j) / 0 \leq i \leq N_c, 0 \leq j \leq N_l \} \longrightarrow R_+^p$$

$$(i, j) \longmapsto t_{ij} = t(i, j) \quad (2)$$

Si  $p = 1$ , il s'agit d'une image "monochrome". Ainsi, pour une image en niveaux de gris (cf. ci-après) la fonction  $t$  représente la Luminance et on a généralement pour  $t_{ij}$ , l'échelle des valeurs entières  $[0, 255]$ ; de la sorte,  $t_{ij}$  admet une représentation codée sur 8 bits. En s'intéressant à une chrominance particulière d'une image couleur, on a également  $p = 1$ .

Si  $p > 1$ , l'image est dite "en couleurs" ou "multi-spectrale". L'ensemble le plus populaire et le plus classique est celui où on retient pour la description d'un pixel, les trois composantes R (Rouge), V (Vert) et B (Bleu). Dans ce cas, où  $p = 3$ , on a :

$$t_{ij} = t(i, j) = ( R(i, j), V(i, j), B(i, j) ) \quad (3)$$

$t$  est ainsi une fonction vectorielle. Comme pour la luminance, R, V et B peuvent chacune se coder sur une échelle de 256 valeurs.

Le tableau  $T$  représente les données brutes de l'image ayant subi un processus de numérisation. Le tableau  $T$  est ensuite représenté - via éventuellement un codage des données brutes - par un nuage de points (au sens de l'analyse des données) dans l'espace géométrique  $R_+^q$  ou bien, dans un espace de type  $E^q$ , où l'échelle  $E$  des valeurs est plus pauvre que  $R_+$ , pour mieux respecter la nature de l'information. Si on considère les exemples évoqués ci-dessus, on peut, le plus directement, retenir un histogramme unidimensionnel (resp. tridimensionnel) si  $p = 1$  (resp.  $p = 3$ ); auquel cas,  $p = q$ . Cependant, les situations les plus variées peuvent se présenter par rapport à l'application. De sorte, qu'à partir du tableau initial  $T$ , on déduit, par un processus parfois complexe, une nouvelle représentation ou codage de l'image. On peut chercher à réduire la dimension initiale  $p$ , on peut chercher à regrouper de façon simple des pixels déterminés (e.g. blocs 8x8) ou résultant d'un algorithme géométrico-statistique, sophistiqué.

Une opération de Quantification résulte directement d'un partitionnement ou régionalisation de l'espace  $R_+^q$  de représentation. Désignons par

$$P = \{ R_1, R_2, \dots, R_i, \dots, R_k \} \quad (4)$$

une telle partition. La Quantification suppose de plus, une représentation adéquate de chacune des classes  $R_i$  par un sommet  $y_i$  de l'espace  $R_+^q$ ,  $1 \leq i \leq k$ . On adopte souvent pour  $y_i$  le centre de gravité de la classe  $R_i$ ; mais, d'autres choix intéressants peuvent également être considérés.

L'ensemble

$$A = \{ y_1, y_2, \dots, y_i, \dots, y_k \} \quad (5)$$

des représentants est ce que l'on appelle communément "alphabet de codage" ou "palette des couleurs". On pourrait dire que  $y_i$  "quantifie" l'ensemble des points contenus dans  $R_i$ ,  $1 \leq i \leq k$ ; et il importe que la partition puisse être déterminée de façon à perdre le moins possible d'"information" (voir l'encadré : "Quantification et critères de qualité").

De façon formelle, l'opération de Quantification se présente comme une application :

$$Q: R_+^q \text{ -----} \rightarrow A ; \tag{6}$$

associant à chaque sommet le  $y_i$  qui le représente.

Il s'agit de quantification scalaire si  $q = 1$  et de quantification vectorielle si  $q > 1$ .

L'ensemble "but" de l'application précédente peut de façon évidente - sans ambiguïté être remplacé par l'ensemble des étiquettes,  $\{ 1, 2, \dots, i, \dots, k \}$ ; qu'on notera  $\underline{k}$ . En fait

$Q$  est une fonction composée :

$$Q = \text{Décodeur} \circ \text{Codeur}$$

où

$$\text{Codeur} : R^q \text{ -----} \rightarrow \underline{k} \tag{7}$$

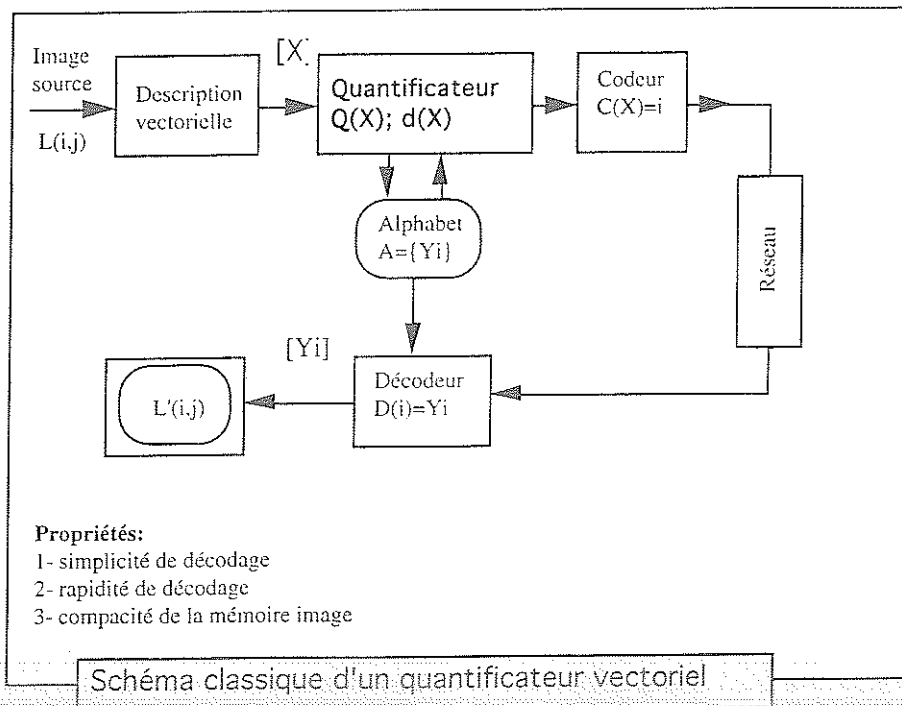
$$y \text{ |-----} \rightarrow \text{Codeur}(y) = i ,$$

si et seulement si  $y$  appartient à la région  $R_i, 1 \leq i \leq k$ .

$$\text{Décodeur: } \underline{k} \text{ -----} \rightarrow A \subset R^q \tag{7'}$$

$$i \text{ |-----} \rightarrow \text{Décodeur}(i) = y_i$$

Les réalisations pratiques sont conformes au synoptique suivant :

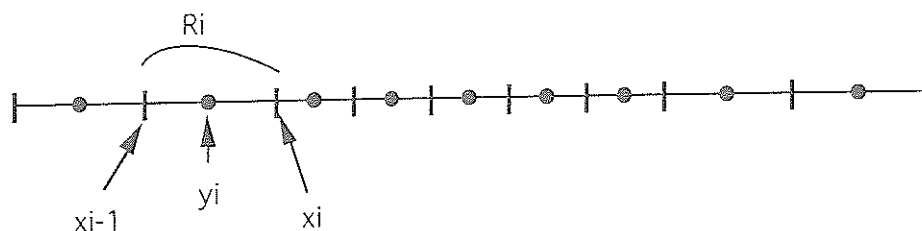


## Quantification et critère de qualité

Un quantificateur scalaire  $Q$  à  $K$ -niveaux est défini par la spécification:

- 1-une partition  $\mathcal{P}$  de l'espace 1-d (la droite) des réalisations de la variable aléatoire et scalaire  $X$ , en régions :  $(R_1, \dots, R_i, \dots, R_k)$
- 2-un ensemble  $\mathcal{Y}$  de  $N$  représentants  $(y_1, \dots, y_i, \dots, y_k)$  en correspondance bi-univoque avec l'ensemble des régions.

Ainsi, pour toute réalisation de la variable aléatoire  $x$  qui survient dans la région  $R_i$  ( $x_{i-1} < x \leq x_i$ ), le quantificateur génère la valeur unique  $y_i$ ,  $1 \leq i \leq K$ .



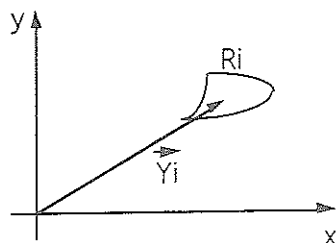
On y associe un "bruit de quantification" représentatif de l'erreur générée et dont la maîtrise est un point central de l'opération de quantification. Sa mesure s'opère sur la base de différentes métriques dont les deux suivantes sont très classiques:

$$E.Q.M. : \text{Moy} (d(x, y)) = \text{Moy} (\|x - y\|^2) \quad (\text{Erreur Quadratique Moyenne})$$

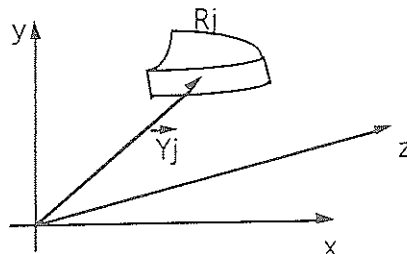
$$S/B : 10 \text{ Log} \frac{\text{Moy} \|x\|^2}{\text{Moy} (\|x - y\|^2)} \quad (\text{Rapport Signal à bruit})$$

La pratique montre que toutes ces mesures ne reflètent que faiblement notre perception des erreurs (ou artefacts) produites sur les informations restituées. Il faut donc mettre en oeuvre pour les images et les sons, une méthode complémentaire de jugement dit "subjectif" (notation sur une échelle de 0 à 5) de la "gêne" réellement perçue par l'appareil psychosensoriel de l'homme.

La Quantification scalaire présentée plus haut se généralise aisément au cas  $n$ -dimensionnel, et l'on parlera cette fois de "Quantification Vectorielle". Elle offre de nombreux intérêts, dont l'un est de considérer d'emblée la nature multidimensionnelle des phénomènes à traiter, et ainsi de mieux coller à la réalité. Des gains sur le couple (compression, qualité) sont généralement obtenus, mais au prix d'un accroissement de complexité.



Quantification 2D: le vecteur  $Y_i$  est le représentant du domaine  $R_i$  du plan



Quantification 3D: le vecteur  $Y_j$  est le représentant du domaine  $R_j$  de l'espace



Il faut noter sur cette figure la place importante du quantificateur sur lequel repose toute "l'intelligence algorithmique". On remarquera que dans notre cas le quantificateur intervient au niveau de l'image source et c'est là que s'applique pour l'essentiel les nombreuses méthodes décrites dans la littérature ("LBG", "median cut", "popularity", "median/residual", "interpolative/residual", "gain/shape", "classified", "finite state", etc. ; cf. [22], [10], [24], [26], [12]).

Les techniques citées sont présentées de façon indépendante, dans le contexte général de l'approche "traitement du signal", alors qu'en fait elles réalisent - certes avec des accents particuliers - des fonctions classiques de l'approche "Classification" en analyse des données. C'est ce dernier point de vue que nous avons développé pour quantifier les données images à compresser. Conformément à ci-dessus, il s'agira de la quantification scalaire si une seule variable (la luminance) est à prendre en compte, la donnée est alors un histogramme unidimensionnel (cf. ci-dessus). Il s'agira de la quantification vectorielle si plusieurs variables sont à prendre en compte, dans le cas par exemple où l'unité statistique est un regroupement de pixels. La donnée est alors un histogramme multidimensionnel. Le cas de l'image couleur sera également considéré.

La méthodologie qui a été la plus expérimentée et développée dans nos travaux est celle d'AVL (cf. [15], [16], [18], [19], [20], [21], [27] et [28]). Néanmoins, la comparaison avec des méthodes plus classiques (cf. §3 ci-dessous) a été pour nous un souci constant.

### 3.3 - Image en niveaux de gris: Traitements et Résultats

Nous commençons par décrire les approches adoptées pour faire de la quantification d'images fixes en  $n$  niveaux de gris (ici  $n = 256$ ). Principalement, nous considérons l'approche de type classificatoire qui repose sur des méthodes de classification aussi bien hiérarchique que non hiérarchique.

Deux études sont élaborées. La première concerne la cas unidimensionnel où l'unité de base est le pixel (quantification scalaire  $QS$ ). La deuxième concerne le cas multidimensionnel où l'unité de base est un bloc de pixels (quantification vectorielle  $QV$ ), l'ensemble à classifier étant celui de tous les pixels de l' image. Différents indices de ressemblance entre les unités de base sont considérés et différentes stratégies de formation des classes sont étudiées.

#### 3.3.1 - Quantification scalaire

On retient pour la représentation statistique d'une image, un nuage unidimensionnel pondéré qui est tout simplement l' histogramme formulé par :

$$N(L) = \{ (l, n_l) / 0 \leq l \leq 255 \} \quad (8)$$

où  $n_l$  est le nombre de pixels possédant la luminance  $l$  et

$$\sum_{l=0}^{255} n_l = N = 256 \times 256, 512 \times 512 \text{ ou } 576 \times 720$$

selon l' image traitée.

Nous appliquons sur cet histogramme des techniques de seuillage résultant de diverses méthodes de classification pour obtenir une subdivision de l'échelle  $L = \{ 0, 1, \dots, 255 \}$  des luminances en  $k$  intervalles :

$$\sigma_k = (l_0, \dots, l_{j-1}, l_j, \dots, l_{k-1}, l_k = 255) \quad (9)$$

On note le  $j$ -ème intervalle

$$I_j = ] l_{j-1}, l_j ], \quad 1 \leq j \leq k \quad (10)$$

C'est à partir de la représentation adéquate de  $I_j$  par l'un de ses points ( $1 \leq j \leq k$ ) que s'effectue la quantification scalaire  $QS$ . Le plus important, est d'obtenir une "bonne" subdivision  $\sigma_k$ .

Pour ce faire, nous considérons deux types d'approches:

1- l'approche **classificatoire** consiste à subdiviser l'échelle  $L$  à l'aide de méthodes de classification. Nous considérons différents codages de la variable luminance, et notamment celui en préordonnance, différents indices de similarité et de dissimilarité entre points-objets et différents critères d'agrégation des classes. Ainsi ont été étudiées :

- la méthode mise en oeuvre par W.D. Fisher [7] avec un codage numérique. Elle rentre dans le cadre des méthodes de partitionnement. Elle ne s'applique que dans le cas où l'ensemble des objets est décrit par une seule variable descriptive et nécessite l'existence d'un ordre sur ces objets. Sous ces hypothèses et pour un nombre fixé de classes, l'algorithme associé à cette méthode fournit la solution globalement optimale au vu du critère de l'inertie (minimisation de l'inertie intra-classe). Le procédé adopté pour la recherche de cette partition optimale se fait de façon séquentielle et est inspiré de la programmation dynamique, approche fondée sur l'utilisation d'équations fonctionnelles et le principe d'optimisation [2].

- la méthode d'agrégation autour des centres mobiles (*k-means* et extensions, [8], [1], [6], [22]) avec un codage numérique.

- la méthode de la vraisemblance du lien (AVL) développée par I.C. Lerman (cf. références déjà citées) avec ici un codage en préordonnance, qui est de nature ordinale au niveau de l'échelle des différences. Cette méthode se conçoit dans un cadre général de représentation des données et s'adresse aussi bien à la classification de l'ensemble des objets que celui des variables. L'idée est de concevoir une mesure de proximité entre deux structures de même type qui ne prend en compte que ce qui est significatif par rapport à une hypothèse d'indépendance en respectant les caractéristiques de cardinalité de ces deux structures, c'est à dire que les liens entre les éléments à classer sont mesurés par rapport à une hypothèse d'absence de lien. L'indice de proximité entre deux classes est calculé à partir de la loi de probabilité du lien maximal; qui est d'ailleurs lui-même mesuré par la valeur observée d'une probabilité.

- la méthode développée par J.H. Ward [29] avec un codage numérique. Elle consiste, rappelons-le, à agréger les paires de classes qui minimisent la perte d'inertie.

2- l'approche **directe** consiste à faire une subdivision  $\sigma_k$  en k intervalles de l'échelle L, vérifiant :

$$\sigma_k = (l_0 = 0, \dots, l_{j-1}, l_j, \dots, l_{k-1}, l_k = 255) \quad (11)$$

Chacun des intervalles  $I_j = ] l_{j-1}, l_j ]$ ,  $1 \leq j \leq k$  est :

- soit de même longueur c'est à dire que :

$$l_{j-1} - l_j = [ 255 / k ] \quad (12)$$

Un tel découpage est appelé *linéaire-longueur* c'est-à-dire que toutes les classes ont le même nombre de valeurs de luminance;

- soit de même poids c'est-à-dire que :

$$nl_{j-1} - nl_j = [ (N_l \times N_c) / k ] \quad (13)$$

où  $nl_j$  (resp.  $nl_{j-1}$ ) représente le nombre de pixels de l'image ayant la luminance  $l_j$  (resp.  $l_{j-1}$ ) et  $N_l$  (resp.  $N_c$ ) représente le nombre de lignes (resp. de colonnes) de l'image.

Un tel découpage est appelé *linéaire-poids* c'est à dire que toutes les classes ont un nombre de pixels identique. Toutefois, les poids des classes sont presque, et non nécessairement exactement, les mêmes pour une raison très simple: l'échelle L est discrète et, par conséquent, ses valeurs sont entières.

Différentes subdivisions issues des deux approches sont analysées par leurs étendues et leurs poids, mais également comparées par plusieurs critères. Tous ces aspects font l'objet de la section 3.3.4. où nous expliciterons les résultats obtenus. Mais avant, nous allons préciser le codage en préordonnance de la variable luminance. Ce dernier s'avère très intéressant vis à vis du critère subjectif d'appréciation de la qualité des images décomprimées (cf. 3.3.4.). Nous allons également rappeler l'indice de la vraisemblance de lien.

L'ensemble des paires de modalités de la variable luminance est muni d'un préordre total traduisant de façon ordinale la ressemblance entre modalités. Une telle structure est appelée préordonnance sur L. Il s'agit d'un préordre total sur l'ensemble H des couples d'éléments de L, tel que:

$$H = \{(l, m) \mid 0 \leq l \leq m \leq 255\} \tag{14}$$

Ce préordre est établi de façon compatible avec la différence (m-l) :

$$\begin{aligned} & (0, 255) < \\ & (0, 254) \sim (1, 255) < \\ & (0, 253) \sim (1, 254) \sim (2, 255) < \\ & \dots\dots\dots \\ & (0, 255 - h) \sim (1, 256 - h) \sim \dots \sim (h, 255) < \\ & \dots\dots\dots \\ & (0, 0) \sim (1, 1) \sim (2, 2) \sim \dots \sim (255, 255) \end{aligned} \tag{15}$$

Ce préordre total comporte 256 classes, la (h+1) ème - en comptant de haut en bas - (0, 255 - h) ~ (1, 256 - h) ~ .... ~ (h, 255), comporte exactement (h+1) éléments dont chacun se trouve défini par un couple de la forme (i, 255 + i - h), 0 ≤ i ≤ h.

On a

$$\text{card}(H) = \sum_{h=0}^{255} (h+1) = 1/2 \text{card}(L)(\text{card}(L) + 1) \tag{16}$$

Nous codons un tel préordre sur H par la notion de rang moyen. Ainsi, si (l, m) appartient à la (h + 1) ème classe; c'est à dire si (l-m) = (255-h), nous avons alors ([9], [27]) :

$$\rho[(l, m)] = 1 + 2 + \dots + h + (1/2)(h+2) \tag{17}$$

En remarquant que  $\sum_{i=1}^h i = (1/2)h(h+1)$ , l'équation (10) devient :

soit

$$\rho[(l, m)] = (1/2) [h(h+1) + (h+2)]$$

$$\rho[(l, m)] = (1/2) [(h+1)^2 + 1] \tag{18}$$

Comme  $(m-1) = (255 - h)$ , on obtient finalement:

$$\rho[(l, m)] = (1/2) [(256 - (m-1))^2 + 1] \quad (19)$$

La préordonnance (15) induit sur l'ensemble des pixels une préordonnance qui est un préordre total sur l'ensemble des  $(1/2)(N)(N+1)$  paires de pixels, où une même paire  $\{i, j\}$  se trouve rangée conformément au rangement de  $(L(i), L(j))$  dans (15), en supposant que  $L(i) \leq L(j)$ .

La distribution d'un tel codage est décrite dans [9].

On prend comme indice brut de similarité entre deux pixels  $i$  et  $j$  celui défini par (12). Il sera ensuite statistiquement normalisé compte tenu de l'histogramme (1).

En effet, soit  $\{i, j\}$  une paire de pixels éléments de  $\Pi$ . On définit l'indice brut de similarité entre  $i$  et  $j$  par :

$$s(i, j) = \rho(L(i), L(j)) \quad (20)$$

où  $\rho[(l, m)]$  a été défini par (12). Ainsi :

$$s(i, j) = (1/2)[(256 - |L(i) - L(j)|)^2 + 1] \quad (21)$$

Cet indice varie entre 1 (si  $|L(i) - L(j)| = 255$ ) et 32768.5 (si  $|L(i) - L(j)| = 0$ ).

Par centrage et réduction de  $s$  par rapport à sa distribution sur l'ensemble  $P_2$  ( $\Pi$ ) des paires de pixels, on obtient :

$$Q(i, j) = (1/\sigma)(s(i, j) - \mu) \quad (22)$$

où  $\mu$  et  $\sigma^2$  sont respectivement, la moyenne et la variance de  $s$  sur l'ensemble  $\Pi$  des paires de pixels.

On considère alors une hypothèse d'absence de lien où à la paire  $\{i, j\}$ , on associe une paire  $\{i^*, j^*\}$  d'objets aléatoires indépendants, où  $Q(i^*, j^*)$  devient une variable aléatoire  $\mathcal{N}(0,1)$ ; c'est à dire, normale centrée et réduite.

L'indice de la vraisemblance du lien est alors :

$$P(i, j) = \Pr [Q(i^*, j^*) \leq Q(i, j)] = \Phi(Q(i, j)) \quad (23)$$

où  $\Phi$  est la fonction de répartition de la loi  $\mathcal{N}(0,1)$ .

Ainsi, la ressemblance entre  $i$  et  $j$  est évaluée d'autant plus grande que  $Q(i, j)$  apparaît invraisemblablement grand, eu égard à la distribution de  $Q(i^*, j^*)$  dans l'hypothèse d'absence de lien.

Étant données deux classes C et D définissant deux parties disjointes de L (cf. (1) ci-dessus). L'indice de comparaison de la "vraisemblance du lien maximal" (cf. références ci-dessus mentionnées) se met ici sous la forme suivante :

$$\text{Pr} [ p (C^*, D^*) \leq p (C, D) ] \quad (24)$$

où

$$p (C, D) = \max \{ P (l, m) \mid (l, m) \in C \times D \} \quad (25)$$

avec conformément à (23),

$$[ \forall (l, m) \in L \times L ], \quad P(l, m) = \Phi [ (1/\sigma)(\rho (l, m) - \mu) ] \quad (26)$$

où  $\rho (l, m)$  est l'indice exprimé dans (17).

On démontre que le critère d'agrégation s'écrit (cf. références citées) :

$$VL (C, D) = [ p (C, D) ]^{c \times d} \quad (27)$$

où  $c = \text{card} (C)$  et  $d = \text{card} (D)$ .

Suite à la démarche de F. Nicolaiu [25] proposant de remplacer la fonction produit  $c \times d$  par une autre fonction  $g (c, d)$ ; dont notamment  $\sqrt{c \times d}$ , nous avons suggéré la famille de critères de la " vraisemblance du lien maximal" sous la forme :

$$VL (C, D) = [ p (C, D) ]^{(c \times d)^\varepsilon} \quad (28)$$

où  $\varepsilon$  est un paramètre réel variant entre 0 et 1. Nous montrons que la valeur  $\varepsilon = 0.5$  a un intérêt formel particulier.

Pour des raisons de précision de calcul, la forme équivalente et définitive du critère, utilise la fonction strictement croissante  $-\text{Log} (-\text{Log})$ .

Précisément :

$$S_\varepsilon (C, D) = -\varepsilon (\text{Log } c + \text{Log } d) - \text{Log} \{ -\text{Log} [ p(C, D) ] \} \quad (29)$$

Rappelons qu'une formule de réactualisation élégante existe pour ce critère [17], qui est, au sens de Bruynooghe [3], "contractant" (propriété des voisinages réductibles).

Signalons que dans le cadre de notre expérience de quantification scalaire, la classification s'applique au nuage unidimensionnel  $\mathbf{N} (L)$  [ cf. (8) ci-dessus ]; les classes initiales sont les pieds d'un histogramme et la matrice initiale des dissimilarités entre classes se met sous la forme:

$$\{ S_{\varepsilon}(l, m) \mid 0 \leq l \leq m \leq 255 \} \quad (30)$$

où, le cardinal de la classe d'étiquette  $l$ , est  $n_l$ ,  $0 \leq l \leq 255$ .

De la sorte, dans l'équation (30) ci-dessus

$$c = \sum \{ n_l \mid l \in C \}, \text{ (resp. } d = \sum \{ n_l \mid l \in D \} \text{)} \quad (31)$$

Une valeur forte de  $\varepsilon$  (par trop voisine de 1) peut entraîner quelques "petites" disconnexions dans la formation hiérarchique des classes de luminance. Mais, il suffit d'une diminution de  $\varepsilon$  pour, très vite, restituer la connexion des classes qui deviennent des intervalles de luminance. D'autre part et curieusement, cette diminution paramétrée améliore la qualité perçue de l'image compressée. Alors que nous démontrons [18] que le critère de l'inertie {méthode de Ward (cf. ci-dessous)} entraîne le respect de la condition de connexité. Cependant, il faut garder à l'esprit que ce n'est pas une propriété mathématique, ni même un critère quantitatif d'évaluation, qui peuvent constituer les garants absolus de la qualité perçue de l'image restituée.

### 3.3. Quantification vectorielle d'une image en niveaux de gris

Une première démarche consiste à découper l'image en blocs carrés adjacents de même taille  $M$  (par exemple  $4 \times 4$ ) et de quantifier scalairement chacun des blocs. On pourrait alors adopter soit l'approche directe soit l'approche classificatoire basée sur l'histogramme des luminances, à condition toutefois, que la taille  $M$  des blocs soit suffisamment grande.

Un autre procédé - visant une compression plus intéressante - consiste à faire la classification de ces blocs d'image. C'est ce type d'approche qui a été retenu.

On découpe l'image en blocs adjacents carrés de taille fixe  $M$  ( $2 \times 2$ ,  $4 \times 4$  ou  $8 \times 8$ ). Ce découpage correspond à un tableau de données à  $M$  colonnes et  $(N/M)$  lignes où  $M$  et  $N$  représentent respectivement la taille de chaque bloc (par exemple  $m = 16$ ) et la taille de l'image (par exemple  $N = 512 \times 512$ ). L'intersection entre la ligne  $h$  et la colonne  $c$  correspond à  $L(i_c^h)$  qui représente la valeur de luminance du pixel  $i$  du bloc  $b_h$  de l'image.

Afin de réduire l'information à transmettre, on applique à chaque bloc une transformation permettant d'obtenir des composantes peu ou pas corrélées. C'est la transformée discrète en cosinus [10] qui est la plus utilisée. Les nouveaux blocs constitués par les vecteurs des coefficients numériques de la transformée sont ensuite organisés en classes et sous-classes. A cette fin, l'algorithme le plus utilisé - dit LBG dans le domaine [22] - est celui des centres de gravité mobiles, bien connu en classification, sous le nom des "k-means". On effectue une classification en cascade; de façon récursive. On commence par une

première partition globale sur l'ensemble total; puis on applique à chacune des classes, le même algorithme; et si nécessaire, ainsi de suite, jusqu'à une profondeur donnée, qui peut dépendre de la classe découpée.

On sait bien que l'un des points les plus vulnérables de cet algorithme consiste en le choix a priori du nombre de classes de l'ensemble à classifier et surtout du système initial des "centres d'attraction" (ou "noyaux"). Dans notre cas, chaque élément (figuré par un sommet de l'espace de représentation) est un bloc de taille  $M$ . Ainsi, par rapport au choix arbitraire des centres, on peut cependant et dans notre contexte mentionner que ces centres peuvent, par exemple, être constitués de blocs *homogènes* (c'est à dire toutes les composantes sont égales), de blocs *textures* (c'est à dire le contraste est inférieur ou égal à un certain seuil), des blocs contours horizontaux, verticaux et obliques.

Notre contribution se situe au niveau de la classification de blocs d'image par le biais de méthodes hiérarchiques. Ainsi, étant donné deux blocs quelconques  $b_h$  et  $b_{h'}$  tels que :

$$b_h = \{L(\dot{\mathbf{i}}_1^h), \dots, L(\dot{\mathbf{i}}_c^h), \dots, L(\dot{\mathbf{i}}_M^h)\} \quad (32)$$

où  $L(\dot{\mathbf{i}}_c^h)$  est la luminance du pixel  $\dot{\mathbf{i}}_c$  du bloc  $b_h$ , et

$$b_{h'} = \{L(\dot{\mathbf{i}}_1^{h'}), \dots, L(\dot{\mathbf{i}}_c^{h'}), \dots, L(\dot{\mathbf{i}}_M^{h'})\} \quad (33)$$

où  $L(\dot{\mathbf{i}}_c^{h'})$  est la luminance du pixel  $\dot{\mathbf{i}}_c$  du bloc  $b_{h'}$ , l'indice de distance entre  $b_h$  et  $b_{h'}$  peut-être défini par :

$$d(b_h, b_{h'}) = \sqrt{\sum_{c=1}^M (\mathcal{L}(\dot{\mathbf{i}}_c^h) - \mathcal{L}(\dot{\mathbf{i}}_c^{h'}))^2} \quad (34)$$

Pour un tel choix d'indice, on peut considérer les quatre critères d'agrégation des classes suivants :

- le saut minimal,
- le diamètre,
- la distance moyenne,
- le critère de Ward.

On peut également considérer la méthode de la vraisemblance du lien (AVL) où chacune des variables est codée en terme de préordonnance et où l'indice de la vraisemblance du lien entre objets est analogue à celui défini dans le cas d'une seule variable ([20] et [9]).

### 3.3.3 Résultats

Nous exposons les résultats obtenus par les approches décrites précédemment ainsi que les critères pour évaluer de tels résultats. L'objectif étant - rappelons-le - d'étudier le



comportement de chacune de ces méthodes vis à vis des données images et d'évaluer les résultats obtenus par la qualité perceptuelle des images reconstruites après compression. Six images test ont été utilisées - choisies parmi celles reconnues au sein de la communauté "codage d'image" pour leurs difficultés. Ce sont des images fixes en 256 niveaux de gris - chaque pixel étant codé sur 8 bits - de taille  $N_l \times N_c$ , où  $N_l$  représente le nombre de lignes et  $N_c$  le nombre de colonnes. Il s'agit de :

- 1- l'image *Barbara* de taille 576 x 720,
- 2- l'image *Bateau* de taille 576 x 720,
- 3- l'image *Lena* de taille 256 x 256,
- 4- l'image *Peppers* de taille 512 x 512,
- 5- l'image *Voiture* de taille 512 x 512,
- 6- l'image *Zelda* de taille 576 x 720.

Les résultats sont analysés pour comprendre l'influence du codage (des objets) et des critères - en termes d'analyse des données - sur la qualité perceptuelle des images reconstruites.

L'évaluation de l'ensemble des résultats obtenus est faite en considérant à la fois :

- a) des critères subjectifs (appréciation directe et humaine des images reconstruites) habituellement utilisés en compression d'image numérique;
- b) des critères définis par le CCIR (Comité Consultatif International des Radiocommunications). Nous avons plus particulièrement considéré la méthode de *comparaison par paires* [5], [11].

Le but d'une telle méthode est de faire un classement des différentes images reconstruites par un certain nombre d'observateurs. On édite alors une bande de test où sont présentées toutes les combinaisons de ces images par paires. L'ordre d'apparition des paires doit être aléatoire mais deux images identiques ne peuvent se suivre. Lors de la visualisation, il est demandé à l'observateur de décider qu'elle est d'un point de vue psychovisuel la meilleure image dans la paire courante.

Les votes du test sont arrangés dans des matrices de dominance triangulaires pour chaque image et chaque observateur.

On compte le nombre de triades intransitives et l'on estime le niveau de signification de ces intransitivités. On analyse ensuite le niveau de cohérence des préférences entre observateurs et ceci pour chaque image. Finalement, on obtient un classement par image avec une probabilité associée [9].

c) Des critères objectifs à savoir :

- l'entropie de l'image des différences donnée par :

$$e = -\frac{1}{N_l \times N_c} \sum_{l=0}^{255} n_l \log_2 \left( \frac{n_l}{N_l \times N_c} \right) \quad (35)$$

où  $N_l \times N_c$  représente le nombre total de pixels de l'image et  $n_l$  correspond au nombre de pixels possédant  $l$  comme niveau de gris de différence (en valeur absolue) entre les deux images. La base 2 du logarithme correspond à l'unité habituellement utilisée en théorie de l'information à savoir le bit;

- L'erreur quadratique moyenne entre l'image originale et l'image reconstruite après quantification, donnée par:

$$EQM = -\frac{1}{N_l \times N_c} \sum_{i=1}^{N \times M} (l_i - \hat{l}_i)^2 \quad (36)$$

où  $l_i$  est le niveau de gris associé au pixel  $i$  dans l'image originale et  $\hat{l}_i$  correspond au niveau de gris associé au pixel  $i$  dans l'image reconstruite;

- le taux de compression défini par le ratio entre le nombre de bits par pixel de l'image originale et le nombre moyen de bits par pixel après compression (ou le ratio entre la quantité d'information avant et après compression).

En quantification scalaire, l'étude porte plus particulièrement sur la quantification en 16, 8 et 4 classes. Ce faisant, nous construisons les images correspondantes à chaque niveau de quantification en représentant chaque classe par la moyenne de la restriction de l'histogramme des luminances à cette classe. L'étape de quantification nous permet de passer d'une résolution sur 256 niveaux de gris à une résolution plus faible. Elle permet de passer d'une image où chaque pixel est codé sur 8 bits à une image où ce dernier est codé sur  $\log_2(q)$  où  $q$  est le nombre de niveaux de reconstruction.

Mentionnons, par ailleurs, que la méthode des *centres mobiles* nécessite le choix a priori de  $k$  centres -  $k$  étant le nombre de classes. La suite des centres, considérée pour toutes les images traitées, est celle issue du découpage linéaire-longueur de l'échelle  $L$ , c'est à dire :

$$\left( \left[ \frac{255}{k} \right], 2 \times \left[ \frac{255}{k} \right], \dots, k \times \left[ \frac{255}{k} \right] \right) \quad (37)$$

Cette suite initiale des centres peut être issue du découpage linéaire de l'échelle L, non pas en longueur, mais en poids. Elle peut être issue d'un algorithme de classification non hiérarchique, hiérarchique ou encore peut être choisie aléatoirement.

La planche 1 (reconstruction de *Lena* sur 16 niveaux) donne les images reconstruites en 16 classes - chaque classe étant représentée sur un niveau de gris - par les différentes méthodes des deux approches classificatoire et directe, pour l' image *Lena*. La différence réside dans la restitution de l'épaule et le front qui est légèrement meilleure pour *Fisher* et *Ward*. On distingue davantage de faux contours pour la *linéaire-longueur* et les *centres mobiles*.

**Tableau 1 : Classement par image, étude comparative.**

images	Classements
Barbara	$a_{0.1}(20) \succ l(20) \succ a_{0.5}(20) \succ a_{0.5}(16) \succ a_{0.1}(16) \succ l(16)$
Bateau	$a_{0.5}(16) \succ a_{0.1}(16) \succ l(20) \succ a_{0.1}(20) \succ a_{0.5}(20) \succ l(16)$
Peppers	$l(20) \succ a_{0.1}(20) \succ a_{0.5}(20) \succ a_{0.5}(16) \succ l(16) \succ a_{0.1}(16)$
Voiture	$a_{0.5}(16) \succ a_{0.1}(16) \succ l(16) \succ a_{0.1}(20) \succ a_{0.5}(20) \succ l(20)$
Zelda	$a_{0.1}(20) \succ l(20) \succ a_{0.5}(20) \succ a_{0.1}(16) \succ a_{0.5}(16) \succ l(16)$

L'étude comparative menée par le CCETT donne le classement, illustré par le tableau 1, obtenu pour chaque image, où " $\succ$ " désigne "préféré à" et où  $a_{\epsilon}(q)$  et  $l(q)$  désignent respectivement  $AVL_{\epsilon}$  et linéaire-longueur pour  $q=20, 16$  classes.

Le traitement par image montre que, d'une part, la quantification sur 20 niveaux de gris n'est pas nécessairement préférée à celle sur 16 niveaux de gris. D'autre part, le codage  $a_{0.1}$  n'est pas toujours préféré à un codage  $a_{0.5}$  ou réciproquement. L'appréciation de l'observateur dépend de la nature et du contenu de l'image. Cependant, le classement global

$$a_{0.1}(20) \succ l(20) \succ a_{0.5}(20) \succ a_{0.5}(16) \succ a_{0.1}(16) \succ l(16)$$

donne un résultat plus clair puisque toutes les quantifications sur 20 niveaux sont préférés à celles sur 16 niveaux.

**Tableau 2 : Entropies de l'image des différences en QS, par méthode.**

image	#classes	Linéaire		Fisher	Ward	centres linéaire	AVL 0.5
		longueur	poids				
Voiture	4	5.98	5.68	5.61	5.66	5.69	5.63
	8	5.07	4.90	4.80	5.04	4.90	4.99
	16	4.04	4.05	3.83	3.96	3.97	4.05
Lena	4	6.03	5.97	5.74	6.00	5.78	5.82
	8	5.07	4.97	4.80	4.93	4.96	5.20
	16	4.08	4.04	3.83	3.94	3.97	4.05
Bateau	16	4.07	3.95	3.63	3.74	3.84	3.82
Zelda	16	4.06	3.86	3.59	3.75	3.94	3.82

PLANCHE 1: Quantification Scalaire (1)



a)



b)



c)



d)



e)



f)

Reconstruction de "Lena" sur 16 niveaux de gris en QS:

a) image origine, b)  $\mathcal{AVL}$  0.5, c) Fisher, d) Ward, e) linéaire-poids, f) linéaire-longueur

Les tableaux 2 et 3 donnent respectivement les entropies des images de différences et l'EQM des images reconstruites, correspondant aux différentes méthodes pour 16, 8 et 4 classes.

**Tableau 3 : EQM des images reconstruites e QS, par méthode.**

image	#classes	Linéaire		Fisher	Ward	centres linéaire	AVL 0.5
		longueur	poids				
Voiture	4	266.7	191.4	179.7	186.5	199.1	184.0
	8	75.4	79.6	55.2	67.1	61.2	68.7
	16	20.1	37.4	13.8	16.4	17.7	18.5
Lena	4	302.1	269.6	194.3	261.5	211.7	211.1
	8	80.3	63.5	49.6	59.3	61.5	92.0
	16	21.0	17.4	13.6	15.0	17.0	17.2
Bateau	16	20.3	23.3	9.5	10.9	13.2	12.5
Zelda	16	20.9	15.6	10.1	11.6	15.5	14.0

L'entropie est une fonction qui mesure la quantité d'information contenue dans l'image. Si l'image était parfaitement reconstruite, l'entropie de l'image des différences serait égale à 0. L'opération de quantification se faisant avec perte d'information, l'entropie de l'image des différences est nécessairement supérieure à 0. Mais cette mesure ne permet en aucun cas de juger la qualité perceptuelle de l'image reconstruite. L'EQM ne le permet pas non plus. Il faut donc faire appel à des tests psychovisuels.

Pour une quantification sur 16 niveaux dont la qualité perceptuelle est *satisfaisante*, on passe de 8 bits/pixel à 4 bits/pixel pour ainsi obtenir un taux de compression égal à 2.

L'étude en quantification scalaire peut constituer un préalable indispensable à celle de la quantification vectorielle.

Ainsi en quantification vectorielle, seule l'approche classificatoire hiérarchique de blocs de pixels de taille 4 x 4 est analysée. Les deux images considérées pour cette étude sont *Lena* et *Voiture* où nous avons classifié les blocs des pixels par les cinq méthodes mentionnées auparavant (saut minimal, saut maximal, saut moyen, vraisemblance du lien -  $\epsilon = 1$  et  $\epsilon = 0.5$  -, et Ward) en 256 classes. Le représentant de chaque classe étant le *bloc moyen* de taille 4 x 4, des blocs appartenant à la classe. Nous nous sommes également intéressés au comportement de ces méthodes et à la vitesse de dégradation de la qualité de l'image en considérant des classifications en 128, 64 et 32 classes.

La figure 2 (reconstruction de Lena en 256 classes en QV) donne la reconstruction de l'image Lena en 256 classes. Ce qui ressort de façon nette de cette quantification est que le saut minimal ne convient pas du tout. Les quatre autres méthodes se caractérisent par *l'effet de bloc* qu'il va falloir atténuer. ce dernier est plus perceptible pour *Lena* que pour *Voiture*.

PLANCHE 2 : Quantification Vectorielle (2)



a)



b)



c)



d)



e)



f)

Reconstruction de "Lena" sur 256 niveaux de gris en QV:

a) image origine, b) AVL 0.5, c) Ward, d) AVL 1, e) diamètre f) distance moyenne

Notons par  $n_c$  le nombre de classes considérées pour faire de la quantification,  $t_b$  la taille d'un bloc et  $N_b$  le nombre total des blocs à classer. Le débit  $D$  est obtenu de la façon suivante:

$$D = \log_2(n_c) + \frac{n_c \times t_b \times 8}{N_b} \quad (38)$$

de sorte que pour une image de taille  $128 \times 128$  où  $t_b = 16$ ,  $N_b$  est alors égal à 1024. Si on prend  $n_c = 256$ , on a alors  $D = 40$  bits par bloc soit un débit de 2.5 bits par pixel.

Ce faisant, nous avons codé chaque élément du bloc représentant d'une classe sur 8 bits. sans détériorer la qualité des images reconstruites après compression, on peut réduire ce débit avec une qualité perceptuelle de l'image reconstruite presque identique en réduisant le nombre de bits nécessaire pour coder un élément du représentant d'une classe, caractérisé par sa valeur de luminance. A cette fin, nous proposons d'utiliser les résultats obtenus en quantification scalaire pour passer de 8 bits à un nombre inférieur, multiple de 2. Le débit devient:

$$D' = \log_2(n_c) + \frac{n_c \times t_b \times 4}{N_b} \quad (39)$$

de sorte que pour les mêmes valeurs de  $n_c$ ,  $t_b$  et  $N_b$  que celles considérées auparavant, on obtient un taux de compression  $D' = 24$  bits par bloc soit un débit de 1.5 bits par pixel.

Nous avons ainsi pu montrer tout l'intérêt de l'approche classificatoire dans la quantification scalaire et vectorielle d'une image fixe en niveaux de gris. En effet, par rapport à une technique naïve (e.g. la "linéaire" pour la QS), la qualité perceptuelle des images reconstruites est nettement meilleure; d'autre part, les mesures de l'entropie et de l'EQM sont globalement plus petites.

Nous allons à présent compléter notre présentation en considérant une autre forme de quantification vectorielle, pour les images en couleur qui a nourri [14] et [9]. Également dans ce cadre, diverses méthodes ont été considérées et comparées; en relation avec la méthode AVL.

### 3.4 - Image en couleur; Traitements et Résultats.

#### 3.4.1 - La donnée de base

Il y a lieu ici de se rappeler du paragraphe 3.2 (cf. expression (3)) où chaque pixel  $(i, j)$  se trouve décrit par un vecteur à 3 composantes (R, V, B); dont chacune est codée sur une échelle à 256 valeurs; c'est à dire, nécessitant 8 bits par pixel. Nous avons ici, afin de diminuer la quantité d'information à traiter - sans nuire notablement à la qualité

psychovisuelle de l'image - opérer une préquantification *linéaire-longueur* sur chacune des composantes R, V et B, de l'image pour passer d'une représentation de 8 bits à 5 bits par pixel (soit donc 32 valeurs possibles au lieu de 256). Le nuage de points à traiter se trouve ainsi limité à  $2^{15}$  points (32 512) dans l'espace (R, V, B).

Dans ces conditions, la donnée de base, au lieu d'être un nuage (ou histogramme) unidimensionnel [cf. (8)], est un nuage (ou histogramme) tridimensionnel, qu'on peut écrire sous la forme :

$$N(J) = \left( (P, n_p) / p \in [0, 31]^3 \right) \quad (40)$$

où J désigne l'image concernée et où P est le sommet de l'espace tridimensionnel (R, V, B), de composantes  $\{R(p), V(p), B(p)\}$  qui représente l'ensemble des pixels (i,j) de l'image J ayant la même représentation spatiale (cf. (3)).  $n_p$  est le nombre d'éléments de cet ensemble.

Nos expériences de classification ont principalement porté sur cet espace non métrique de représentation des couleurs; et ce en associant à chaque composante R, V, B, une variable préordonnance. L'indice de similarité adopté pour la classification AVL de N(J) est celui développé dans [20] et implanté dans CHAVL [21].

La classification a également porté sur des représentations dans d'autres espaces de représentation des couleurs, bien connus - approximativement métrique tel que "L\*uv" (défini par la CIE (Commission Internationale pour l'Éclairage)), ou non métrique tel que "Y CrCb", très utilisé en télévision.

### 3.4.2 - Méthodes et Résultats

L'effort a porté sur la comparaison des différentes méthodes de classification suivantes :

- La méthode autour de la recherche des K modes de la distribution tridimensionnelle ("popularity algorithm" de Heckbert [12]), avec usage de la métrique définie dans l'espace colorimétrique CIE Luv.
- Segmentation hiérarchique par dichotomies successives. Chaque dichotomie concerne le couple formé par une classe et un axe (parmi les trois de l'espace colorimétrique) : (Classe, Axe). A chaque fois, on choisit le couple (Classe, Axe) pour lequel la variation de la classe sur l'axe est maximale. Cette variation peut-être mesurée par l'écart maximum pondéré [en tenant compte des poids  $n_p$  (cf. [40]) ou bien, par l'inertie.
- La méthode AVL (Analyse de la vraisemblance du lien) dont il était question ci-dessus, avec un codage des variables en termes de "préordonnances" [18], [20], [9].
- La "QVpar bloc" ("Color Cell Compression" de Campbell [4]) par :



(1) division de l'image en blocs 4 x 4, (2) partition des couleurs de l'image en 2 classes P0 et P1 (inf (resp. sup) à la luminance du bloc considéré) pour tous les blocs de l'image, (3) génération des alphabets A0 et A1 correspondants, (4) codage avec A1 et A0 selon que le pixel appartient à P0 ou P1. La méthode utilise l'espace CIELuv. Le débit moyen obtenu est de l'ordre de 1 à 2 bits/pixel.

- La méthode dite "LBG" (de Linde, Buzo et Gray, [22]) qui n'est autre que celle des K-means [8], [1].

- La méthode "QV moyenne/variance" de Murakami [23] par :

(1) division de l'image en blocs 4 x 4 pour chacune des composantes Y, Cr et Cb, (2) centrage et normalisation, (3) codage par alphabets de formes élémentaires (8192 pour l'ensemble des composantes Y, Cr et Cb prises séparément, et obtenues sur séquence d'apprentissage, par un algorithme de type "K-means"), (4) codage en QS de la moyenne et de la variance des blocs de l'image à coder. Débit inférieur à 1 bit/pixel pour des images couleurs.

Les conclusions de l'analyse expérimentale effectuée (cf. [14]) sont les suivantes :

- l'approche AVL associée à un codage en préordonnance de la variable, s'est toujours mieux comportée.

- le codage en préordonnance est très adapté à l'expression de nos sensations psychosensorielles.

- le codage à bas débit (< 1 bit/pixel) en QV directe s'est avérée impossible, sans y adjoindre des techniques de compression complémentaires (codage entropique, changement d'espace source pour réduire la corrélation spatiale, etc.....). D'où le besoin d'intégrer la QV dans un schéma de codage à réduction de débit plus complet.

- la QV couleur directe présente toutefois une propriété de maintien des contours de l'image - à faible nombre de classes ( 8 ou 16) très intéressante (feuilletage de catalogue d'images à haute résolution, ...),

- globalement, les débits obtenus varient de 1 bit/pixel ("algorithme Color Cell Compression" - bloc de 4x4) à environ 3 bit/pixel ( QV couleur directe, sans regroupement de pixels ).

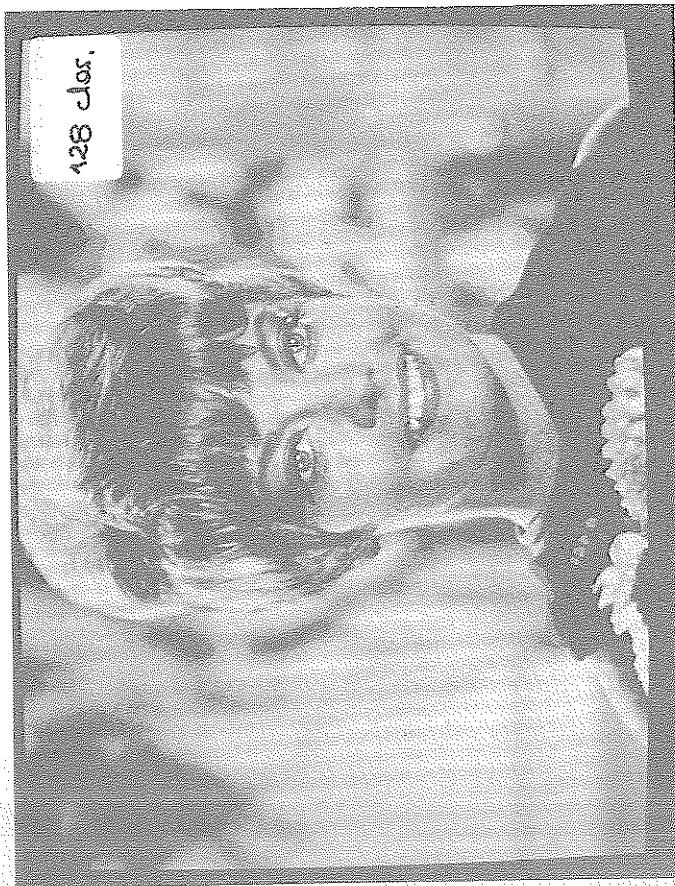
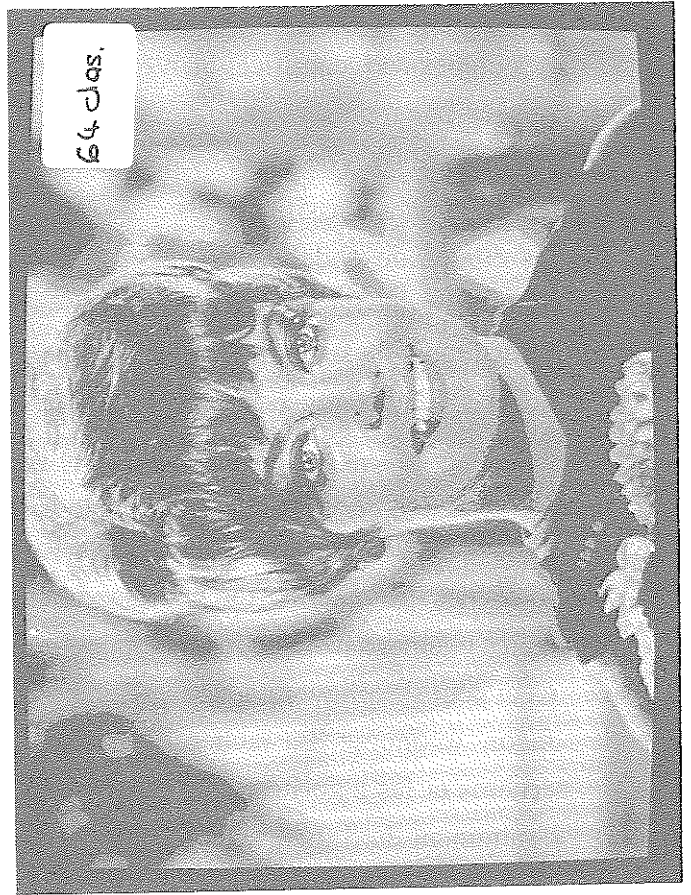
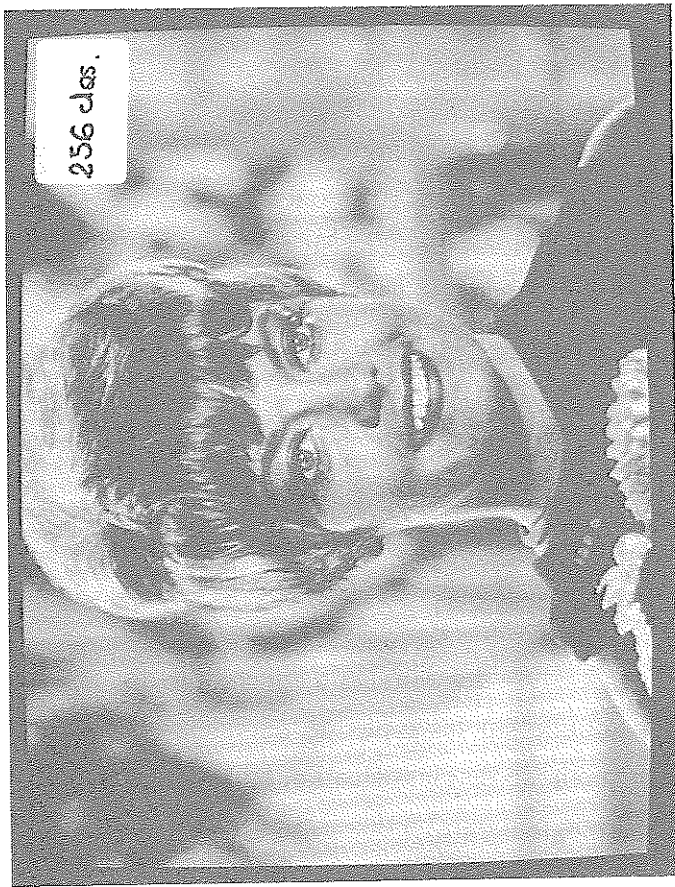
- voir planche Quantification Vectorielle (3): ZELDA (32, 64, 128, 256 classes "popularity")

- voir planche Quantification Vectorielle (4): ZELDA (16, 48, 137, 250 classes "AVL")

PLANCHE 3 : Quantification Vectorielle (3) - algorithme "Popularity"  
(32, 64, 128, 256 classes)



PLANCHE 4 : Quantification Vectorielle (4) - algorithme "AVL préordonnance R, V, B.  
(16, 48, 137, 250 classes)



#### 4 - CONCLUSIONS

Les études de Quantification Vectorielle appliquées au codage à réduction de débit des images fixes, présentées ci-dessus - assez brièvement sur les aspects méthodologiques, pour des raisons de place évidentes - ont fourni un champ d'étude nouveau et très riche aux méthodes classificatoires classiques de l'analyse des données. Elles ont montré que les études de codage des images, "appartenant" traditionnellement au domaine du traitement du signal, pouvaient-être aisément vues sous l'angle de l'analyse des données statistiques - le souci de synthétiser l'information étant le même, les méthodes, sous des noms différents, procèdent de la même démarche.

Des résultats concrets ont été obtenus sur les possibilités de l'approche "analyse des données" qui "résistent très bien au choc des données images", puisque après quelques retouches des critères, donnent souvent des appréciations psychovisuelles supérieures à celles obtenues par d'autres méthodes.

Ceci peut s'expliquer parce que la méthodologie en analyse des données a été étudiée de façon autonome. La recherche est ainsi plus profonde sur les types de données, les critères et les algorithmes. Alors que lorsqu'une méthode de classification apparait en Traitement du Signal, elle apparait un peu "en fin de parcours" à des fins applicatives directes et où le souci n'est plus une analyse approfondie, notamment d'un point de vue mathématique, du comportement de l'algorithme.

Dans les différentes études présentées, nous avons en vue la démonstration du "bon" comportement de la méthode AVL; et cela, en comparaison avec les méthodes les plus célèbres appliquées dans le domaine. Dans ce cadre, ainsi d'ailleurs, dans celui plus général, des interactions entre l'analyse des données et la quantification, les auteurs sont conscients que dans le travail exposé, le sujet est loin d'être épuisé. A cet égard il suffit de remarquer que les études sur ce thème sont toujours très présentes dans les colloques internationaux sur le traitement du signal, qu'une technique de compression "VQ" pour la TVHD a été proposée par des laboratoires Américains, que la QV est souvent utilisée dans des algorithmes de codage de l'image - non sur des données images brutes, mais sur des images prétraitées (après analyse en cosinus, en ondelettes, ... ).

*Les auteurs tiennent à remercier le CCETT et l'IRISA pour leur soutien dans ces études exploratoires.*

*Une reconnaissance toute particulière est adressée à Thierry ALPERT (CCETT/TAV/CEP) pour sa contribution centrale aux tests subjectifs d'évaluation de la qualité perceptuelle des images, ainsi qu'à Thierry FAUCONNIER (IRISA/Projet Themis) pour son aide constante et sa disponibilité.*

## 5 - RÉFÉRENCES

- [1] G.H. BALL AND D.J. HALL, "ISODATA, a novel method of data analysis and pattern classification". Stanford Research Institute, Menlo Park, California, Technical report, SRI project 5533, 1965.
- [2] R.E. BELLMAN ET S.E. DREYFUS, "La programmation dynamique et ses applications". Dunod, Paris, 1965.
- [3] M. BRUYNOOGHE, "Nouveaux algorithmes en classification automatique applicables aux très gros ensembles de données, rencontrés en traitement d'images et en reconnaissance des formes". Thèse d'État, Université de Paris VI, 1989.
- [4] G. CAMPBELL et al., "Two bit/pixel full color image encoding", SIGGRAPH, Vol. 20, N°4, Dallas, Aug. 1986.
- [5] CCIR, "A paired comparison procedure for picture quality assessments. Documents CCIR, Subject : interim working party 11/4, Doc. 11/4 - 156, Period 1986-1990.
- [6] E. DIDAY, "La méthode des nuées dynamiques". Rev. Stat. App., Vol. 19, N°2, pp. 19-34, 1971.
- [7] W.D. FISHER, "On grouping with maximum homogeneity". J.A.S.A., Vol. 1, pp. 4-29, 1958.
- [8] E.W. FORGY, "Cluster analysis of multivariate data : Efficiency versus Interpretability of classifications". Biometrics, Vol. 21, pp. 768-769.
- [9] N. GHAZZALI, "Comparaison et réduction d'arbres de classification, en relation avec des problèmes de quantification en imagerie numérique", thèse de Doctorat de l'Université de Rennes I, 14 mai 1992.
- [10] A. GHERSHO and R.M. GRAY, "Vector quantization and signal compression". Kluwer Academic publishers, 1992.
- [11] D.J. GRANRATH, "The role of human visual models in image processing", Proceedings of the IEEE, Vol. 67, pp. 552-561, May 1981.
- [12] P. HECKBERT, "Colorimage quantization for frame buffer display". Computer Graphics, Vol. 16, N° 3, July 1982.
- [13] D.A. HUFFMAN, "A method for the construction of minimum redundancy codes", Proceedings IRE Vol. 40, pp. 1098-1101, September 1952.
- [14] A. LÉGER, "Introduction à la quantification vectorielle des images fixes". Thèse de Docteur-Ingénieur, Université de Rennes I, 18 décembre 1987.
- [15] I.C. LERMAN, "Sur l'analyse des données préalable à une classification automatique (proposition d'une nouvelle mesure de similarité)", Rev. Math & Sc. Hum., 8ème année, n°32, 1970.

- [16] I.C. LERMAN, "Classification et analyse ordinale des données", Dunod, Paris 1981.
- [17] I.C. LERMAN, "Formules de réactualisation en cas d'agrégations multiples", Rairo série R.O., Vol. 23, n°2, pp. 151-163, 1989.
- [18] I.C. LERMAN, N. GHAZZALI, "Quantification image par AVL, avec codage préordonnance", Rapport interne IRISA.
- [19] I.C. LERMAN, N. GHAZZALI, "What do we retain from a classification ? An experiment in image coding". Actes des Journées Internationales Analyses des Données et Apprentissage Symbolique-Numérique. Édition E. Diday et Y. Lechevallier, Nova Sciences Publishers, Inc. (New York), pp. 18-20 septembre 1991.
- [20] I.C. LERMAN, P. PETER, "Élaboration et logiciel d'un indice de similarité entre objets d'un type quelconque. Application au problème du consensus en classification", Publication interne n° 262, Irisa, Rennes, juillet 1985, 72 pages. Cette publication a conduit à un article dans : Rev. Stat. Appl. , n° 35, pp. 39-60, 1987.
- [21] I.C. LERMAN, P. PETER, H. LEREDDE, "Principes et calculs de la méthode implantée dans le programme CHAVL (Classification Hiérarchique par Analyse de la Vraisemblance des Liens)", La revue Modulad, n° 12 et n° 13.
- [22] G. LINDE, A. BUZO, R.M. GRAY, "An algorithm for vector quantizer", IEEE transactions on communications, Vol. COM-28, N° 1, pp. 84-95, 1980.
- [23] T. MURAKAMI et al., "Vector Quantization of color images", ICASSP 86, Tokyo.
- [24] N. M. NASRABADI, R.A. KING, "Image coding using vector quantization: a review". IEEE transactions on communications, Vol. 36, N° 8, pp. 957-971, August 1988.
- [25] F. NICOLAÛ, "Critérios de análise classificatoria hierárquica baseados na função de distribuição", Thèse, Faculté des sciences de Lisbonne, février 1981.
- [26] M.T. ORCHARD, "Color quantization of images". IEEE transactions on signal processing, Vol. 39, N° 12, pp. 2677-2721, December 1991.
- [27] M. OUALI-ALLAH, "Analyse en préordonnances des données qualitatives. Applications aux données numériques et symboliques". Thèse de l'Université de Rennes I, décembre 1992.
- [28] P. PETER, "Méthodes de classification hiérarchique et problèmes de structuration et de recherche d'information assistées par ordinateur". Thèse de l'Université de Rennes I, mars 1987.
- [29] J.H. Jr WARD, "Hierarchical grouping to optimize an objective function". J.A.S.A., Vol. 58, pp. 236-244, 1963.
- [30] J. ZIV, A. LEMPEL, "Compression of individual sequences via variable-rate coding". IEEE Transactions on information theory, Vol. IT-24, N°5, September 1978.

#### OUVRAGES:

- [1] N. AHMED and RAO, "Orthogonal Transform for Digital Signal Processing", Springer-Verlag, New York, 19...
- [2] T. BERGER, "Rate distortion theory", Prentice-Hall, Engelwood Cliffs, NJ, 1971.
- [3] M. KUNT, M. BELLANGER, C. GUEGUEN, F. DE COULON, "Traitement de l'information", Presses polytechniques et universitaires romandes, 1991.
- [4] B. MARTI et al., "TÉLÉMATIQUE, techniques, normes, services", collection scientifique des télécommunications, DUNOD, 1990.
- [5] E. ROUBINE, "Introduction à la théorie de la communication", Masson, Paris, 1970.
- [6] C.E. SHANNON, "The mathematical theory of communication, Bell System technical journal, july 1948, pp. 379-423 et oct. 1948, pp. 623-656.

#### REVUES:

- Le courrier du CNRS, dossiers scientifiques, "Signaux et Images", N°77, juin 1991.
- W.D. PENNEBAKER, J.L. MITCHELL, G. LANGDON, R.B. ARPS, "An overview of the basic principles of the Q-Coder Binary Arithmetic Coder", IBM J. Res. Develop. 32, n°6, pp. 717-726, november 1988.
- A. D. WYNER, "Fundamental limits in information theory", Proceedings of the IEEE, vol. 69, N)2, pp. 239-251, february 1981.

#### NORMES:

- ISO/IEC JTC1/SC29/WG12, IS 10918-1, "Technologies de l'information - Compression numérique et codage des images fixes de nature photographique - partie 1: Spécification et directives"; DIS 10918-2, "partie 2: Test de conformité". AFNOR, Tour Europe, 92180 La Défense.
- ISO/IEC JTC1/SC29/WG11, DIS 11172, "Coding of moving pictures and associated audio - for digital storage media at up to about 1.5 Mbit/s. ("MPEG")", AFNOR, Tour Europe, 92180 La Défense.
- ISO/IEC JTC1/SC29/WG9, IS 11544, "Progressive bi-level Image compression, revision 4.1, ("JBIG")", AFNOR, Tour Europe, 92180 La Défense.

