



UN ALGORITHME DE COMPRESSION A REDONDANCE STATISTIQUE ET PERCEPTUELLE MINIMALE

Thierry EUDE, Hocine CHERIFI*

LIESIB Université de Bourgogne 6 rue Boulevard Gabriel BP138 21004 DIJON CEDEX

*TSI URA 842 23 Rue Paul Michelon 42023 St-Etienne CEDEX 02 e-mail:cherifi@univ-st-etienne.fr

RESUME

Afin d'accroître les performances des algorithmes de codage d'images traditionnels, il est essentiel d'aborder le problème sous l'angle statistique mais en tenant compte de l'aspect perceptuel du système visuel humain. Dans ce papier on présente un schéma de codage basé sur la DCT qui exploite ces deux aspects. En combinant un seuillage psychovisuel et une quantification statistique, on améliore le taux de compression sans dégradation de la qualité de l'image.

1. INTRODUCTION

Actuellement un certain nombre de standards de compression d'images fixes et animées sont finalisés (JPEG, MPEG). Ces standards définissent une sémantique et une structure de codage. Dans le cadre de ces standards il est possible de développer de nouveaux algorithmes en vue d'améliorer la qualité du processus.

D'une façon générale, les techniques de compression non conservatives s'intéressent à la redondance de l'information contenue dans l'image. Celle-ci peut être abordée en considérant qu'elle est de nature statistique ou perceptuelle.

Si l'on considère la redondance sous l'angle statistique, la quantification est l'élément prépondérant qui permet d'améliorer les performances du processus de compression. De nombreux auteurs ont proposé des schémas de quantification utilisant différents critères d'optimisation [1][2]. Dans [3], nous avons montré qu'une modélisation appropriée de la distribution des coefficients TCD permet d'améliorer de façon sensible les performances de la chaîne de compression.

ABSTRACT

To improve the performance of traditional coding schemes one should not only remove the statistical aspect of redundancy but also incorporate perceptual models in order to take care of subjective redundancy. In this paper we present a DCT coding scheme combining a psychovisual thresholding and a statistical quantisation. It appears to improve the compression rate without loss in subjective image quality.

Par ailleurs, l'approche perceptuelle privilégie les mécanismes du système visuel humain. En effet dans bon nombre d'applications le juge en dernier ressort de la qualité de reconstruction d'une image est le récepteur humain. Les algorithmes de codage utilisent alors des mesures perceptuellement significatives déduites de la modélisation du système visuel. Le but est alors de retenir dans une image toute l'information discernable par l'observateur. Le concept clef de l'approche perceptuelle est la notion de masquage qui traduit l'incapacité du système visuel à distinguer deux signaux. Ceci permet d'introduire la notion de distorsion perceptuelle comme critère d'optimisation. Ces dernières années un certain nombre de modèles de masquage perceptuels a été introduit dans les systèmes de compression [4][5]. Bien que s'inspirant d'hypothèses différentes, l'approche perceptuelle et l'approche statistique ont permis d'apporter des améliorations notables. Il semble donc évident que la prise en compte de ces deux aspects complémentaires est la seule voie qui permette de conserver l'information "pertinente" contenue dans l'image, tout en augmentant le taux de compression.



Dans ce papier, une nouvelle approche est proposée pour développer un algorithme de type JPEG qui prend en compte de façon originale la réduction d'information tant statistique que perceptuelle. Pour ce faire on combine un pré traitement basé sur un modèle perceptuel à un schéma de quantification optimal au sens de la distribution des coefficients TCD.

Après avoir présenté succinctement le codeur JPEG, nous décrivons les étapes de quantification et de pré traitement perceptuel. Le processus de codage intégrant ces deux éléments est ensuite évalué expérimentalement.

2. PRINCIPE DE L'ENCODAGE

Le processus d'encodage utilisé correspond au processus séquentiel de base défini dans le standard JPEG. Il comprend trois étapes essentielles: passage de l'image dans le domaine transformé, Quantification et codage entropique. La transformation de l'image est effectuée par la TCD opérant sur des blocs 8x8. Le rôle de cette opération est de concentrer l'énergie afin de réduire le nombre d'échantillons significatifs à transmettre. L'étape de quantification est réalisée en utilisant une table de valeurs correspondant aux différents coefficients. Ces coefficients quantifiés sont ensuite traités par un processus d'encodage entropique. La figure 1 illustre la procédure principale pour tous les processus d'encodage basés sur la TCD.

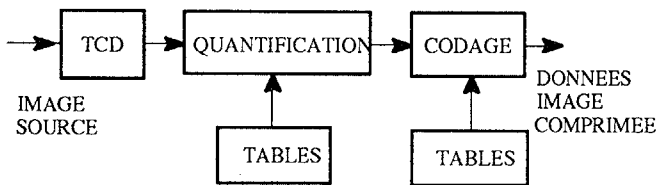


Figure 1 : Diagramme simplifié du codeur basé sur la TCD

Aucune valeur par défaut n'est spécifiée par le standard pour les tables de quantification. Ces valeurs sont choisies en fonction de la qualité de restitution et selon les caractéristiques particulières des images, des systèmes et des conditions de visualisation de l'application. La norme indique néanmoins des valeurs initiales de tables de quantification usuellement dénommées matrices de visualisation.

3. QUANTIFICATION STATISTIQUE

L'étape primordiale dans l'optimisation du processus de compression réside dans la définition de la table de

quantification. Pour ce faire on exploite la distribution statistique des données. Des travaux antérieurs ont permis de montrer que la distribution la plus adaptée pour modéliser les coefficients résultant de la TCD d'image est une loi mélange de gaussienne [3] :

$$\Phi(\sigma_i, p_i, x) = \sum_{i=1}^k \frac{p_i}{\sqrt{2\pi} \sigma_i} \exp\left(-\frac{x^2}{2\sigma_i^2}\right) \quad (3.1)$$

où p_i, σ_i sont respectivement la probabilité a priori, et l'écart type de la i ème composante du mélange. Ce résultat permet, en utilisant un algorithme d'identification des paramètres de la loi, de définir une table de quantification optimale au sens de la distribution statistique des coefficients [6].

La quantification des coefficients est alors obtenue de la manière suivante :

$$I_{pq}(u, v) = \text{int} \left[\frac{I_p(u, v)}{Fe \times q_{u,v}} \right] \quad (3.2)$$

$u, v = 0, 1, 2, \dots, 7$

Fe est un coefficient qui permet d'ajuster le taux de compression

$q_{u,v}$ est le pas de quantification défini de la façon suivante :

$$q_{u,v}^{-1} \text{ est tel que } \int_{-\infty}^{q_{u,v}^{-1}} \Phi(x, \hat{p}_i, \hat{\sigma}_i) dx = c \quad (3.3)$$

Cette procédure de quantification est schématisée par le diagramme de la figure suivante.

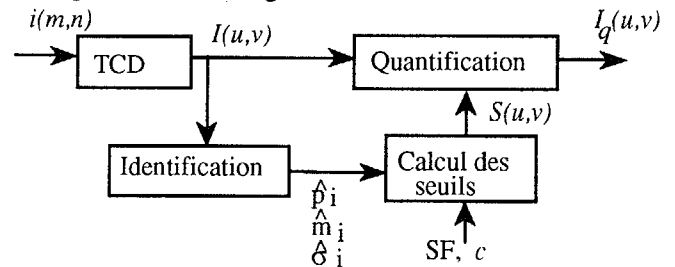


Figure 2 : Principe de la quantification probabiliste

4. MODELE PERCEPTUEL

Il est possible d'exploiter les imperfections du système visuel humain dans les processus de codage d'image. En effet si l'oeil humain est incapable de discerner la distorsion introduite par le processus de compression décompression, la perte introduite est alors nulle au sens perceptuel. Des modèles de masquage visuel ont été développés [7][8]. A partir de ces modèles, il est possible de déterminer des seuils de masquage dans le domaine originel ou transformé de l'image. L'étape de pré traitement que nous employons utilise le concept de la distorsion visuellement discernable (DVD)

récemment introduit par Jayant [7]. La DVD d'une image est une fonction des propriétés locales du signal, telle que la luminance du fond ou l'activité des changements de luminance.

$$DVD(x, y) = \max\{f_1[m_{bg}(x, y), \sup(m_{el}(x, y, v_{x,y}))], f_2[m_{bg}(x, y)]\} \quad (4.1)$$

où $m_{bg}(x, y)$ représente la luminance moyenne du fond
 $m_{el}(x, y)$ représente la moyenne pondérée des différences de luminance dans le voisinage du pixel x, y .

$f_1(x, y)$ modélise l'effet de masquage spatial

$f_2(x, y)$ modélise le seuil de visibilité dû à la luminance du fond

La signification perceptuelle de chaque pixel d'une image est évaluée par la DVD. Le seuil de masquage pour un bloc TCD est donné par:

$$dvd(u, v) = TCD\{DVD(x, y)\} \quad (4.2)$$

L'estimation de cette quantité permet de déterminer pour chacun des blocs TCD de l'image à comprimer des seuils de signification perceptuels des coefficients TCD.

5. PROCESSUS D'ENCODAGE

L'introduction du modèle perceptuel peut être utilisé pour déterminer une table de quantification optimale pour chacun des blocs de l'image considéré. Bien entendu une telle méthodologie est inadaptée au standard JPEG. La stratégie utilisée consiste à effectuer une étape de pré quantification de l'image. Tous les coefficients TCD dont la valeur est inférieure au seuil de quantification sont mis à zéro alors que les autres coefficients restent inchangés. Ainsi l'on conserve la qualité perceptuelle de l'image tout en réduisant l'information transmise. Le principe du processus de compression proposé est globalement schématisé dans la figure 3

Ce schéma de compression bien que relativement complexe reste compatible au standard JPEG. Il correspond en fait à une pré quantification des coefficients TCD suivi d'une quantification standard. Les éléments de la matrice de quantification se déduisant de l'identification des paramètres de la distribution des coefficients pré quantifiés.

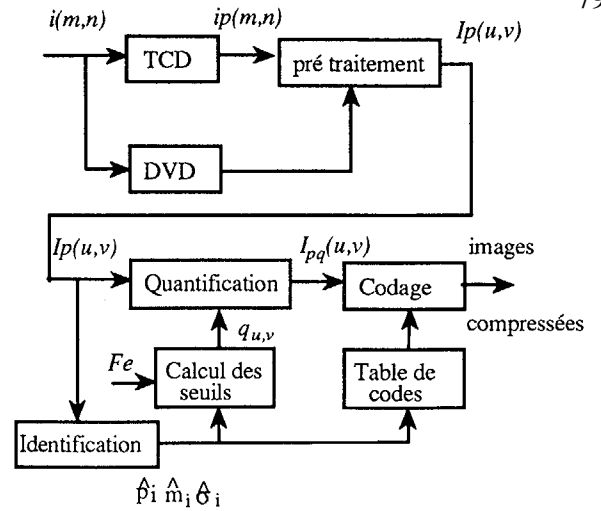


Figure 3 : Schéma de compression avec pré traitement perceptuel et matrice de quantification probabiliste

6. RESULTATS EXPERIMENTAUX

Afin d'évaluer les performances du processus proposé nous avons simulé la compression/décompression sur une banque d'images médicales et naturelles.

Nous avons comparé les performances obtenues à l'aide de trois processus de compression différents :

- JPEG avec matrice de visualisation
- JPEG avec matrice de quantification probabiliste
- JPEG avec pré traitement perceptuel et matrice de quantification probabiliste

D'une façon générale, les résultats de ces simulation ont fait apparaître un accroissement des performance du processus de compression. Ainsi à qualité visuelle comparable, le taux de compression est toujours supérieur. Ce comportement est illustré par l'exemple représentatif donné par le tableau 1. Celui-ci représente les résultats obtenus pour une image donnée figure 4.

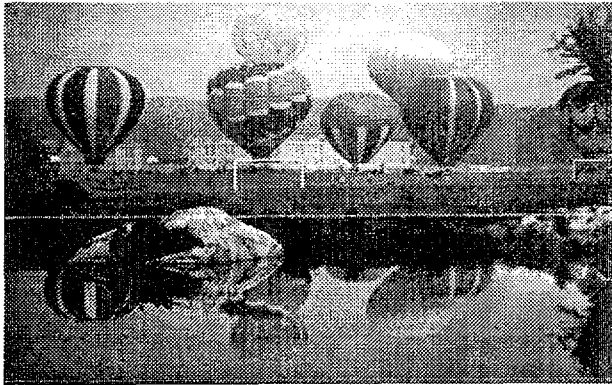
Processus utilisé	Taux de compression bit/pixel	RSB (dB)	RSBP (dB)
a	0,45	29,24	34,24
b	0,39	29,36	35,15
c	0,34	29,69	36,42

Tableau 1: a) JPEG avec matrice de visualisation, b) JPEG avec matrice de quantification probabiliste, c) JPEG avec pré traitement perceptuel et matrice de quantification probabiliste

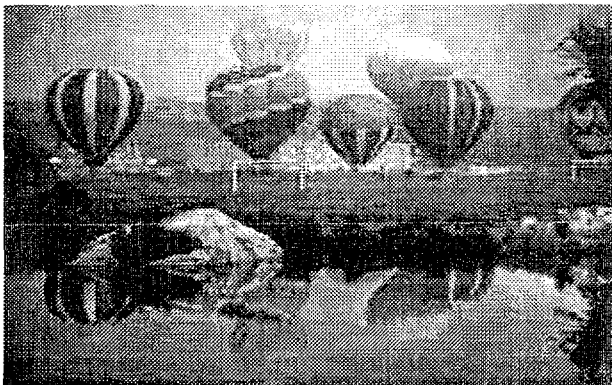


On donne par ailleurs les valeurs du RSB perceptuel (RSBP). Cette quantité est une mesure de la capacité de cette méthode à exploiter les caractéristiques perceptuelles de l'image. Elle est définie par :

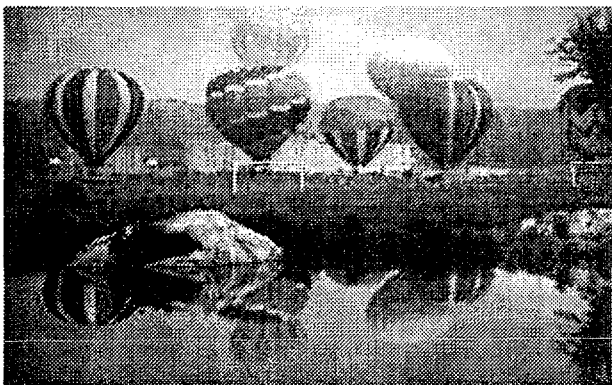
$$RSBP = 20 \log_{10} \left(\frac{255}{\sqrt{E \left\{ \left[|I(x,y) - \hat{I}(x,y)| - DVD(x,y) \right]^2 \delta(x,y) \right\}}} \right) \quad (6.1)$$



(a) image originale



(b) JPEG avec matrice de visualisation



(c) avec pré traitement perceptuel et matrice de quantification probabiliste

Figure 4

7. CONCLUSION

Le schéma de codage d'image proposé utilise une mesure de la sensibilité au contraste et de l'effet de masquage spatial du système visuel humain.

Notre stratégie consiste à effectuer un pré traitement de l'image originelle à compresser afin de mettre en valeur ses éléments perceptuellement significatifs. Le but de cette étape est de s'affranchir de la redondance perceptuelle. A l'issue de ce traitement, on utilise les informations statistiques sur la distribution des coefficients TCD pour réduire la redondance au niveau statistique. La combinaison de critères d'optimisation statistique et perceptuel permet d'exploiter de façon plus efficace les processus de compression standards. Bien entendu le gain en performance est aux prix d'une complexité accrue de l'encodeur.

BIBLIOGRAPHIE

- [1] E. Viscito and C. Gonzales, "A video compression algorithm with adaptative bit allocation and quantization", *Proc. SPIE Visual Communications and Image Processing*, pp. 58-71, Vol. 1605, 1991.
- [2] S.W. Wu and A. Gersho, "Enhanced video compression with standardized bit stream syntax", *Proc. of IEEE ICASSP*, pp. I.103-I.106, april 1993.
- [3] T. Eude, R. Grisel, H. Cherifi and R. Debrie, "On the distribution of the DCT coefficients", *Proc. of IEEE ICASSP*, Adelaïde (Australia), pp. V.365-V.368, 19-22 april 1994.
- [4] A. Watson, "DCT quantization matrices visually optimized for individual images", *SPIE Conference on human Vision, Visual Processing and Digital Display IV* ., Vol 1913 pp. 202-216, February 1993.
- [5] D. McLaren and D. Thong Nguyen, "Removal of subjective redundancy from DCT_coded images", *IEE proceedings-I*, Vol 138, N°5 ., pp. 345-350, October 1991.
- [6] T. Eude, H. Cherifi "Une méthode de compression adaptée JPEG basée sur la modélisation statistique", *La revue de modulad INRIA* ., Vol2 N°14 pp. 23-49, Décembre 1994.
- [7] N. Jayant, "Signal compression: technologie targets and research directions", *IEEE J. Select. Areas Commun.*, pp. 314-323, Jun. 1992.
- [8] N. Jayant, J. Johnston and R. Safranek "Signal compression: based on human perception", *Proceeding of the IEEE J. Vol 81 N°10*., pp. 1385-1422, October 1993.