

Compression d'images par modélisation des coefficients TCD en lois mélange et quantification adaptative dans l'environnement JPEG

Image Compression using Modelling of DCT Coefficients with Mixture Laws and Adaptive Quantization in JPEG Environment

par Richard GRISEL, Florence ASTRADE

CPE – Lyon, Laboratoire LISA – CNRS EP 0092
43, boulevard du 11 Novembre 1918, BP 2077
F-69616 Villeurbanne cedex

résumé et mots clés

Dans cet article sont présentés des résultats obtenus sur la recherche des lois de distribution suivies par les coefficients de la Transformée en Cosinus Discrète (TCD) d'images ainsi qu'une première approche de quantification scalaire adaptative.

Dans la littérature, de nombreux travaux ont déjà été menés sur l'étude des distributions des coefficients TCD. Les auteurs trouvent des gaussiennes, des laplaciennes ou des lois de Cauchy en basant leur décision sur le test d'adéquation de Kolmogorov–Smirnov et en choisissant la loi qui donne le plus petit résultat au test. Afin de pouvoir justifier de la décision sur la loi, nous avons utilisé une batterie de tests et comparé les résultats à des seuils. Finalement, en utilisant le test de Cramer qui s'est avéré le plus sélectif, nous montrons que les coefficients TCD suivent une loi gaussienne ou somme de deux à quatre gaussiennes.

Nous pouvons alors calculer une matrice de quantification adaptative et comparer les images reconstruites après quantification adaptative et après quantification par la norme JPEG. Pour une qualité comparable (PPSNR et histogrammes équivalents), ces images ont un meilleur taux de compression par quantification adaptative.

Compression d'image, quantification adaptative, JPEG, lois mélange.

abstract and key words

In this paper, we introduce a new modelling technique for statistical distributions of DCT coefficients and its application to image compression. By modelling the statistics of the AC DCT coefficients we derive an approach for scalar adaptive quantization.

A lot of work can be found in the literature concerning these distributions. Depending on authors, they have been modeled as Gaussian, Laplacian or Cauchy laws. The previous results are mainly based on the use of the Kolmogorov–Smirnov fit-test, the decision being taken for the law giving the smallest result to the test.

The paper describes the building of statistic tools created in order to define thresholds for the decision. Several fit-tests are computed and finally, the Cramer test appears to be the most selective one.

The results we obtain show that statistical distributions of AC DCT coefficients are either Gaussians or mixtures of two to four Gaussians. We can then compute an adaptive quantization matrix and compare images reconstructed with adaptive quantization and with JPEG norm.

For identical quality (equivalent PPSNR and histograms), reconstructed images have a higher compression rate when adaptive quantization is used.

Image compression, adaptive quantization, JPEG, finite mixture distribution.

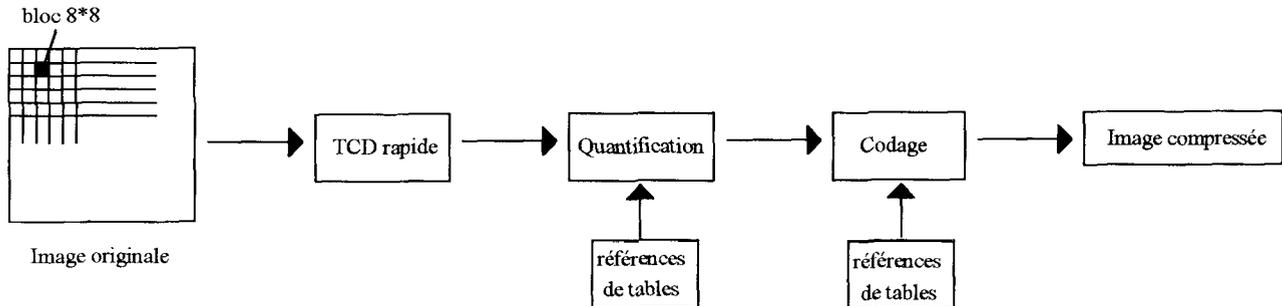


Figure 1. – Principe général de la compression d'images basée sur la TCD.

1. introduction

De nos jours, les images constituent un support d'information très puissant. Cependant, leur utilisation intensive amène d'importants problèmes d'archivage et de transmission, vu la quantité d'information qu'elles contiennent.

Plusieurs méthodes de compression d'image ont déjà été développées pour résoudre ce problème. On peut citer par exemple la décomposition fractale, des méthodes utilisant une transformation de l'image dans le domaine spatio-temporel (transformée en ondelettes) ou dans le domaine fréquentiel comme la transformée de Karhunen-Loeve ou la transformée cosinus discrète (TCD) [Jain][Mathieu]. C'est cette dernière méthode qui va être développée ici.

Nous donnerons, pour illustrer les résultats obtenus sur la modélisation, des exemples d'application sur des images test de différentes modalités de type médical. Les images de type médical ont été choisies en liaison avec les travaux de recherche appliquée menés par le laboratoire dans le domaine des PACS (Picture Archiving and Communication Systems) [Abouchi][Eude92]. Pour ces travaux dans lesquels il y a mise en œuvre de système d'information, il n'est pas question pour l'instant d'envisager des méthodes de compression non réversibles, sachant que pour JPEG, seule la compression réversible avec codage entropique a reçu l'approbation des comités de standardisation pour son emploi en transmission et archivage de l'imagerie médicale (norme DICOM 3). Notre but est de démontrer qu'il est possible, à partir d'une modélisation précise des coefficients, d'envisager des taux de compression élevés, donc avec perte d'information, en restant dans le standard JPEG avec quantification scalaire.

Afin d'améliorer la compression proposée par JPEG, nous devons réaliser une quantification adaptée à l'image étudiée, c'est-à-dire aux lois de distribution suivies par les coefficients obtenus après transformation de l'image. Cette étude commence donc par la détermination de ces lois dans la mesure où les résultats trouvés dans la littérature ne nous paraissent pas satisfaisants. Afin d'obtenir des résultats tout à fait fiables, nous avons eu recours à une batterie de tests d'ajustement de lois qui nous a permis

de montrer que les coefficients TCD suivent soit une loi gaussienne soit une loi somme de deux à quatre gaussiennes.

Nous avons alors pu déterminer une matrice de quantification adaptée à ces lois et comparer les images reconstruites après quantification adaptative avec celles obtenues par JPEG.

Parallèlement à cette étude, d'autres travaux sont en cours pour utiliser les lois statistiques des coefficients TCD afin d'établir d'autres méthodes de quantification adaptatives, scalaire [Commaniciu95a] ou vectorielle [Commaniciu95b].

2. compression d'image par la TCD

Toutes les méthodes de compression basées sur l'utilisation de la Transformée Cosinus Discrète (TCD) suivent trois étapes : transformation de l'image ou de sa différence prédite, quantification des coefficients transformés et codage (figure 1).

La transformation est appliquée sur des blocs de 8*8 pixels; elle est définie par :

$$F(u, v) = \frac{2c(u)c(v)}{N} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) \cos\left(\frac{u\pi}{2N} (2x+1)\right) \cos\left(\frac{v\pi}{2N} (2y+1)\right)$$

avec $c(0) = \frac{1}{\sqrt{2}}$ et $c(u) = 1$ si $u \neq 0$.

N est la taille des blocs (on choisit $N = 8$); x, y sont les coordonnées dans le domaine spatial et u, v les coordonnées dans le domaine des fréquences.

Chaque bloc est donc composé de 64 coefficients. Le coefficient $(0,0)$ est noté DC et représente la moyenne des intensités du bloc, les autres sont notés AC . Les coefficients comportant l'information globale de l'image sont localisés en basse fréquence alors que ceux situés dans les plus hautes fréquences sont souvent assimilés à du bruit (figure 2).

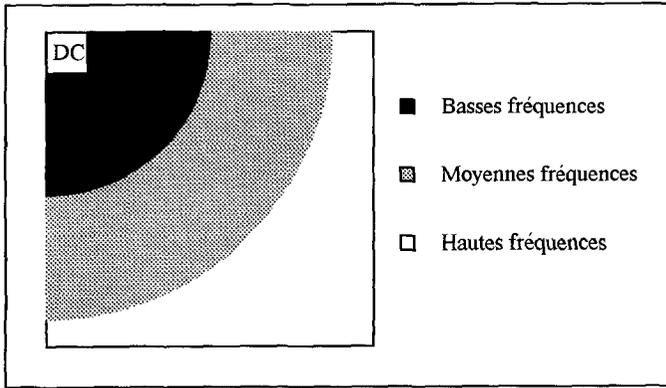


Figure 2. – Distribution fréquentielle des coefficients TCD.

L'étape de quantification est la seule partie irréversible du processus de compression. C'est elle qui induit les pertes subies lors de la compression. Ainsi, elle doit être optimisée pour à la fois améliorer le taux de compression et garder une bonne restitution de l'image. Une méthode d'optimisation pourrait être d'adapter la quantification à l'image, plutôt que de définir une quantification identique pour toutes les images. Cette adaptation peut être réalisée en utilisant les caractéristiques statistiques des coefficients TCD.

La dernière étape, le codage, ne sera pas développée ici.

3. modélisation des lois suivies par les coefficients TCD

3.1. distributions statistiques possibles

Jusqu'à présent, la littérature donne beaucoup d'informations différentes sur la modélisation des distributions statistiques suivies par les coefficients TCD. Tous les auteurs sont d'accord sur le fait que le coefficient *DC* suit une loi gaussienne, mais plusieurs théories différentes ont été écrites sur les coefficients *AC*. Ainsi, par exemple, Tesher, Murakami, Reininger et Bellifemine proposent des laplaciennes tandis que Baskurt parle de lois de Cauchy. De même, Lee, Kim, Rowberg et Riskin pensent que les coefficients *AC* obtenus par différence d'images suivent des distributions gaussiennes généralisées [Baskurt] [Bellifemine] [Lee] [Murakami] [Reininger] [Tescher]. Tous ces résultats ont été obtenus en utilisant le test d'adéquation de Kolmogorov-Smirnov.

Ainsi, dans la littérature, il a été écrit que les coefficients TCD peuvent être modélisés par les trois lois suivantes :

– une gaussienne $N(m, \sigma)$:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right)$$

– une laplacienne $L(m, \beta)$: $f(x) = \frac{1}{2\beta} \exp\left(-\frac{|x-m|}{\beta}\right)$

– une loi de Cauchy $C(\alpha, \beta)$: $f(x) = \frac{1}{\pi\beta} \frac{1}{1 + \frac{(x-\alpha)^2}{\beta^2}}$

Cependant, une étude précédente a montré que, pour des applications liées à la segmentation d'images, les images peuvent être modélisées par des mélanges de gaussiennes [Cherifi94]. Ainsi, puisque la TCD est une transformation orthogonale, nous avons pensé que les coefficients TCD pouvaient également suivre une loi mélange de gaussiennes. De plus, comme le montre la figure 3, certains coefficients TCD n'ont pas une distribution symétrique, ils ne peuvent donc pas être modélisés par une loi symétrique telle qu'une gaussienne ou une loi de Laplace. C'est pourquoi il a été envisagé d'utiliser des sommes de lois mélange qui, grâce aux différentes moyennes des lois, permettent de modéliser des distributions dissymétriques.

Afin de ne pas privilégier les lois gaussiennes, nous avons décidé de considérer également des lois mélange de laplaciennes.

Ces deux lois sont définies par :

$$f(x) = \sum_{i=1}^M p_i f_i(x)$$

avec p_i : probabilité et f_i : gaussienne ou laplacienne.

3.2. tests d'adéquation

Une manière de savoir si une loi empirique correspond ou non à une loi théorique donnée est d'utiliser des tests appelés tests

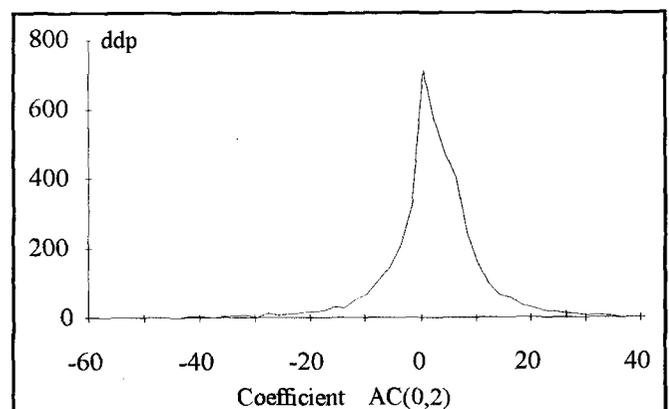


Figure 3. – Densité de probabilité empirique du coefficient $AC(0,2)$ d'une image d'os. Sur cet exemple, on peut constater que la distribution statistique n'est pas symétrique, elle ne peut donc pas être modélisée par une loi symétrique telle qu'une gaussienne ou une laplacienne, d'où le choix de la modélisation par des lois mélange de gaussiennes.

d'adéquation ou d'ajustement de lois. Il s'agit d'estimer les paramètres des différentes lois théoriques possibles afin de pouvoir calculer leur fonction de répartition puis d'utiliser les tests pour déterminer la meilleure loi théorique associée à l'échantillon considéré.

Ainsi, pour vérifier si un échantillon $F(x)$ provient d'une distribution de fonction de répartition connue $F_0(x)$, on pose les hypothèses suivantes :

$H_0 : F_0(x) = F(x) \Rightarrow$ la loi empirique F est assimilable à la loi théorique F_0 .

$H_1 : F_0(x) \neq F(x) \Rightarrow$ les deux lois ne sont pas comparables.

Pour déterminer l'hypothèse vraie parmi ces deux avec un risque α de rejet à tort de l'hypothèse H_0 , on dispose de tables statistiques. Elles donnent un seuil dépendant de α . Pour accepter l'hypothèse H_0 , le résultat x du test ne doit pas dépasser ce seuil :

si $x > \text{seuil}_\alpha$ alors on rejette H_0 avec un risque α de se tromper :

→ loi empirique \neq loi théorique

si $x < \text{seuil}_\alpha$ alors on accepte H_0 :

→ loi empirique = loi théorique

Les études précédentes basaient leurs résultats uniquement sur le test d'adéquation de Kolmogorov-Smirnov. Les décisions sur les lois étaient prises en considérant que la loi théorique suivie par le coefficient TCD est celle qui donne le résultat le plus faible à ce test.

Afin d'obtenir des résultats plus fiables, nous avons décidé d'utiliser une batterie de tests d'ajustement de lois paramétriques (test D , Kurtosis et test de symétrie) et non paramétriques (tests de Kolmogorov-Smirnov, de Cramer-von-Mises, d'Anderson-Darling et de Watson) [Cherifi87] [Stephens]. Le choix de la loi est alors réalisé en considérant qu'une loi empirique suit une loi théorique donnée si les résultats aux tests sont inférieurs à des seuils. Ces seuils sont donnés pour chaque test dans des tables. N'ayant pu trouver toutes les tables de seuils dont nous avons besoin dans la littérature, nous avons dû construire des outils statistiques pour les établir. Ainsi, nous avons simulé les différents tests par une méthode Monte Carlo pour pouvoir ajuster les seuils pour une loi gaussienne et une loi de Laplace.

Nous avons pour cela généré 1000 échantillons (chacun de taille 500) de lois normales centrées réduites. Nous avons alors calculé les résultats aux tests pour trois valeurs du risque de rejet à tort α (5%, 10% et 15%). Comme on peut le voir dans le tableau 1, les valeurs de seuils ainsi obtenues sont tout à fait comparables à celles que l'on a pu trouver dans la littérature [Stephens]. Les différences qui apparaissent sont dues à la différence de taille entre les échantillons utilisés. Nous en avons alors conclu que la méthode utilisée est fiable. Ceci nous a permis d'établir une table de seuils pour la loi de Laplace (tableau 2), table inexistante dans la littérature.

Tableau 1. – Valeurs de seuil pour une loi normale centrée réduite $N(0, 1)$ (1000 échantillons) et valeurs trouvées dans la littérature (1000 échantillons)

Nom du test d'adéquation	Valeur α	Seuils obtenus	Seuils trouvés dans la littérature
KOLMOGOROV	5%	$x = 0,039498$	0,042943
	10%	$x = 0,035403$	
	15%	$x = 0,033962$	0,020807
CRAMER	5%	$x = 0,133408$	0,126
	10%	$x = 0,107436$	0,104
	15%	$x = 0,094768$	0,091
ANDERSON	5%	$x = 0,768421$	0,787
	10%	$x = 0,626385$	0,656
	15%	$x = 0,552691$	0,576
WATSON	5%	$x = 0,128064$	0,116
	10%	$x = 0,100178$	0,096
	15%	$x = 0,087248$	0,085
D TEST	5%	$x = 1,194004$	1,411
	10%	$x = 0,907934$	1,119
	15%	$x = 0,709708$	
KURTOSIS	5%	$x = 3,363713$	3,26
	10%	$x = 3,264188$	
	15%	$x = 3,189641$	

Tableau 2. – Valeurs de seuils pour une loi de Laplace $L(0, 1)$ (1000 échantillons)

Nom du test d'adéquation	Valeur α	Seuils obtenus
KOLMOGOROV	5%	$x = 0,0533726$
	10%	$x = 0,048159$
	15%	$x = 0,045110$
CRAMER	5%	$x = 0,299904$
	10%	$x = 0,228705$
	15%	$x = 0,192915$
ANDERSON	5%	$x = 1,470545$
	10%	$x = 1,185662$
	15%	$x = 0,990446$
WATSON	5%	$x = 0,217766$
	10%	$x = 0,175658$
	15%	$x = 0,150642$
KURTOSIS	5%	$x = 8,190862$
	10%	$x = 7,372080$
	15%	$x = 6,904776$

Dans la mesure où, pour les coefficients TCD, la taille des échantillons est de 4096 et non de 500, nous avons dû regarder si les tests évoluent en fonction de la taille de l'échantillon.

Notre analyse donne les résultats suivants :

- les valeurs des tests de Cramer-von-Mises, d'Anderson et de Watson ne changent pas quelle que soit la taille de l'échantillon;
- le test de Kolmogorov-Smirnov varie en \sqrt{N} (N étant la taille de l'échantillon);
- le test D varie lui aussi en \sqrt{N} et n'est, de toute façon, valable que pour une loi normale.

3.3. estimation des paramètres des distributions

Afin d'estimer les différents paramètres des distributions des coefficients transformés (moyenne, écart type et probabilité), nous avons utilisé l'algorithme EM stochastique [Redner][Celeux]. Cet algorithme itératif calcule les paramètres au sens du maximum de vraisemblance.

4. résultats

4.1. introduction

Nous avons validé la méthode en utilisant des images provenant du CERF (Collège des Enseignants en Radiologie Français) correspondant à diverses modalités (os, angiographie, poumon, scanner, IRM) [Images Types], les résultats obtenus montreront que la méthode est extensible à n'importe quel type d'image.

Lorsque nous avons appliqué sur les images transformées l'algorithme SEM et les tests avec un taux de rejet de 5%, il est apparu qu'aucun coefficient ne pouvait être modélisé par une laplacienne ou une loi de Cauchy. Les seules lois que nous ayons obtenues étaient des lois mélange d'une à quatre gaussiennes. Ceci peut être justifié par les valeurs données par le Kurtosis. En effet, la moyenne des 64 valeurs obtenues au test du Kurtosis est toujours bien supérieure à six (valeur du Kurtosis pour une laplacienne). Ainsi, les 64 distributions ne peuvent pas être uniquement des laplaciennes ou des gaussiennes.

Dans la mesure où tous les tests donnent les mêmes résultats, i.e. un mélange de gaussiennes, nous avons décidé, dans une seconde étape, de n'utiliser que trois tests :

Nous avons tout d'abord utilisé le test de Kolmogorov-Smirnov car c'est le test le plus connu. La fonction de répartition de ce test est tracée sur la figure 4. Sa formulation discrète est la suivante :

$$t = \max_{1 < i < n} |F_X(x_i) - F(x_i)|$$

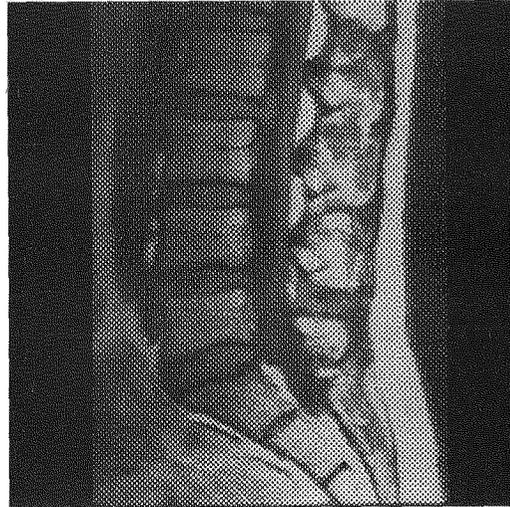


Image IRM (10).

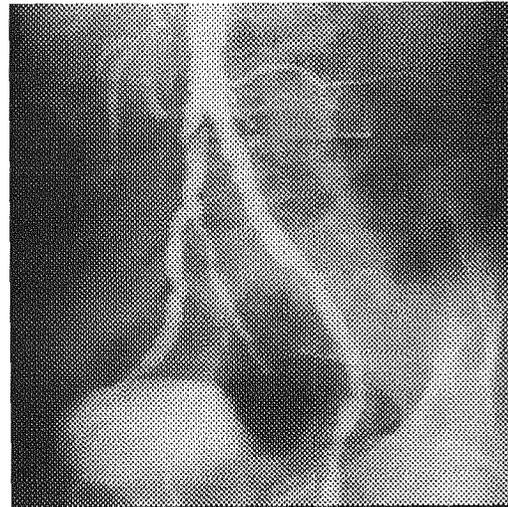


Image d'Angiographie (12).

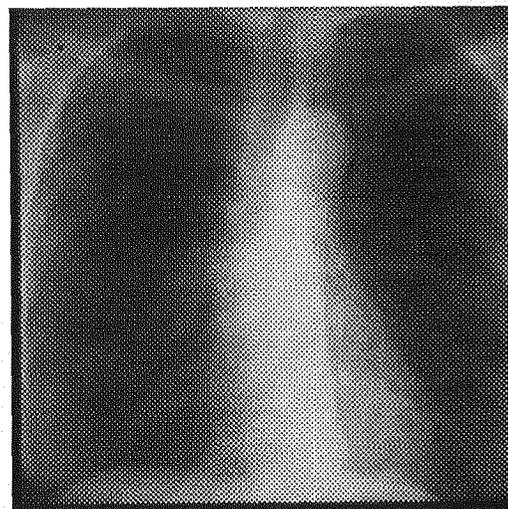


Image de poumon (15).

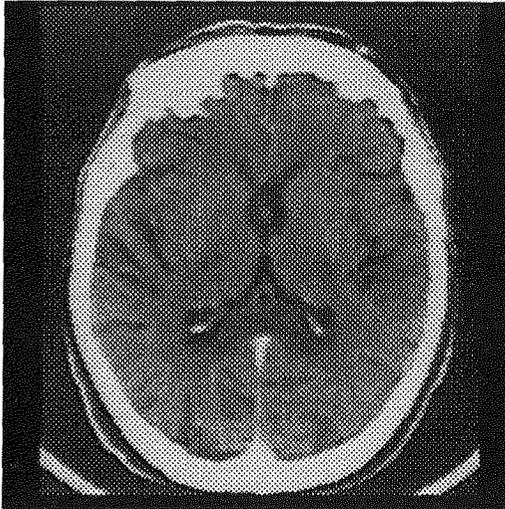


Image de scanner (17).

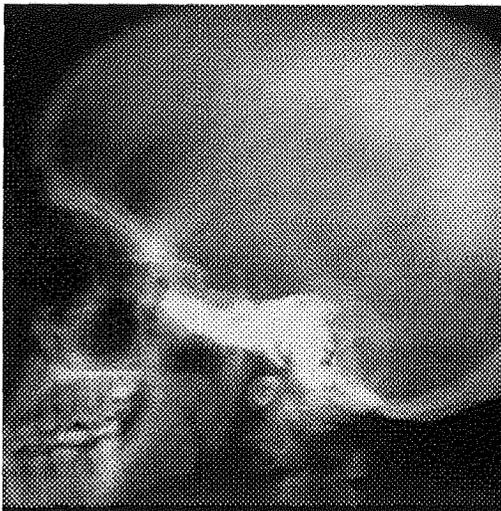


Image d'os (24).

Images types : Les numéros entre parenthèses sont les repères utilisés dans le tableau 8.

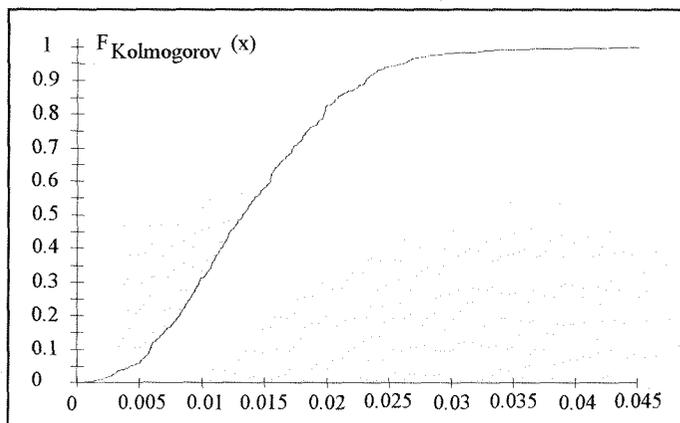


Figure 4. – Fonction de répartition du test de Kolmogorov-Smirnov.

n représente la taille de l'échantillon. Pour ce calcul, les valeurs x_i de l'échantillon doivent être ordonnées ($x_1 < x_2 < \dots < x_n$).

Ce test mesure une distance, il donne pour résultat la plus grande différence existant entre une fonction de répartition empirique F_X et une fonction théorique F .

Le deuxième test que nous utilisons est le test de Cramer-von-Mises. Si x est une variable aléatoire continue, ce test est défini par :

$$n\omega^2 = \int_{\mathcal{R}} [F_n^*(x) - F(x)]^2 dF(x).$$

$n\omega^2$ calcule la surface entre une fonction de répartition empirique F_n^* et une répartition théorique F .

On peut montrer que pour une variable aléatoire discrète, le calcul du test de Cramer devient :

$$n\omega^2 = \frac{1}{12n} + \sum_{i=1}^n \left(\frac{2i-1}{2n} - F(x_i) \right)^2$$

La figure 5 montre la fonction de répartition du test de Cramer. Ce test a été choisi pour les quatre raisons suivantes :

- il ne varie pas avec la taille de l'échantillon ;
- il a donné de très bons résultats à l'étude qui a été réalisée précédemment sur la puissance des tests ;
- dans la mesure où le test de Kolmogorov-Smirnov mesure la plus grande distance entre deux fonctions de répartition, il dépend beaucoup des valeurs extrêmes. Au contraire, le test de Cramer-von-Mises donne un résultat qui prend en compte tous les points en même temps. Il est ainsi moins sensible aux points aberrants ;
- c'est le test le plus sélectif parmi tous ceux que nous avons utilisés.

Nous allons enfin utiliser le test du Kurtosis. Il est en effet très intéressant puisque sa valeur donne une idée sur la loi suivie par un échantillon : si le Kurtosis vaut 3 alors l'échantillon suit une loi

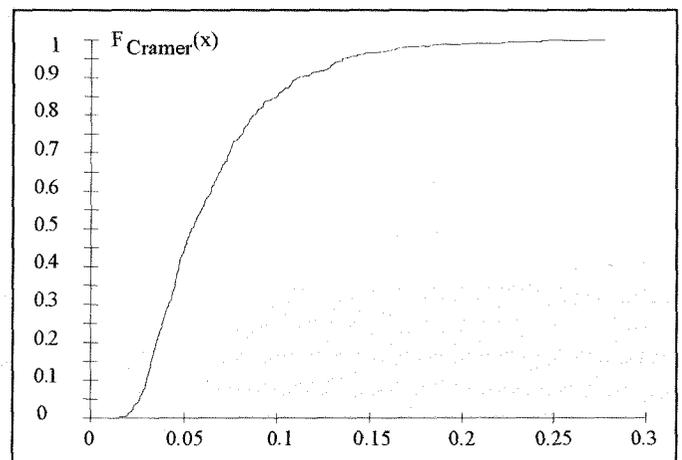


Figure 5. – Fonction de répartition du test de Cramer.

gaussienne, s'il vaut 6 il suit une laplacienne. Pour une variable aléatoire continue ce test est défini par :

$$\text{Kurtosis} = \frac{m_4}{(m_2)^2} \text{ avec } m_n = \int_{\mathcal{R}} f(x) \cdot x^n dx$$

Sa fonction de répartition est tracée figure 6.

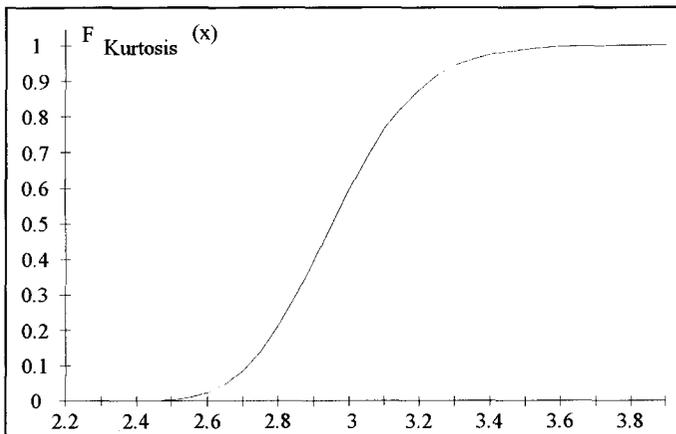


Figure 6. – Fonction de répartition du test de Kurtosis.

Il est important de préciser que l'utilisation de ces tests d'adéquation doit être réalisée sur des variables aléatoires indépendantes. Une décorrélation complète serait obtenue en utilisant la transformée de Karhunen–Loeve [Neagoe], cependant, l'utilisation de cette transformation est limitée par le fait qu'elle demande une quantité de calculs très importante. On ne peut alors qu'utiliser des transformations orthogonales sous-optimales comme la TCD sur des blocs de faible dimension. Les coefficients obtenus par la TCD sont donc partiellement décorrélés, la corrélation restante n'est pas exploitée.

Ainsi, l'hypothèse de base utilisée est que les coefficients suivent des lois statistiques indépendantes. Une telle hypothèse est couramment employée dans la littérature traitant de la compression d'image par la TCD [Aizawa] [Bellifemine94] [Saito].

4.2. description de la démarche suivie

Ayant remarqué que les coefficients TCD ne suivaient ni des laplaciennes ni des lois de Cauchy, nous nous sommes focalisés sur des lois mélange de gaussiennes [Eude94].

Ainsi, pour chaque échantillon, nous avons appliqué trois fois l'algorithme SEM afin d'estimer les paramètres des trois lois possibles : une, deux ou trois gaussiennes (nous nous sommes limités à trois gaussiennes afin d'avoir moins de calculs à réaliser). Puis, pour chacune de ces lois, nous avons calculé les résultats des tests de Cramer et de Kolmogorov.

Nous avons finalement choisi la loi qui répondait aux contraintes suivantes :

- les résultats des deux tests doivent être inférieurs aux seuils correspondant à $\alpha = 5\%$ (i.e. inférieure à 0,133408 pour Cramer et 0,039498 pour Kolmogorov);

- si plusieurs lois acceptent la première contrainte, la loi choisie est celle qui possède le plus faible nombre de gaussiennes.

Pour beaucoup de coefficients, les deux tests (Cramer et Kolmogorov) donnaient la même loi. Dans les cas où deux lois différentes étaient choisies, nous avons décidé de faire confiance au test le plus limitant qui était dans tous les cas le test de Cramer. Ceci est justifié par les histogrammes tracés sur la figure 7. En effet, dans le cas de ce coefficient (coefficient 9 d'une image de scanner), le test de Kolmogorov indiquait qu'une somme de 2 gaussiennes était suffisante pour modéliser la distribution empirique du coefficient tandis que d'après le test de Cramer il fallait une somme de 3 gaussiennes. La figure 7 montre que le choix du test de Cramer est nettement plus satisfaisant que celui du test de Kolmogorov.

4.3. résultats caractéristiques (figures 8 à 10)

Les résultats obtenus pour une image de chaque modalité sont présentés figure 8.

On peut constater que très peu de coefficients ne satisfont pas aux conditions voulues, c'est-à-dire au plus à une somme de trois gaussiennes. Cependant, dans tous les cas, quatre gaussiennes ont toujours été suffisantes pour les modéliser. On peut remarquer que les images os, angiographie, poumon appartiennent à la même famille d'imagerie par projection directe, par opposition aux coupes IRM et scanner de la famille des images reconstruites. La modélisation obtenue pour l'image d'angiographie présente 2 gaussiennes pour certains coefficients haute fréquence, ceci provient probablement du fait que les hautes fréquences des blocs de cette image représentent un mélange de bruit quantique, de structures vasculaires très fines et de transition au niveau des os (à droite dans l'image). Les distributions relatives aux images de projection (*a, b, c*) sont plutôt dissymétriques alors que les coupes (*d et e*) ont une répartition quasi-symétrique.

Nous pouvons aussi noter que pour toutes les images, les coefficients basses fréquences suivent plutôt une somme de deux à quatre gaussiennes alors que presque tous les coefficients hautes fréquences ont pu être modélisés par une gaussienne : ils représentent le bruit de l'image. Nous donnons à titre indicatif dans le tableau 3, les valeurs obtenues pour la modélisation des coefficients 3 (u, v) = (0, 3), 13 (u, v) = (1, 6) et 17 (u, v) = (2, 3) de l'image angiographie. Nous donnons également figures 9 et 10 quelques courbes qui présentent le Kurtosis (figure 9) et les différentes valeurs du test de Cramer comparées au seuil à 5% pour une image de poumon (figure 10). La figure 9 confirme le fait que les coefficients TCD appartenant aux basses fréquences ne peuvent pas être modélisés par une loi gaussienne puisque la valeur du Kurtosis obtenue pour ces coefficients est très supérieure à 3.

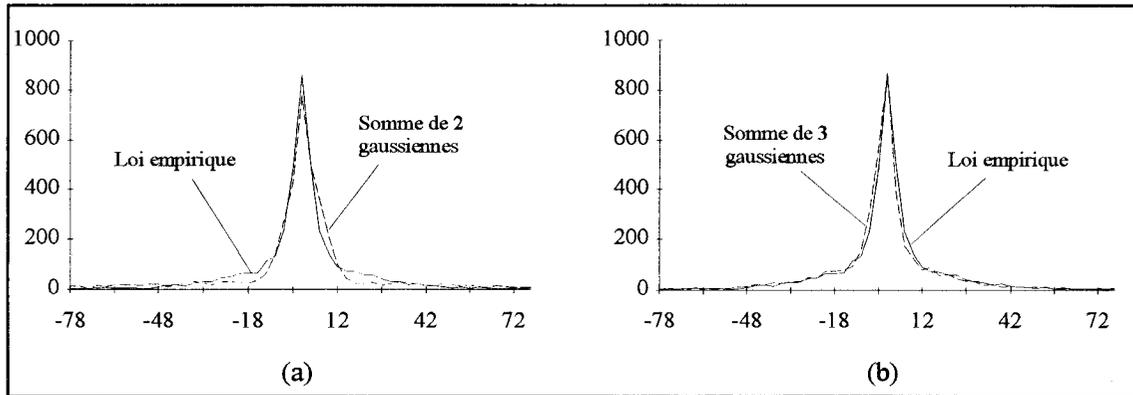


Figure 7. – Histogramme du coefficient 9 d'une image scanner : (a) comparé à la loi théorique somme de 2 gaussiennes choisie par le test de Kolmogorov (b) comparé à la loi théorique somme de 3 gaussiennes choisie par le test de Cramer. On s'aperçoit que le test de Cramer est plus restrictif que celui de Kolmogorov, la loi qu'il choisit pour la modélisation est nettement meilleure que celle choisie par le test de Kolmogorov.

Tableau 3. – Valeurs des paramètres des lois mélange de gaussiennes pour les coefficients 3 (4 classes), 17 (3 classes) et 13 (2 classes) du bloc TCD de l'image angiographique, les coefficients sont présentés par «distance» ou fréquence radiale croissante par rapport au coefficient DC.

Fréquence radiale = $f_r = \sqrt{u^2 + v^2}$.

Coefficient	Fréquence radiale		Classe 1	Classe 2	Classe 3	Classe 4
$3(u, v) = (0, 3)$	3	p	0,49	0,41	0,08	0,02
		m	-1,03	-0,58	-2,19	63,21
		σ	13,6	12,72	78,5	869
$17(u, v) = (2, 3)$	3,61	p	0,51	0,37	0,12	
		m	-0,16	0,08	0,38	
		σ	26,28	7,86	196	
$13(u, v) = (1, 6)$	6,08	p	0,51	0,49		
		m	-0,38	0,39		
		σ	8,09	12,71		

Au contraire, pour les coefficients situés en haute fréquence, les valeurs du Kurtosis peuvent être assimilées à la droite représentant le Kurtosis pour une gaussienne. La comparaison réalisée dans la figure 10d entre les résultats du test de Cramer et le seuil de Cramer à 5% montre que tous les coefficients ont pu être modélisés par une loi qui répond aux critères demandés.

5. quantification

La quantification est l'étape la plus importante de la compression d'image. En effet, c'est cette partie du processus qui permet de réduire la taille des images à stocker. Cependant, c'est également durant cette partie que les pertes d'information sont générées. La norme JPEG (Joint Photographic Expert Group) [ISO] est la norme la plus connue pour la compression d'images fixes. Elle

propose une méthode de quantification identique pour toutes les images.

Afin d'améliorer le taux de compression obtenu par JPEG tout en diminuant les pertes, nous devons optimiser l'étape de quantification en l'adaptant à l'image à traiter.

5.1. quantification proposée par JPEG

Dans JPEG, la quantification est réalisée sur les blocs 8*8 de coefficients transformés : chacun des 64 coefficients TCD est quantifié par un quantificateur uniforme défini par :

$$FQ(u, v) = \text{arrondi} \left(\frac{F(u, v)}{Q(u, v)} \right)$$

où $FQ(u, v)$ est le coefficient TCD quantifié, $Q(u, v)$ est l'élément (u, v) de la matrice de luminance (tableau 4), et où « arrondi » représente l'entier le plus proche.

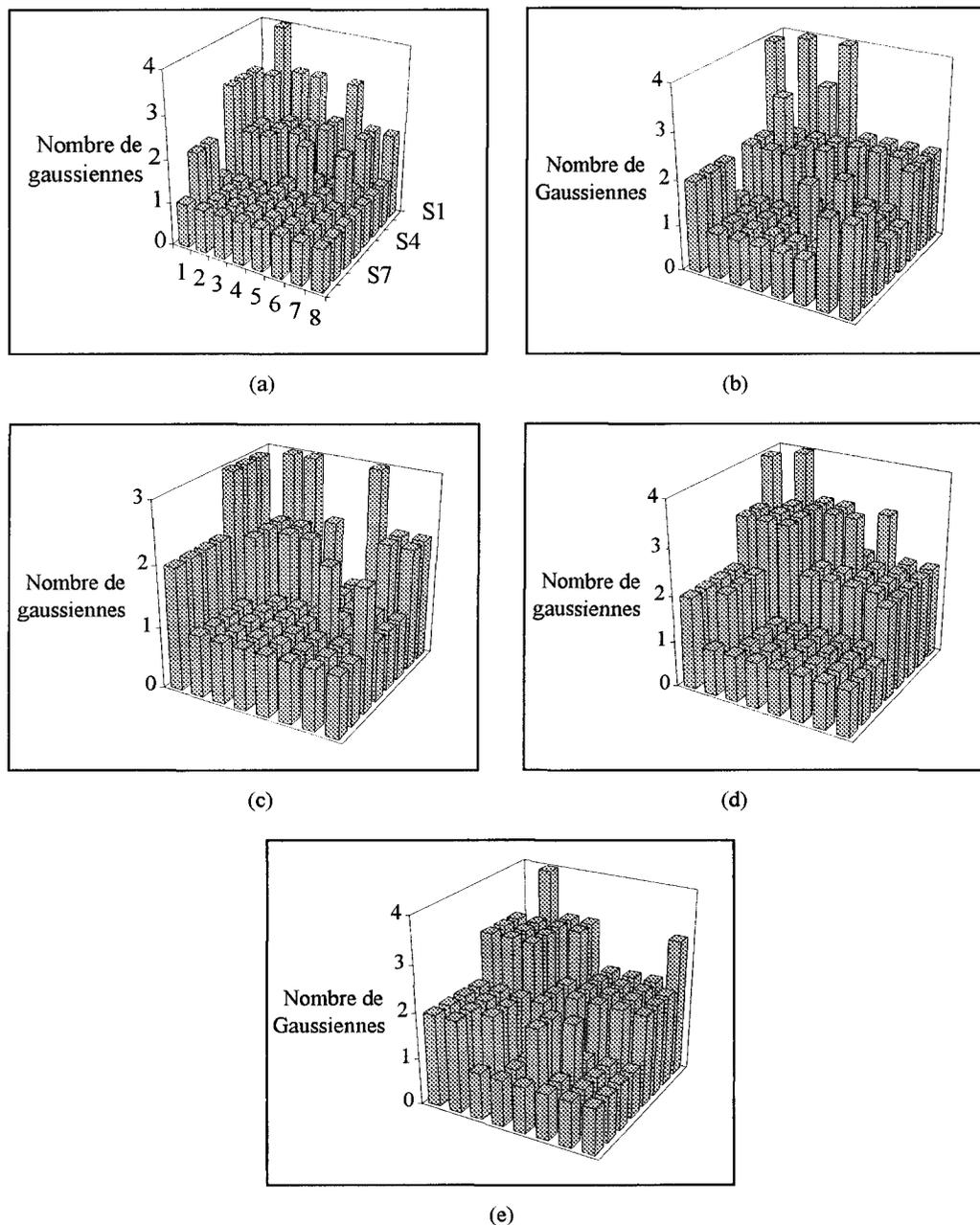


Figure 8. – Lois statistiques suivies par les coefficients TCD pour différentes images : (a) image d'os, (b) angiographie, (c) image de poumon, (d) IRM, (e) image de scanner. Chaque graphique représente le bloc 8*8 de coefficients TCD de l'image traitée. Les barres donnent le nombre de gaussiennes formant la distribution suivie par les coefficients, chaque gaussienne étant caractérisée par une moyenne, un écart-type et une probabilité d'apparition. Pour toutes ces images, 4 gaussiennes suffisent à modéliser tous les coefficients TCD. On peut remarquer que les coefficients basse-fréquence nécessite une modélisation pas une somme d'au moins 3 gaussiennes, tandis qu'il suffit d'une gaussienne pour modéliser les coefficients haute fréquence.

La déquantification est définie par :

$$FD(u, v) = FQ(u, v) \times Q(u, v),$$

où $FD(u, v)$ est le coefficient TCD déquantifié.

La matrice de luminance a été définie de manière expérimentale. Elle est utilisée de la même manière pour toutes les images.

Il est possible d'introduire un facteur appelé facteur de qualité qui permet de visualiser plus ou moins de défauts sur l'image reconstruite. Ce facteur, compris dans l'intervalle [1; 100], divise ou multiplie la matrice de luminance par une certaine valeur. Cependant, les dégradations qui apparaissent limitent le taux de compression. Adapter la quantification à l'image permet d'améliorer ce taux tout en gardant la qualité de l'image reconstruite.

Tableau 4. – Table de quantification de luminance de JPEG.

16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

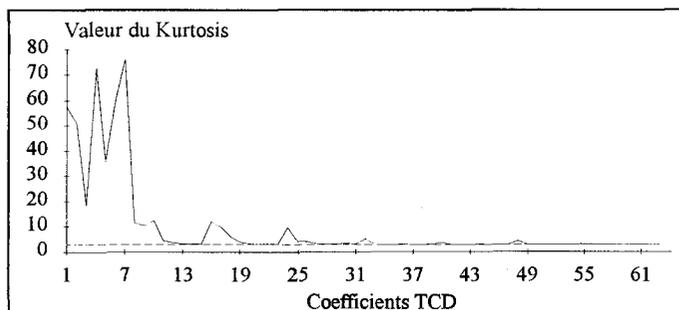


Figure 9. – Résultats du test de Kurtosis pour une image d'os. La ligne droite représente la valeur de Kurtosis pour une gaussienne (kurtosis = 3). On s'aperçoit donc que tous les coefficients haute-fréquence (coefficients environ supérieurs à 35) ont un Kurtosis très proche de 3, leur modélisation par une gaussienne est donc justifiée. Plus on s'approche des basses fréquences, plus le Kurtosis augmente, on vérifie ainsi que les coefficients proches du DC doivent être modélisés par un nombre élevé de gaussiennes.

5.2. quantification adaptative

Quantifier les coefficients TCD de manière uniforme consiste à diviser tous les coefficients du bloc par un facteur. Cependant, cette quantification doit tenir compte du fait que plus les coefficients sont dans les basses fréquences, plus ils sont riches en information. Ainsi, le facteur de quantification doit être différent selon la position du coefficient dans le bloc : il sera d'autant plus fort que les coefficients seront pauvres en information. Il sera donc plus grand en haute fréquence qu'en basse fréquence.

La matrice de luminance de JPEG répond à ce principe, cependant elle est la même pour toutes les images. Elle ne peut donc pas être optimale.

Adapter l'étape de quantification consiste à définir une table de quantification différente pour chaque image. Cette table doit être calculée à partir des lois de distribution des coefficients TCD déterminées dans les deux parties précédentes.

A partir des lois théoriques trouvées, nous avons déterminé un seuil S qui sépare l'histogramme d'un coefficient AC en trois parties (figure 11) :

- $\alpha/2\%$ des valeurs du coefficient sont supérieures à S ou inférieures à $-S$;
- $(1 - \alpha)\%$ des valeurs du coefficient sont comprises dans l'intervalle $[-S, S]$.

Ces seuils sont faibles pour les coefficients dont les valeurs sont fortement concentrées autour de zéro, donc pour les hautes fréquences. Ils sont plus élevés pour les autres.

Pour chaque coefficient AC qui suit la distribution $p(x)$, S est défini par :

$$\int_{-\infty}^{\infty} p(x)dx = 1 - \frac{\alpha}{2}$$

Le seuil S peut être calculé en considérant le fait que la surface entre 0 et S est égale à $(1 - \alpha)/2$ (en effet, la surface sous une fonction de distribution vaut 1) :

$$\int_0^S p(x)dx = \frac{1 - \alpha}{2}$$

Cette équation est résolue grâce à l'algorithme itératif de Newton et à la fonction d'erreur $erf(x)$.

La valeur de α dépend de l'image mais aussi de la position des coefficients dans le bloc. En effet, presque tous les coefficients hautes fréquences suivent une distribution gaussienne d'écart type très faible alors que les coefficients basses fréquences suivent une loi mélange de gaussiennes dont l'écart type total est beaucoup plus élevé (figure 12) (tableau 3). Ainsi, puisque les formes des distributions sont différentes selon la position des coefficients dans le bloc, nous avons décidé de choisir trois valeurs de α différentes selon la plage de fréquences dans laquelle est situé le coefficient (figure 3). Cette technique de différenciation des coefficients par plage « spectrale » est couramment utilisée dans les méthodes de quantification vectorielle pour la génération des codebooks [Commaniciu95b].

La matrice obtenue est appelée matrice des seuils, un exemple en est donné au tableau 5.

Tableau 5. – Matrice des seuils obtenue en adaptant la quantification aux distributions statistiques des coefficients TCD pour une image d'os.

DC	6.36	4.91	5.60	3.21	4.61	4.19	3.62
15.77	6.45	5.30	4.38	4.32	3.99	3.75	3.81
6.08	4.42	3.63	4.18	3.34	3.27	3.11	3.54
3.59	2.57	2.72	2.48	2.17	2.20	2.12	2.97
2.51	2.06	1.93	1.78	1.69	1.69	2.21	2.26
2.08	1.62	2.08	1.93	1.83	1.87	1.86	1.81
1.38	1.16	1.20	1.13	0.88	1.14	1.11	1.10
0.96	0.90	0.89	0.88	0.87	0.89	0.87	0.88

Le passage de la matrice des seuils $S(u, v)$ à la table de quantification adaptative $Q_{ad}(u, v)$ se fait par la relation suivante :

$$Q_{ad}(u, v) = \frac{Fe}{S(u, v)} \quad Fe \text{ est un facteur d'échelle.}$$

Cette relation permet d'avoir une valeur Q_{ad} faible lorsque S est élevé. Elle permet donc de conserver les valeurs significatives d'un coefficient et de mettre les autres à zéro.

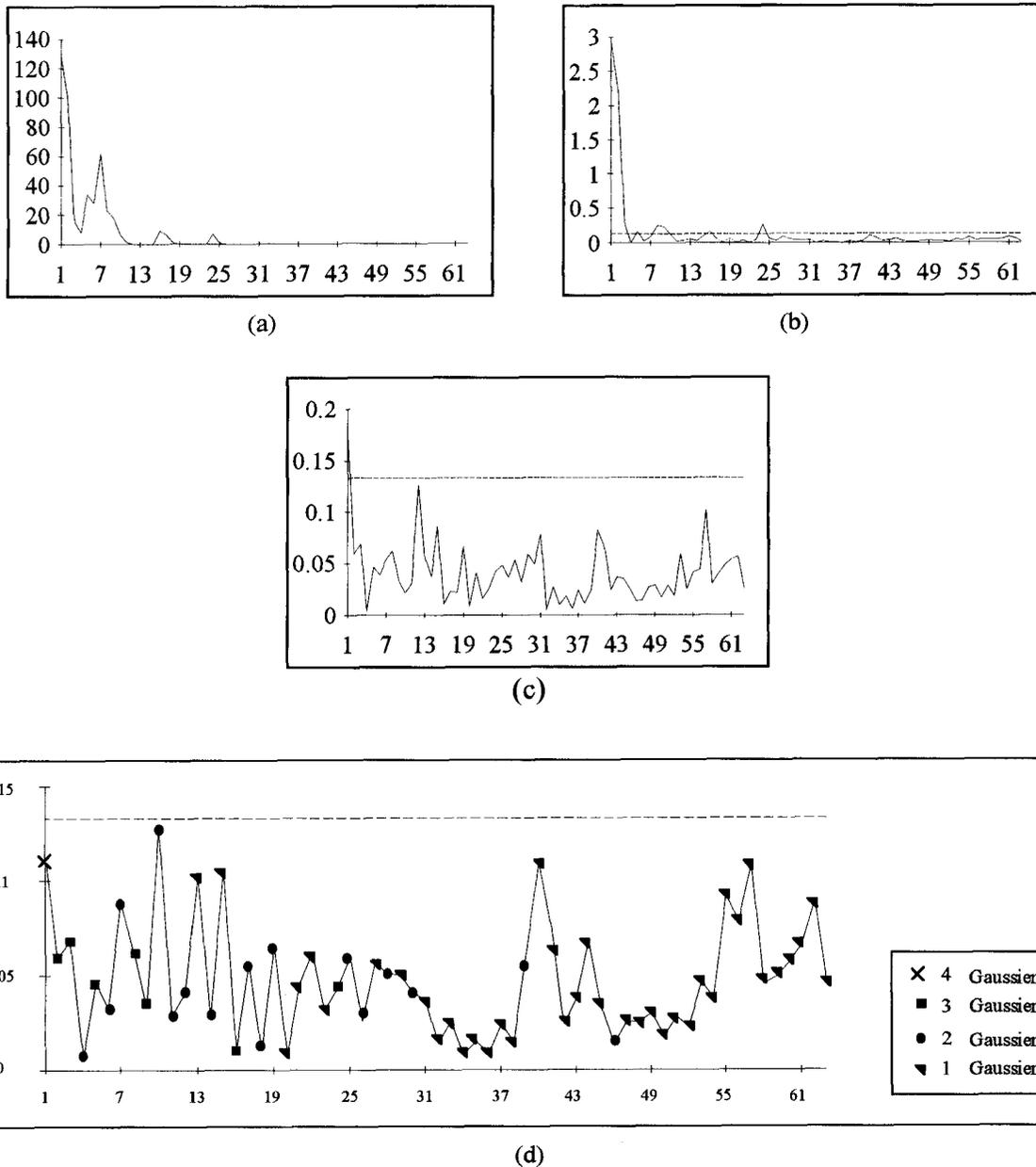


Figure 10. – Courbes « test de Cramer = $f(\text{coefficient TCD})$ » pour une image d'os pour :
 (a) une gaussienne, (b) deux gaussiennes, (c) trois gaussiennes.

La ligne droite représente le seuil de Cramer à $\alpha = 5\%$ (cramer = 0,133408). Le graphe (d) représente les différentes lois choisies pour chaque coefficient. On peut noter que le premier coefficient AC doit être modélisé par quatre gaussiennes pour que le résultat de Cramer soit en-dessous du seuil.

Le facteur Fe est déterminé de manière à avoir la table de quantification adaptative du même ordre de grandeur que celle de JPEG. Ainsi, nous avons choisi de fixer le maximum de la matrice adaptative à la même valeur que le maximum de JPEG, c'est-à-dire à 121. Le coefficient AC ayant le seuil le plus faible aura donc un pas de quantification égal à 121.

Nous obtenons alors notre facteur d'échelle :

$$Fe = \max(\text{JPEG}) \times \min(S(u, v)) \text{ avec } \max(\text{JPEG}) = 121$$

La table de quantification correspond à la matrice des seuils du tableau 5 est donnée au tableau 6.

La comparaison entre l'image originale et les deux images compressées (celle obtenue par JPEG et celle obtenue par quantification adaptative) est réalisée de deux manières différentes : d'une part par le calcul du rapport signal sur bruit crête à crête noté PPSNR (Peak to Peak Signal to Noise Ratio), d'autre part par la comparaison des histogrammes des trois images (figure 13). Le PPSNR est défini par :

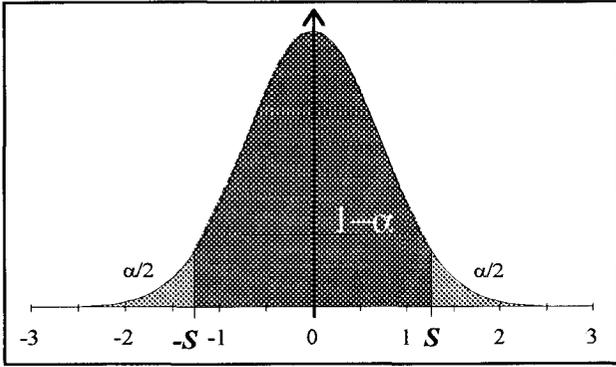
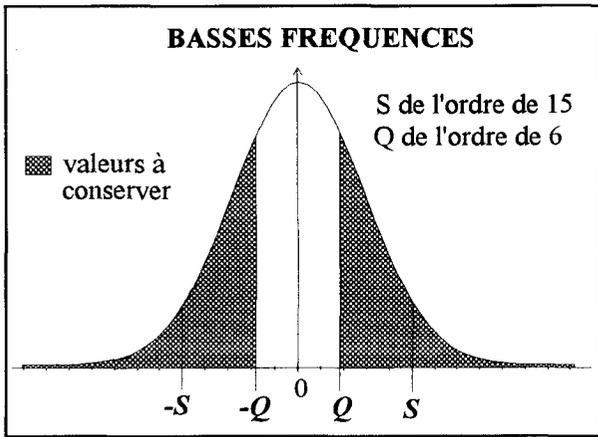
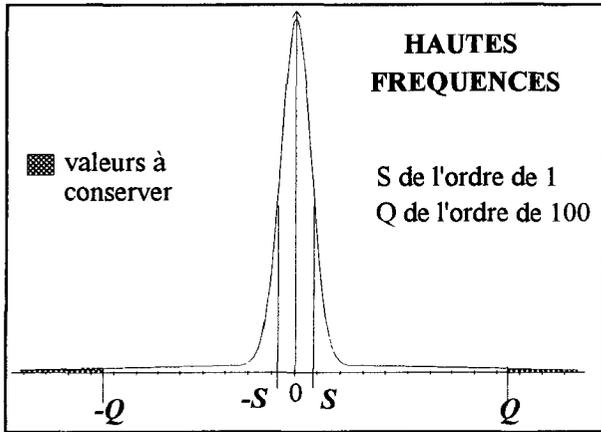


Figure 11. – Séparation de l'histogramme d'un coefficient en trois parties par seuillage à la valeur S .



(a)



(b)

Figure 12. – Valeurs d'un coefficient à conserver selon sa position dans le bloc.

S représente le seuil, il est donc tel que $\alpha\%$ des valeurs du coefficient sont supérieures à S ou inférieure à $-S$. Q représente la valeur du pas de quantification choisi, il vaut environ $100/S$.

La courbe (a) représente les coefficients situés en basse fréquence. Ils ont une distribution assez aplatie (grand-écart-type) donc la valeur de S est plutôt élevée. Au contraire, la courbe (b) représente les coefficients haute fréquence dont la ddp a un écart-type très faible, d'où le seuil S également faible, et un pas de quantification très élevé.

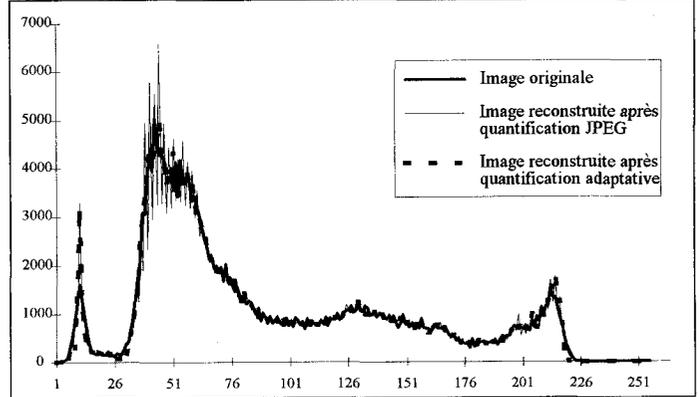


Figure 13. – Comparaison des histogrammes d'une image de poumon : Image originale et Images reconstruites après quantification JPEG et quantification adaptative.

On constate que dans les faibles niveaux de gris les deux images quantifiées sont identiques, mais différentes de l'image originale. Dans les niveaux de gris autour de 50, l'histogramme de l'image reconstruite après quantification JPEG fluctue beaucoup alors que l'histogramme de l'image obtenue après quantification adaptative suit celui de l'image originale. Enfin, dans les niveaux de gris élevés, les trois histogrammes sont semblables

Tableau 6. – Table de quantification obtenue en adaptant la quantification aux distributions statistiques des coefficients TCD pour une image d'os.

16	19	23	19	34	27	29	34
7	16	20	24	24	26	28	28
17	24	29	25	31	32	34	30
29	41	38	42	48	48	49	35
42	51	54	59	62	62	47	46
51	65	50	54	57	56	56	58
76	90	87	93	119	92	95	95
109	116	118	119	121	118	121	120

$$PPNSR = 10 \log_{10} \frac{255^2}{\sum_{x=0}^{N-1} \sum_{y=0}^{N-1} [f(x, y) - \hat{f}(x, y)]^2}$$

avec

$$\begin{cases} N \times N : \text{taille de l'image} \\ f(x, y) : (x, y)^{\text{ème}} \text{ pixel de l'image originale} \\ \hat{f}(x, y) : (x, y)^{\text{ème}} \text{ pixel de l'image reconstruite} \end{cases}$$

5.3. résultats

Le tableau 7 montre les résultats obtenus sur l'image os de taille 262 144 octets. Sur ce tableau :

Tableau 7. – Résultats obtenus pour une image d'os.

Première ligne : l'image est passée par la chaîne de traitement de JPEG mais n'a pas été quantifiée (la table de quantification était remplie de 1). Les pertes obtenues (taux de compression de 2.3) sont dues aux erreurs d'arrondi.

Deuxième ligne : les résultats sont obtenus après quantification avec la matrice de luminance.

Troisième ligne : l'image a été quantifiée par la matrice de quantification adaptative. on s'aperçoit par ce tableau que pour un fort gain en compression, on garde un PPSNR plus élevé en utilisant la matrice de quantification adaptative.

Type de traitement	PPSNR (db)	Taille de l'image compressée	Taux de compression
JPEG sans qualification	60.6	115 609	2.3
JPEG avec quantification	41.7	10006	26.2
Quantification adaptative	42.4	8181	32

Tableau 8. – Taux de compression obtenu sur chaque image testée avec JPEG et avec quantification adaptative.

Ces images ont été calculées de telle sorte que les PPSNR des images obtenues avec les deux quantification soient sensiblement équivalents. On constate alors que les taux de compression obtenus sont toujours supérieurs avec la quantification adaptative (gain en moyenne de 2.3 koctets). On constate que les valeurs de α (BF, MF, HF pour basses, moyennes et hautes fréquences) choisies diffèrent peu selon les modalités et selon les plages de fréquence.

Modalité	Numéro de l'image	Quantification adaptative						Quantification JPEG		
		α (%)			taille de l'image compressée	taux de compression	PPSNR	taille de l'image compressée	taux de compression	PPSNR
		BF	MF	HF						
IRM	10	40	20	5	11 895	22	38,96	15 691	16,7	37,65
Angiographie	12	20	20	5	12 927	20,3	39,61	14 739	17,8	39,13
Poumon	15	20	20	5	10 393	25,2	40,06	12 513	20,9	39,56
	16	20	20	5	9 996	26,2	39,2	13 175	19,9	38,6
	26	20	20	5	7 956	32,9	39,5	10 708	24,5	39
	27	30	20	5	11 005	23,8	40,1	13 779	19	39,2
	36	40	20	10	8 212	31,9	40	10 708	24,5	39,5
	47	40	20	5	9 659	27,1	41,2	12 209	21,5	40
Scanner	17	40	20	5	14 193	18,5	38,8	16 569	15,8	37,9
	19	40	20	5	19 933	13,2	38	21 499	12,2	37,4
	28	40	20	5	15 086	17,4	39,1	17 586	14,9	38
Os	24	20	20	5	8 181	32	42,4	10 006	26,2	41,7
	25	30	20	10	11 100	23,6	40,3	13 430	19,5	39,6
	44	30	20	10	9 971	26,3	40	13 132	20	39
	45	30	20	10	9 131	28,2	40,7	11 725	22,4	39,9

– la première ligne donne les résultats obtenus par JPEG sans quantification. Les pertes qui apparaissent représentent les erreurs d'arrondis. Cette image permet d'avoir une idée de la signification du PPSNR .

– la deuxième ligne donne les résultats obtenus par la matrice de luminance.

– la troisième ligne donne les résultats obtenus par quantification adaptative. Pour cette image, la matrice des seuils a été calculée avec $\alpha = 30\%$ pour les coefficients basses fréquences, $\alpha = 20\%$ pour les coefficients moyennes fréquences et $\alpha = 10\%$ pour les coefficients hautes fréquences.

Ce tableau montre que la quantification adaptative permet d'obtenir à la fois un gain en compression (gain de plus de 1.8 koctets) et un gain en qualité (amélioration de 0.7 dB du PPSNR).

Le tableau 8 présente tous les résultats obtenus sur les images test, pour un PPSNR sensiblement équivalent dans les deux méthodes de quantification.

Ce tableau ainsi que la comparaison des histogrammes des deux images quantifiées avec l'image originale permettent de conclure que pour une même qualité de l'image reconstruite, le taux de compression obtenu est meilleur par quantification adaptative que par la matrice de luminance.

6. conclusion

Dans cet article, de nouveaux outils statistiques ont été introduits pour déterminer les distributions statistiques des coefficients TCD. En effet, de tels outils nous permettent de prouver la validité des tests. Ainsi, utiliser le test de Cramer-von-Mises nous a permis d'obtenir des résultats fiables sur les distributions statistiques des coefficients TCD. Pour les images test utilisées, les coefficients TCD peuvent être modélisés par des lois mélange d'une à quatre gaussiennes : trois et parfois quatre gaussiennes pour les coefficients basses fréquences et une gaussienne pour les hautes fréquences.

Les exemples donnés montrent que cette modélisation permet d'améliorer l'étape de quantification de la compression proposée par la norme JPEG. On gagne en moyenne 2 octets pour chaque image, pour des PPSNR sensiblement équivalents.

L'inconvénient est, pour l'instant, que cette méthode est très longue au niveau de son exécution puisque pour les 64 coefficients du bloc il faut appliquer trois, voire quatre fois l'algorithme SEM, calculer trois ou quatre fois le résultat du test de Cramer et enfin utiliser l'algorithme itératif de Newton.

De prochaines recherches concerneront l'évaluation des images reconstruites utilisant des moyens de comparaison subjectifs et objectifs, l'introduction de la modélisation des coefficients TCD pour l'étape de codage. De plus, l'approche quantification scalaire donnée ici étant pour l'instant heuristique, nous avons à l'étude d'autres méthodes de quantification comme par exemple une quantification scalaire basée sur une méthode de type haute résolution [Commaniciu95a], et une quantification vectorielle adaptative [Commaniciu95b].

Remerciements

Nous voudrions remercier M. Hocine Cherifi et M. Thierry Eude pour leurs conseils à propos de la théorie statistique. Les auteurs souhaitent aussi signaler l'apport des différents rapporteurs qui ont permis, par leurs remarques et conseils, d'améliorer la présentation de ce travail.

BIBLIOGRAPHIE

- [Abouchi] N. Abouchi, M. Jourlin, O. Bohbot, C. Faurie, R. Grisel, «Network data bank image prototype : the R.S.I. project», *European Symposium on Advanced Networks and Systems (EOS/SPIE)*, vol. 2451, Amsterdam, mars 1995, pp. 360–369.
- [Aizawa] K. Aizawa, T. Natarajan, K.R. Rao, «Adaptive Discrete Cosine Transform coding with vector quantization for color images», *Proc. IEEE ICASSP'86*, Tokyo, Japon, avril 1986, pp. 985–988.
- [Baskurt] A. Baskurt, «Compression d'images numériques par la transformation cosinus discrète», *Ph.D de l'Université de Lyon*, juin 1989.
- [Bellifemine92] F. Bellifemine, A. Capellino, A. Chimienti, R. Picco, R. Ponti, «Statistical analysis of the 2D-DCT coefficients of the differential signal

for images», *Signal Processing : Image Communication*, vol. 4, n°6, 1992, pp. 477–488.

[Bellifemine94] F. Bellifemine, R. Picco, «Video signal coding with DCT and vector quantization», *IEEE trans. Commun.*, Vol. 42, n°2/3/4, février/mars/avril 1994, pp. 200–207.

[Celeux] Celeux, Diebolt, «Reconnaissance de mélange de densité et classification, un algorithme d'apprentissage probabiliste : l'algorithme SEM», *Rapport de recherche INRIA*, n°349, 1984.

[Cherifi87] H. Cherifi, «Evaluation des tests d'hypothèse», *GRETSI*, Nice, mai 1987, pp. 81–84.

[Cherifi94] H. Cherifi, R. Grisel, «Filter Expectation Maximisation Algorithm for Image Segmentation», *Proc. IEEE ICASSP'94*, Adelaïde, Australie, avril 1994, pp. V.133–V.136.

[Commaniciu95a] D. Commaniciu, R. Grisel, F. Astrade, «Medical image compression using mixture distributions and optimal quantization», *IASTED 95, Signal and Image Processing*, Las Vegas, USA, novembre 1995, pp. 89–92.

[Commaniciu95b] D. Commaniciu, R. Grisel, «Transform vector quantization with training set synthesis», *3rd Conf. Digital Image Computing, DICTA'95*, Brisbane, Australie, décembre 1995, pp. 139–144.

[Eude92] T. Eude, R. Grisel, «Commission archives : rapport final», *Centre Régional d'Informatique Hospitalière*, Rouen, 1992.

[Eude94] T. Eude, R. Grisel, H. Cherifi, R. Debrie, «On the distribution of the DCT coefficients», *Proc. IEEE ICASSP'94*, Adelaïde, Australie, avril 1994, pp. V.365–V.368.

[ISO] ISO/IEC JTC1/SC2/WG 10, «Initial draft for adaptative Discrete Cosine Transform technique for still picture data compression standard», *International Organisation for Standardization (ISO)*, février 1991.

[Jain] A.K. Jain, «Image data compression : A review», *Proc. IEEE*, vol. 69, mars 1981, pp. 349–389.

[Lee] H. Lee, Y. Kim, A.H. Rowberg and E.A. Riskin, «Statistical Distributions of DCT coefficients and their application to an interframe compression algorithm for 3-D medical images», *IEEE Trans. Medical Imaging*, vol. 12, n°3, septembre 1993, pp. 478–485.

[Mathieu] P. Mathieu, M. Barlaud, M. Antonini, «Compression d'image par la transformée en ondelette et quantification vectorielle», *Traitement du Signal*, vol. 7, n°2, 1990, pp. 101–115.

[Murakami] H. Murakami, Y. Hatori, H. Yamamoto, «Comparison between DPCM and Hadamard transform coding in the composite coding of the NTSC color TV signal», *IEEE Trans. Commun.*, vol. COM-30, n°6, mars 1982, pp. 469–479.

[Neagoc] V.E. Neagoc, O. Stanasila, *Teoria recunoasterii formelor*, ed. Academiei Române, Bucharest, 1992.

[Redner] R.A. Redner, H.F. Walker, «Mixture densities, Maximum likelihood and the EM algorithm», *SIAM Review* 26, 1984, pp. 195–240.

[Reininger] R.C. Reininger and J.C. Gibson, «Distribution of two-dimensional DCT coefficients for images», *IEEE Trans. Commun.*, vol. COM-31, n°6, juin 1983, pp. 935–938.

[Saito] T. Saito et al., «Adaptive Discrete Cosine Transform image coding using gain/shape vector quantizers», *Proc. IEEE ICASSP'86*, Tokyo, Japon, avril 1986, pp. 129–132.

[Stephens] M.A. Stephens, «EDF statistic for goodness of fit tests and some comparisons», *Journal of the American Statistical Association*, vol. 69, n°347, septembre 1974.

[Tescher] A.G. Tescher, «Transform image coding», *Advance in Electronics and Electron Physics*, Supl. 12, New York Academic, 1979, pp. 113–115.

Manuscrit reçu le 11 juillet 1995

LES AUTEURS

Richard GRISEL



Richard Grisel a soutenu en 1987 à l'Université de Rouen une thèse de troisième cycle et en 1994 une habilitation à diriger des recherches à l'Université de Saint-Etienne. De septembre 1991 à août 1997 il a dirigé le groupe de recherche sur les architectures et systèmes électroniques et microélectroniques au sein du laboratoire LISA (CNRS EP0092) de l'École Supérieure de Chimie Physique Électronique de Lyon (CPE Lyon). Depuis septembre 1997 il est professeur à l'Université de Picardie Jules Verne.

Florence ASTRADE



Florence Astrade est titulaire d'un DEA Signal Image Parole (INPG Grenoble) soutenu en 1994. Elle prépare actuellement un doctorat traitant de la détection de signaux non stationnaires sur les images temps-fréquence par reconnaissance de formes au sein du laboratoire LISA de l'École Supérieure de Chimie Physique Electronique de Lyon (CPE Lyon).