



COMPRESSION D'IMAGES RADIOLOGIQUES PAR TRANSFORMATION
DISCRETE EN COSINUS

F. J. García Ugalde et I. Magaña M.

División de Estudios de Posgrado
Facultad de Ingeniería, UNAM
A. P. 70-256; 04510 México, D. F.

RESUME

Pour la mémorisation et la transmission d'images numériques, en particulier dans le cas d'images radiologiques, il est nécessaire d'utiliser la mémoire et la largeur de bande du canal disponibles de la manière la plus efficace. Pour diminuer la quantité de mémoire des équipements de calcul de centres hospitaliers et pour réduire aussi la largeur de bande requise pour transmettre ces images entre les différents services, il est nécessaire de les coder avant leur mémorisation ou leur transmission. Une des techniques dites de la première génération [1], ayant permis d'obtenir de bons résultats pour le codage d'images radiologiques, est basée sur l'utilisation de la transformation discrète en cosinus bidimensionnelle. Dans cet article, nous présentons les résultats de simulations d'un algorithme de codage pour de telles images utilisant cette transformation. Une table d'assignation de bits calculée en fonction d'une expression empirique [2], est utilisée pour effectuer la quantification des coefficients transformés. L'image numérique originale représentée par une matrice de $n \times n$ points est divisée en sous-blocs disjoints de dimensions $m \times m$. Dans notre cas, les simulations effectuées avec $n=256$ et, $m=16$, $m=32$ et $m=64$ montrent une forte amélioration de la qualité des images traitées au fur et à mesure de l'augmentation du paramètre m . Ainsi sur plusieurs images, le taux maximum de compression obtenu avec une qualité acceptable est de 12 pour $m=64$, 6 pour $m=32$ et 4 pour $m=16$. En opérant un traitement supplémentaire après le décodage, une compression de l'ordre de 16 pour $m=64$ pourra être obtenue prochainement.

SUMMARY

For digital images storing and transmitting, especially in the case of radiological images, it is necessary to use in the better way the storage memory and the bandwidth of the communication channel. In order to reduce memory requirements for medical-purpose computer equipments, and also to reduce bandwidth requirements for the transmission of these images between hospital services, it is necessary to code the digital images before storing and transmitting them. One of the so-called first generation techniques for coding [1], which has enabled to achieve good performances in coding radiological images, is based on the two-dimensional discrete orthogonal transform utilisation. In this paper, we present results of simulations for coding such images using the two-dimensional Discrete Cosine Transform (2D-DCT). An empirical expression based bit-allocation table [2], has been used to code the quantized transformed image. The original digital image represented by an $n \times n$ matrix, is divided into $m \times m$ subimages. In this work, $n=256$ and simulations have been done with $m=16$, $m=32$ and $m=64$. It is shown that the quality reconstruction improves as m increases. Among several images, the maximum compression ratio obtained, keeping a good image quality, was 12:1 with $m=64$, 6:1 with $m=32$, and 4:1 with $m=16$. With post-processing after decoding, a compression ratio of 16:1 with $m=64$ is going to be achieved soon.

1. INTRODUCTION

Le développement accéléré des nouvelles technologies dans le domaine de l'électronique et des communications a provoqué un changement progressif des méthodes d'acquisition, de mémorisation, de visualisation, et de transmission d'images radiologiques, utilisés dans le diagnostic. Parallèlement aux avantages inhérents aux images radiologiques numérisées, tel que le fait de pouvoir utiliser des techniques de traitement numérique pour assister le diagnostic. Il existe le besoin de compter avec des méthodes efficaces pour réduire le grand nombre de bits nécessaires pour les représenter,

ceci se traduira par une meilleure utilisation de la mémoire et de la largeur de bande du canal de transmission.

Dans ce travail, nous avons effectué le codage ou compression d'images radiologiques, en utilisant une des méthodes dites de la première génération [1], le codage par transformation, qui est basé sur le fait que la distribution de l'énergie dans la version transformée d'une image, est plus adéquate pour le codage que sa représentation spatiale.

Parmi les différentes méthodes de codage par transformation, nous avons étudié les performances de la transformation discrète en cosinus bidimensionnelle, elle est particulièrement



intéressante pour la compression d'images qui ne comportent pas beaucoup de fréquences hautes [1], tel est le cas de la plupart d'images de type radiologique [3]. En effet, lorsqu'une image peut être modélisée par un processus de Markov d'ordre inférieur, avec une grande corrélation entre ses points, la transformation discrète en cosinus a la propriété de se rapprocher de la transformation optimale Karhunen-Loève [4] (surtout quand les dimensions du bloc à coder sont grandes), en conséquence le taux de compression obtenu est assez élevé, tout en conservant un nombre de calculs nettement inférieur.

Une table d'assignation de bits calculée en fonction d'une expression empirique [2], qui permet de coder chaque coefficient transformé avec un nombre de bits proportionnel à son contenu d'information, est utilisée pour effectuer la quantification. Aussi grâce à cette table, la structure bidimensionnelle de l'image peut se transformer, après l'opération de quantification, dans une suite unidimensionnelle de coefficients transformés. L'envoi de cette table au récepteur permet de reconstruire la structure bidimensionnelle pour effectuer le décodage.

Les travaux publiés dans [3] montrent les résultats obtenus avec des taux de compression de 4, 8, 12 et 16. Les images originales utilisées dans ces travaux sont représentées par des matrices de 512x512 points images et chaque point est quantifié à 256 niveaux. Le traitement est réalisé en prenant l'image complète, ce qui le rend plus lourd.

En comparaison, nous avons obtenu une qualité des images traitées tout à fait équivalente, mais avec une structure de traitement plus adaptée à l'implantation des algorithmes sur une carte de traitement numérique de mémoire limitée. Cette implantation fait partie de la phase actuelle du travail. La structure de traitement que nous avons utilisée consiste à diviser l'image originale représentée par une matrice de $n \times n$ points images, où $n=256$, en sous-blocs disjoints de dimensions $m \times m$. Pour des valeurs de m égaux à 16, 32, et 64, le taux de compression obtenu sur plusieurs images est respectivement de 4, 6 et 12. Comme un paramètre de comparaison avec la plupart des travaux publiés nous avons utilisé l'erreur NMSE (normalized mean-square error). Sous forme de photographies, on présente aussi quelques résultats.

Le présent article se décompose en trois parties: la première partie donne quelques notions de base sur la compression d'images par transformation discrète en cosinus, la deuxième partie décrit le système de codage utilisé, et la dernière partie donne les résultats d'essais subjectifs effectués en vue d'apprécier les dégradations introduites par le codage.

2. COMPRESSION D'IMAGES PAR TRANSFORMATION DISCRETE EN COSINUS (DCT)

Les méthodes utilisées pour effectuer la compression, ou le codage d'image, peuvent être des techniques orientées soit à la réduction de redondance, soit à la réduction d'entropie. Il existe bien sûr la possibilité d'effectuer une combinaison des différentes méthodes.

Puisque l'entropie est définie comme l'information moyenne d'une source, une réduction d'entropie entraîne une réduction d'information. Il n'est pas possible de récupérer cette réduction d'information, donc en conséquence, c'est une opération irréversible. C'est dans cette catégorie que nous trouvons les méthodes de codage par transformation.

En contrepartie, les méthodes de réduction de redondance éliminent, ou au moins réduisent, la redondance des images, de telle façon qu'elle pourrait être réincorporée plus tard. Ils sont donc en conséquence des procédés réversibles.

Malgré la perte d'information introduite par les méthodes de réduction d'entropie, elles sont plus intéressantes puisqu'elles permettent d'atteindre des taux de compression beaucoup plus élevés.

Le codage par transformation se réalise en effectuant séquentiellement deux sortes de traitement: la première opération consiste à prendre dans l'image

originale un ensemble de points et à leur appliquer une transformation linéaire de telle sorte que l'ensemble initial, où les points étaient statistiquement dépendants soit transformé en un ensemble de coefficients plus indépendants; la seconde opération est le codage proprement dit qui est ainsi opérée, non pas dans l'espace image, mais dans l'espace transformé en quantifiant les coefficients ainsi calculés: c'est dans cette opération que l'on peut faire intervenir des propriétés psychovisuelles pour déterminer la loi de quantification et le nombre d'éléments binaires à utiliser pour chaque coefficient de manière à minimiser la visibilité subjective globale de l'erreur associée.

Les transformations d'images ont pour but d'obtenir dans le plan transformé des caractéristiques plus favorables pour la réduction de débit. L'observation de la répartition d'énergie, c'est à dire des amplitudes des différents coefficients obtenus après transformation, permet de préciser ces caractéristiques. Les coefficients de forte amplitude sont beaucoup moins fréquents que ceux de faible amplitude et cette distribution est pratiquement stationnaire d'une image à l'autre. Les coefficients de forte amplitude doivent être quantifiés avec plus de précision, tandis que ceux qui apportent peu d'information sont négligés, on obtient de cette manière la compression.

Pour effectuer la transformation d'une image numérique représentée par une matrice de $n \times n$ points images, généralement on utilise des sous-blocs de dimensions $m \times m$, après on obtient la transformation bidimensionnelle pour chacun de sous-blocs. Cette opération de division est faite principalement pour réduire le nombre de calculs. Généralement le paramètre m est une puissance de deux, et les sous-blocs sont disjoints.

Dans ce travail les coefficients de la transformation bidimensionnelle d'un sous-bloc j sont représentés par $F_j(u,v)$, où $0 \leq u \leq m-1$, $0 \leq v \leq m-1$, et $1 \leq j \leq (n/m)^2$.

Les composantes de fréquence qui sont contenues dans un sous-bloc de l'image originale sont représentées par les coefficients transformés, ceci puisque la transformation est en réalité une décomposition du bloc en fréquences généralisées. Les coefficients $F_j(u,v)$, avec de petites valeurs de u et de v , représentent les composantes de basses fréquences dans l'image, tandis que les effets de haute résolution spatiale sont représentés par les coefficients $F_j(u,v)$, avec de grandes valeurs de u et de v .

Généralement la plupart de l'énergie de l'image est contenue dans les coefficients $F_j(u,v)$ qui caractérisent les composantes de basses fréquences [5].

Le codage par transformation discrète en cosinus est défini de la façon suivante [6].

Soit

$$f(k) \text{ avec } k=0,1,\dots,m-1 \tag{1}$$

une séquence de dimension m .

La transformation, $F(u)$, est donnée par:

$$F(u) = 2/m C(u) \sum_{k=0}^{m-1} f(k) \cos \left[\frac{(2k+1) u \pi}{2m} \right]; \tag{2}$$

$$u = 0,1,\dots,m-1$$

$$\text{où } C(u) = \begin{cases} 1/\sqrt{2} & \text{pour } u=0 \\ 1 & \text{pour } u=1,2,\dots,m-1 \end{cases} \tag{3}$$

et la transformation inverse est définie par:

$$f(k) = \sum_{u=0}^{m-1} C(u) F(u) \cos \left[\frac{(2k+1) u \pi}{2m} \right]; \tag{4}$$

$$k = 0,1,\dots,m-1$$

Pour le cas bidimensionnel, la transformation cosinus est définie par:

$$F(u,v) = 4/m^2 C(u)C(v) \sum_{k=0}^{m-1} \sum_{l=0}^{m-1} f(k,l) \cos \left[\frac{(2k+1) u \pi}{2m} \right] \cos \left[\frac{(2l+1) v \pi}{2m} \right]; \quad u,v = 0,1,\dots,m-1 \quad (5)$$

où les valeurs de $C(u)$ et $C(v)$ sont définies par (3). La transformation inverse de (5), est définie par:

$$f(k,l) = \sum_{u=0}^{m-1} \sum_{v=0}^{m-1} C(u) C(v) F(u,v) \cos \left[\frac{(2k+1) u \pi}{2m} \right] \cos \left[\frac{(2l+1) v \pi}{2m} \right]; \quad k,l = 0,1,\dots,m-1 \quad (6)$$

Dans les dernières années on a donné beaucoup d'importance à cette transformation et on compte à l'heure actuelle avec plusieurs algorithmes rapides [7]. En plus de sa caractéristique de compactation de l'énergie, elle introduit moins de faux contours par rapport à la transformation de Fourier, principalement lorsque le taux de compression est assez élevé.

Les coefficients de la transformation cosinus ont tendance à suivre une distribution Gaussienne, à l'exception du coefficient $F_j(u,v)$, lequel est modélisé d'une meilleure manière par une fonction de densité de Rayleigh.

En considérant une variable aléatoire Gaussienne avec une variance égale à σ^2 , lorsqu'elle est quantifiée avec une distorsion maximale D , elle a besoin d'au moins $1/2 \log(\sigma^2/D)$ bits [8]. On peut observer que si le numéro de bits assigné à un des coefficients $F_j(u,v)$, est proportionnel à son apport d'information, ce numéro devrait être proportionnel au logarithme de la variance du coefficient.

3. SYSTEME DE CODAGE UTILISE

La compression ou le codage d'image est réalisé dans les quatre phases fondamentales suivantes:

- . calcul de la transformation discrète en cosinus
- . normalisation des coefficients transformés
- . calcul de la table d'assignation de bits
- . quantification des coefficients transformés

Dans la première phase l'image originale, représentée par une matrice de $m \times m$ points images, est divisée en sous-blocs de $m \times m$ points, ensuite en utilisant les expressions (5) et (3), on calcule la transformation cosinus bidimensionnelle de chacun des sous-blocs.

La deuxième phase consiste à normaliser les coefficients $F_j(u,v)$ par rapport à leur écart type, de façon à ce qu'il en résulte une variance unité. La variance $\sigma_{u,v}^2$ de chaque coefficient est estimée, moyennant les carrés des éléments qui représentent une même fréquence, pour l'ensemble des $(n/m)^2$ sous-blocs (il faut noter que $E\{F(u,v)\}=0$, pour des images à moyenne nulle) [5]:

$$\sigma_{u,v}^2 = E \left\{ |F(u,v)|^2 \right\} = \frac{1}{m} \sum_{j=0}^{m-1} |F(u,v)|^2; \quad u,v = 0,1,\dots,m-1 \quad (7)$$

Les coefficients normalisés $F_{Nj}(u,v)$ sont:

$$F_{Nj}(u,v) = \frac{F_j(u,v)}{\sigma_{u,v}}; \quad \text{pour } u,v = 0,1,2,\dots,m-1 \quad \text{et } j = 1,2,\dots,(n/m)^2 \quad (8)$$

La phase suivante consiste à calculer la table d'assignation de bits. Cette table de $m \times m$ éléments, nous donne le nombre de bits avec lesquels on va quantifier chacun des coefficients normalisés $F_{Nj}(u,v)$. Elle a été obtenue par une expression qui minimise l'erreur de quantification moyenne total dans

les blocs, lorsqu'on utilise des quantificateurs uniformes:

$$B(u,v) = b_{\text{moy}} + \frac{2}{\ln(10)} \ln(\sigma_{u,v}^2) - \frac{1}{m^2} \sum_{u=0}^{m-1} \sum_{v=0}^{m-1} \ln(\sigma_{u,v}^2) \quad (9)$$

pour $u,v = 0,1,\dots,m-1$

où b_{moy} est le nombre de bits moyen espéré, pour représenter les coefficients d'un sous-bloc et $B(u,v)$ correspond au nombre de bits avec lesquels on va représenter le coefficient respectif, $F_j(u,v)$, où $j = 1, \dots, (n/m)^2$. Puisque $B(u,v)$ est un nombre réel, sa valeur est arrondie à la valeur de l'entier le plus proche, dans le cas où il est négatif il est mis égal à zéro. En faisant varier b_{moy} on obtient les différents taux de compression.

Dans la dernière phase du codage, on effectue la quantification des coefficients transformés normalisés, $F_{Nj}(u,v)$.

Ces coefficients sont encore normalisés par rapport à leurs valeurs maximums $m(u,v)$, de façon à ce qu'ils prennent uniquement des valeurs entre -1 et +1. On a donc:

$$m(u,v) = \max \left\{ |F(u,v)| \right\}; \quad \text{avec } u,v = 0,1,\dots,m-1 \quad (10)$$

et $j = 1,2,\dots,(n/m)^2$

Ensuite, on obtient les coefficients normalisés $F'_{Nj}(u,v)$, définies par:

$$F'_{Nj}(u,v) = \frac{F_{Nj}(u,v)}{m(u,v)}; \quad u,v = 0,1,\dots,m-1 \quad \text{et } j = 1,2,\dots,(n/m)^2 \quad (11)$$

Finalement, chaque coefficient $F'_{Nj}(u,v)$, est représenté par un nombre entier $S_j(u,v)$, donné par:

$$S_j(u,v) = \left[F'_{Nj}(u,v) (2^{B(u,v)} - 1) \right] \quad (12)$$

où $S_j(u,v)$ est arrondie à la valeur de l'entier le plus proche.

La représentation binaire en complément à deux de $S_j(u,v)$, est la valeur codée du coefficient $F_j(u,v)$.

Pour la reconstruction de l'image, on doit transmettre au décodeur: les valeurs $S_j(u,v)$, les valeurs maximums $m(u,v)$, les variances $\sigma_{u,v}^2$, et la table d'assignation des bits $B(u,v)$.

Avec cette information l'opération de décompression se réalise en calculant d'abord la valeur reconstruite des coefficients normalisés $F'_{Nj}(u,v)$, par :

$$F'_{Nj}(u,v) = \frac{S_j(u,v)}{2^{(B(u,v)-1)} - 1} \quad (13)$$

Ensuite on recupere $F_j(u,v)$, par:

$$F_j(u,v) = \sigma_{u,v} m(u,v) F'_{Nj}(u,v); \quad u,v = 0,1,\dots,m-1 \quad \text{et } j = 1,2,\dots,(n/m)^2 \quad (14)$$

Avec $F_j(u,v)$ et les équations (6) et (3), on calcule la transformation inverse de chacun des $(n/m)^2$ sous-blocs. Nous obtenons ainsi l'image reconstruite.

4. RESULTATS OBTENUS

Pour étudier la compression d'images, nous avons utilisé principalement quatre types d'images radiologiques qui représentent: des angiographies (A1, A2), une cellule (CE), des tissus (T1,T2), et une radiographie des vertèbres (V). Ces images ont été fournis par l'Université Technologique de Compiègne, sous la forme des matrices de 256x256 points images et chaque point est quantifié à 256 niveaux.

A fin de donner à l'algorithme de codage une structure



adéquate pour sa future implantation sur une carte prévue autour d'un microprocesseur arithmétique de la famille TMS de Texas Instruments. Nous avons traité les images en les divisant en sous-blocs disjoints de dimensions $m \times m$. Plusieurs simulations ont été effectuées, pour m égal à 16, 32 et 64. Nous nous sommes limités à ces trois valeurs pour conserver un temps de calcul faible, et surtout pour limiter la quantité de mémoire qui sera nécessaire lors de l'implantation de l'algorithme. Actuellement l'ensemble des programmes peut être exécuté sur un micro-ordinateur de type PC-AT. Le temps de traitement, qui comprend le codage et le décodage, est respectivement pour les différentes valeurs de m de: 8, 10 et 12 minutes. Le micro-ordinateur est équipé d'une carte d'interfaçage qui permet la liaison avec un écran de visualisation, dont la résolution est du type d'images de télévision.

Les tests effectués sur les images traitées pour évaluer leur qualité subjective ont consistés à visualiser les images et donner une note de qualité. Malheureusement les images ont été évaluées uniquement par des ingénieurs non-spécialisés en radiologie, nous devons donc compléter notre étude en faisant intervenir ces derniers.

Les résultats de simulations sont donnés sous forme de photographies. Par manque d'espace nous présentons seulement quelques photographies composées où on montre une partie de l'image originale et la reproduction de la même partie de l'image, traitée avec les différentes dimensions de sous-blocs sélectionnés. Le taux de compression pour chaque ensemble d'images est indiqué sous chaque photographie, il est maximum lorsque $m=64$. Comme un paramètre de comparaison avec la plupart des travaux publiés nous utilisons l'erreur NMSE (normalized mean-square error), définie dans [3]. Sous forme de courbes nous présentons les mesures de ces erreurs sur quelques images traitées.



Image CE, compression 8:1. Taux maximum obtenu 12.



Image T1, compression 8:1. Taux maximum obtenu 12.

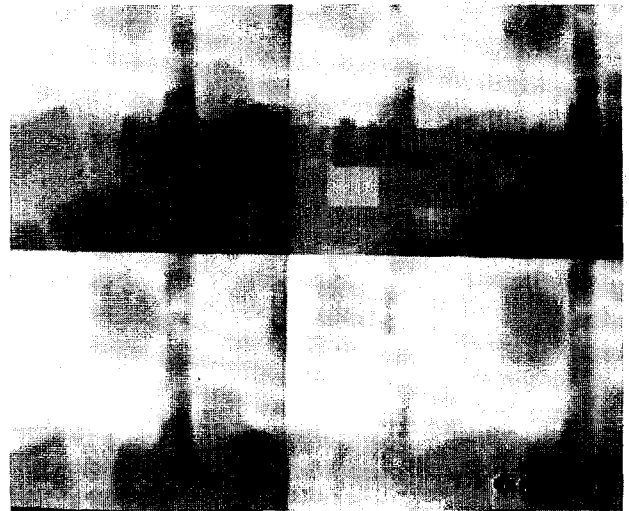
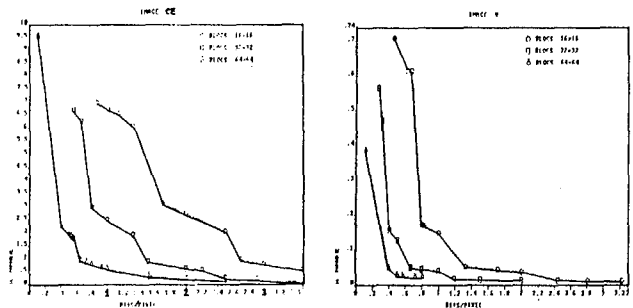


Image V, compression 10:1. Taux maximum obtenu 16.



Erreur NMSE. vs. bits/pixel pour les images CE et V.

5. CONCLUSIONS

Cet étude nous a permis de définir la structure d'un algorithme de codage d'images radiologiques et les dimensions des sous-blocs de l'image originale. Ceci permet d'obtenir un taux de compression supérieur à 10, tout en conservant un faible nombre de calculs, de façon à ce que l'implantation de l'algorithme sur une carte prévue autour d'un microprocesseur arithmétique soit tout à fait réalisable avec peu de composantes électroniques. Les résultats obtenus montrent que pour des sous-bloques de 64×64 la qualité subjective d'images traitées est acceptable et la quantité de mémoire qui sera nécessaire lors de l'implantation reste encore faible.

REFERENCES

- [1] Kunt M., Ikonomopoulos A., "Second-generation image-coding techniques", Proc. IEEE, vol. 73, pp. 543-574, april 1985.
- [2] Wintz P. A., Kurtenbach A. J., "Waveform error control in PCM telemetry", IEEE Trans. in Inf. Theory, vol. IT-14, pp. 650-651, 1968.
- [3] Shih-Chung Lo, Huang H. K., "Radiological image compression: Full-frame bit allocation technique" Radiology, vol. 155, pp. 811-817, june 1985.
- [4] Netravali A. N., Limb J. O., "Picture coding: A review", Proc. IEEE, vol. 68, pp. 366-406, march 1980.
- [5] Rosenfeld A., Kak A. C., "Digital picture processing", Academic Press, vol. 1, 2nd ed., 1982
- [6] Ahmed N., Natarajan T., Rao K. R., "Discrete cosine transform", IEEE Trans. Comput., vol C-23, pp. 90-93, Jan. 1974.
- [7] Chen W. H., Smith C. H., Fralick S. C., "A fast computational algorithm for the discrete cosine transform", IEEE Trans. on Commun., vol. COM-25, pp. 1004-1008, sept. 1977.
- [8] Berger T., "Rate distortion theory", Englewood Cliffs, Prentice-Hall, N. J., 1971.