

Optimisation conjointe d'un codeur source et d'une modulation codée pour la transmission numérique d'images sur canal bruité

O. Aitsab*, R. Pyndiah* et B. Solaiman**

TELECOM BRETAGNE

Technopôle Brest Iroise, B.P. 832, 29285 Brest Cedex, France.

* Département. S.C., ** Département. I.T.I.

Email : Omar.Aitsab@enst-bretagne.fr

(Tél : 98 00 10 70, Fax : 98 00 10 98)

RESUME

ABSTRACT

Dans cet article, nous proposons un système de transmission d'images fixes qui minimise la distorsion introduite par les erreurs résiduelles de transmission. Ce système comprend un codeur source combinant la transformée en cosinus discret et la technique de Quantification Vectorielle (QV). Le taux de compression est égal à 16.7 avec un PSNR de 29.2dB pour l'image Lenna. Les informations issues du codeur source sont ensuite codées par un code en bloc BCH(63,57,3) et transmis par une modulation MAQ-64 (Modulation d'Amplitude en Quadrature). Les paramètres du dictionnaire utilisé pour la QV ainsi que ceux du codeur canal ont été optimisés pour la MAQ-64. Pour le décodage canal, nous avons utilisé un décodage suivant le maximum de vraisemblance (décodage pondéré). Le PSNR de l'image Lenna reconstruite après transmission est égale à 28.6dB pour un SNR canal de 11dB. Ceci représente une faible dégradation du PSNR (0.6dB) en présence d'un Taux d'Erreur Symbole (TES) canal de 10% à l'entrée du décodeur de canal.

In this paper, we propose a complete transmission system for still images which minimizes the distortion introduced by residual transmission errors. This system comprises a source encoder combining DCT (Discrete Cosine Transform) and VQ (Vector Quantization). The compression rate is equal to 16.7 for an average PSNR (Peak Signal to Noise Ratio) of 29.2 dB for Lenna image. The data is then encoded by a BCH(63,57,3) block code and is mapped on a 64-QAM constellation. The parameters of the codebook and channel encoder have been optimized for 64-QAM. For channel decoding, we use maximum likelihood sequence decoding. The PSNR of the reconstructed Lenna image in the receiver is equal to 28.6 dB for a channel SNR of 11 dB. This represents a very small degradation of the PSNR (0.6 dB) in presence of a SER (Symbol Error Rate) of 10% at the input of the channel decoder.

1. Introduction

Durant ces dernières années, les techniques de codage de source et de canal ont enregistré d'énormes progrès. Actuellement, plusieurs systèmes utilisent, d'une part, les modulations codées [1-2] permettant d'augmenter l'efficacité spectrale et, d'autre part, un codeur source puissant pour accroître le nombre d'images transmises dans une bande de fréquence donnée. Cependant, les images transmises deviennent très sensibles aux erreurs de transmission si aucune précaution n'est prise pour atténuer leur effet.

Les exigences requises pour les systèmes de transmission numériques sont de plus en plus sévères. Les images reçues doivent avoir une bonne qualité subjective et rester exploitable même dans un contexte de taux d'erreur canal très élevé. Il n'est donc plus possible d'optimiser séparément les différentes fonctions du système.

Plusieurs auteurs [3-5], ont traité le problème du codage conjoint source canal. Ils ont montré que les performances du système global peuvent être améliorées en utilisant les interactions entre les différents codeur/décodeur du système.

Cependant, ces papiers ont traité des systèmes de transmission simples qui ne sont plus compatibles avec les systèmes de transmission actuels qui utilisent des modulations à grande efficacité spectrale et des codes correcteur d'erreurs à haut pouvoir de correction.

Dans ce papier, nous proposons un nouveau système complet pour la transmission d'images qui minimise la distorsion en présence des erreurs de transmission. Toutes les fonctions de ce système ont été optimisées conjointement afin d'augmenter la tolérance aux erreurs.

Dans la deuxième partie de cet article, nous présenterons un nouveau codeur source qui combine la transformée en cosinus discret (DCT) et la Quantification Vectorielle (QV). La troisième partie est consacrée à l'optimisation conjointe du dictionnaire utilisée pour la QV et de la modulation. Dans la quatrième partie, nous introduisons le codeur canal dans la chaîne de transmission et nous présentons une solution pour l'optimisation conjointe de l'ensemble afin de garder une bonne qualité subjective des images reçues même en présence d'erreurs résiduelles très importantes.

2. Codage source

La compression d'une image peut être obtenue par la décomposition de celle-ci sur une base de fonctions orthogonales en utilisant la DCT [6], suivi d'un codage des paramètres exprimés dans la nouvelle base. Cette transformation projette l'image dans le domaine fréquentiel. Toute l'énergie est concentrée autour des composantes basses fréquences. Les composantes hautes fréquences sont souvent négligeables et, de plus, sont imperceptibles par l'oeil humain. Une bonne quantification des composantes basses fréquences suffit pour représenter l'image.

La quantification vectorielle [7] s'est avérée également très performante pour la compression d'images. Son principe consiste à coder une séquence d'échantillons (vecteur) au lieu de coder individuellement les échantillons. Le vecteur d'échantillons est présenté à un dictionnaire de vecteurs modèles. Le codage s'effectue en approximant le vecteur d'échantillons par le vecteur modèle le plus proche en terme de la distance euclidienne. Chacun des vecteurs du dictionnaire est indexé ; lors du codage, c'est l'indice du vecteur modèle associé qui servira pour représenter le vecteur d'entrée.

Nous avons utilisé les cartes auto-organisatrices de Kohonen [8] pour la construction du dictionnaire. L'intérêt de cette technique réside dans le fait que le dictionnaire construit à base de ces cartes présentent un ordonnancement topologique interne. Deux vecteurs voisins dans le dictionnaire sont similaires (au sens de la distance euclidienne). Cette propriété est exploitée lors de la modulation et du codage de canal pour immuniser le système contre les erreurs de transmission.

Dans cette étude, nous avons construit un dictionnaire bidimensionnel ayant 64 vecteurs modèles, chaque vecteur représente une imagerie de 2x2 pixels. Afin d'améliorer le taux de compression tout en sauvegardant une bonne qualité de l'image quantifiée, les deux techniques : "la DCT et la QV" sont combinées comme le montre la figure 1.

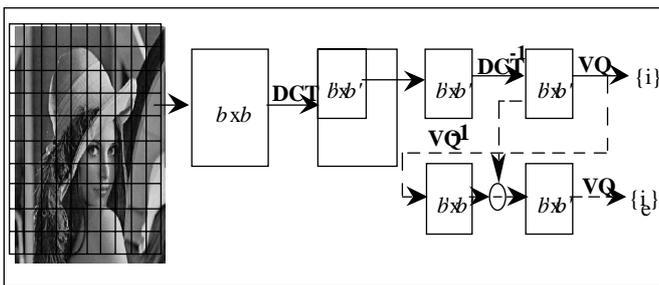


Figure 1 : Compression d'image combinant DCT et QV

L'image est partitionnée en imageries de $b \times b$ pixels. Pour chaque imagerie, on calcule la DCT pour la projeter dans le domaine fréquentiel. Les composantes hautes fréquences sont ensuite éliminées pour former un bloc de taille réduite $b' \times b'$ ($b' < b$). On opère une DCT^{-1} sur ce bloc pour obtenir une nouvelle imagerie de taille réduite $b' \times b'$ dans le domaine spatial. Les nouvelles imageries sont ensuite concaténées. L'image obtenue, qui n'est en fait qu'une réduction de l'image originale, sera compressée par la QV et ensuite transmise via le canal. A la réception, le décodeur reconstruit l'image réduite puis la divise en imageries de taille

$b' \times b'$. Pour chaque imagerie, il calcule la DCT, rajoute des zéros pour les composantes hautes fréquences éliminées au codage pour former un bloc de taille $b \times b$, puis opère une DCT^{-1} pour obtenir une imagerie de taille $b \times b$. Les imageries sont ensuite concaténées pour reconstruire l'image originale. Le système ainsi combiné peut atteindre des taux de compression allant jusqu'à 50.

L'image réduite apporte une première compression de l'image originale, la QV (seconde compression opérée) doit donc être très performante pour ne pas altérer l'image reconstruite (effet de blocs dû à la DCT). Pour cela, l'utilisation d'une QV supervisée ou classifiée s'impose.

Pour notre exemple, nous avons utilisé une QV non supervisée avec un dictionnaire ayant un petit nombre de vecteurs modèles (64) ce qui justifie l'utilisation du second étage (en pointillés) dans lequel on opère une seconde QV sur l'image erreur due à la première QV. Les deux dictionnaires utilisées pour l'image et l'image erreur sont de même taille. Les deux indices $\{i\}$ et $\{i_e\}$ représentant respectivement l'image et l'image erreur sont transmises et sont utilisées par le décodeur lors de la reconstruction. L'image Lenna quantifiée est montrée sur la figure 4-b, le taux de compression est égal à 16.77 avec un PSNR de 29.2dB.

3. Transmission d'image sur un canal gaussien avec une modulation MAQ-64

Chaque indice du dictionnaire est codé sur 6 bits (3 bits pour l'abscisse et 3 pour l'ordonnée). L'ensemble des indices associés à l'image quantifiée sont transmis sur un canal de Gauss en utilisant une modulation MAQ-64. Nous avons choisi le nombre d'états de la modulation égale à celui des vecteurs du dictionnaire. Ainsi chaque vecteur du dictionnaire est directement associé à un point de la constellation comme le montre la figure 2.

La modulation MAQ-64 peut être interprétée comme une superposition de deux MA-8 (Modulation d'Amplitude) qui représentent respectivement la voie en phase et la voie en quadrature. Dans le cas d'un canal de Gauss, la sortie d'un récepteur cohérent pour la voie en phase (ou en quadrature) est donnée par $r_j = e_j + w_j$ où e_j est le symbole émis appartenant à $\{\pm 7, \pm 5, \pm 3, \pm 1\}$ et w_j un bruit blanc gaussien. Les échantillons de bruit sur les voies en phase et en quadrature sont indépendants. Le décodage optimal des symboles reçus est obtenu en utilisant les seuils $\{\pm 6, \pm 4, \pm 2, 0\}$ sur les deux voies.

Pour les deux voies, la probabilité de transmettre un symbole e et de prendre une décision d avec $d \cdot e$ est donnée par la relation :

$$P_{ed} = \left(\frac{M-1}{M} \right) \operatorname{erfc} \left(\frac{d_E(e, d)}{2\sqrt{N_0}} \right) \quad (1)$$

où M est le nombre de symboles ($M=8$) et $d_E(e, d)$ est la distance euclidienne entre les symboles e et d . La probabilité de transition entre deux points de la constellation décroît exponentiellement avec la distance entre ces deux points (car la fonction $\operatorname{erfc}()$ est une fonction exponentielle décroissante).

Pour minimiser l'impact des erreurs de transmission sur l'image reçue, les imageries reçues erronées doivent ressembler visuellement aux images émises. Autrement dit une

imagette reçue erronée doit rester dans le voisinage topologique de celle émise.

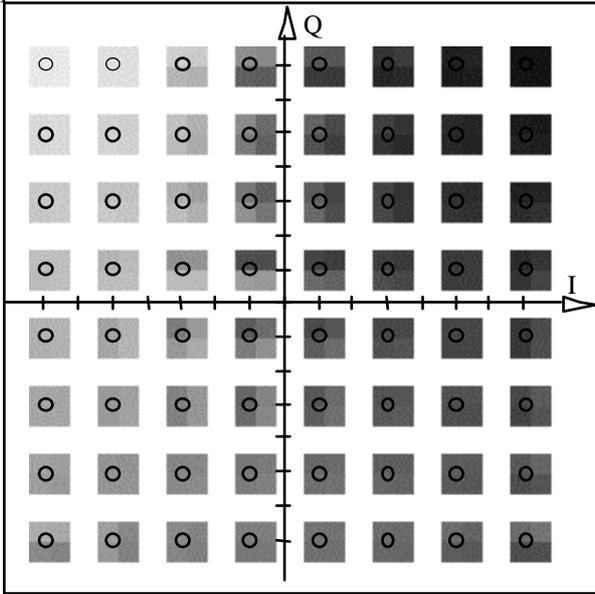


Figure 2 : Association des points de la constellation d'une MAQ-64 à un dictionnaire bi-dimensionnel de 64 vecteurs modèles utilisé pour la QV d'image.

En associant chaque point de la constellation à un vecteur modèle du dictionnaire comme le montre la figure 2, l'ordonnement interne du dictionnaire est sauvegardé puisque les erreurs de transmission les plus probables introduisent une faible distorsion de l'image : les vecteurs voisins sont similaires. Les erreurs qui produisent de fortes distorsions sont les moins probables au niveau de la modulation. Le système ainsi construit est très robuste aux erreurs de transmissions. Le résultat d'une transmission avec une MAQ-64 sur un canal de Gauss avec un Taux d'Erreur Symbole (TES) = 10% est donné dans la figure 4. L'image 4-c représente l'image reçue avec un dictionnaire non ordonné pour la MAQ-64, la qualité de l'image est très dégradée, le PSNR est égal à 19.93dB. L'image 4-d représente l'image reçue avec le dictionnaire ordonné décrit ci-dessus. On note que la qualité subjective de l'image est très bonne malgré un fort TES. L'image comporte 1238 erreurs, et le PSNR est égal à 27.58dB.

Dans le cas où le nombre de vecteurs dans le dictionnaire est différent de celui du nombre de points de la constellation (impossibilité d'association directe comme décrit ci-dessus), on peut adapter le dictionnaire lors de la construction au type de la modulation pour conserver la propriété de masquage des erreurs [10].

4. Introduction du codage canal

Le codage de canal est une technique très répandue dans les systèmes de transmission actuels afin d'améliorer la qualité de transmission. Nous avons introduit un code en bloc BCH dans la chaîne de transmission comme le montre la figure 3. Les codes BCH sont des codes en blocs cyclique qui consistent à concaténer k bits d'information avec $(n-k)$ bits de redondance. Les bits de redondances sont obtenus par une fonction linéaire sur les bits d'informations. Les n bits ainsi obtenus forment un mot de code. Un paramètre très important pour les codes en blocs est d (sa distance minimale de Hamming) ainsi que $t = \lfloor (d-1)/2 \rfloor$ son pouvoir de correction. Le mot de code est transmis via le canal et à la réception, le

décodage consiste à trouver le mot de code à distance minimale du mot reçu.

Le choix du code doit également préserver l'ordonnement établie dans le codeur source et la modulation afin d'avoir une bonne qualité subjective de l'image reçue entachée de quelques erreurs. Nous avons opté pour un code BCH (63,51,3) de pouvoir de correction $t=1$. Les coordonnées du vecteur modèle restent ainsi groupés dans le même symbole de la modulation et dans le même mot de code BCH ($57 = 19 \times 3$). Une erreur de transmission entachera qu'une seule imagette.

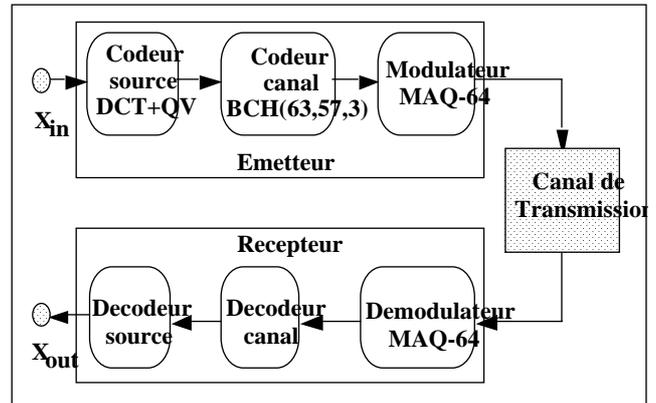


Figure 3 : Système de transmission d'images

Nous avons utilisé deux stratégies de décodage : un décodage algébrique (hard) [11] et un décodage pondéré (soft). Le décodage algébrique est basé sur la distance de Hamming, le mot le plus proche des échantillons reçus en terme de distance de Hamming est le mot décodé. Ce décodage ne respecte pas l'ordonnement établi auparavant et on peut noter la détérioration de l'image reçue après un décodage de ce type sur la figure 4-e. Le décodeur rajoute des erreurs générant de fortes distorsions quand il se trompe. Nous avons donc opté pour un décodage pondéré type Chase [9]. Le décodage cette fois-ci est basé sur la distance Euclidienne et de ce fait respecte l'ordonnement du dictionnaire et de la modulation. On remarque la nette amélioration de l'image reçue après décodage soft (l'image 4-f), un gain de 1dB par rapport à l'image sans codage canal est observé.

5. Conclusion

Nous présentons dans ce papier un nouveau système complet pour la transmission d'images. Ce système peut atteindre des taux de compression élevés tout en demeurant robuste aux erreurs de transmission ce qui est très difficile à obtenir avec les codeurs utilisés actuellement (exemple : JPEG). Les images reçues restent exploitables même si le TES est élevé. Le système proposé peut supporter des pointes de TES allant jusqu'à 30% [10]. L'optimisation conjointe de toutes les fonctions du système nous a permis d'accroître la robustesse vis à vis des erreurs dues au canal de transmission. Nous avons testé aussi la robustesse de notre système dans le cas d'une transmission sur un canal de Rayleigh, les performances obtenues demeurent comparables à celles obtenus sur un canal gaussien. Cependant, on note la présence de quelques erreurs isolées dues aux évanouissements sur le canal. Pour cela nous avons élaboré un post-traitement se basant sur le contexte spatial de l'image pour détecter et corriger ces quelques erreurs isolées [12].



Figure 4 : Transmission sur un canal de Gauss en présence d'une modulation MAQ-64 avec un Taux d'erreur symbole = 10^{-1} et un SNR = 11dB. (a) : Image originale. (b) : Image Quantifiée, taux de compression = 16.77, PSNR = 29,17dB. (c) : Image reçue avec un dictionnaire non ordonné sans codage canal, PSNR = 19.93dB. (d) : Image reçue avec un dictionnaire ordonné sans codage canal, PSNR = 27.58dB. (e) : Image obtenue avec un dictionnaire ordonné après un décodage canal hard, PSNR = 25.87dB. (f) : Image obtenue avec un dictionnaire ordonné après un décodage canal soft, PSNR = 28.58dB.

Remerciements

Ce travail a été financé par le Centre Commun d'Études de Télédiffusion et Télécommunications (CCETT) et le Centre National d'Études des Télécommunications (CNET).

Références

- [1] G. Ungerboeck, "Treillis-coded modulation with redundant signal sets," *IEEE Comm. magazine*, vol 25, Feb. 1987, pp. 5-21.
- [2] R. Pyndiah, A. Picart and A. Glavieux, "Performance of block turbo coded 16-QAM and 64-QAM modulations," in proc. of *IEEE GLOBECOM '95 Conference*, vol. 2/3, Nov. 1995, Singapour, pp. 1039-1044.
- [3] E. Ayanoglu and R. M. Gray, "The design, of joint source and channel treillis waveform coders," *IEEE Trans. Inform. Theory*, vol IT-33, No. 6, Nov. 1987, pp. 855-865.

- [4] N. Favardin, "A study of vector quantization for noisy channels," *IEEE Trans. Inform. Theory*, vol. IT-36, No. 4, July. 1990, pp.799-809.
- [5] K. Sayood, F. Liu and J. D. Gibson, "A constraint joint source/channel coder design," *IEEE Journal on Selected Areas in Communications*, vol 12, No. 9, Dec. 1994, pp. 1584-1592.
- [6] N. Ahmed, T.Natarajan and K. R. Rao, "Discrete cosine transform," *IEEE Trans. Comput.*, vol. C-23, pp. 90-93, 1974.
- [7] R. M. Gray, "Vector quantization," *IEEE Acoustic, Speech and Signal Processing Magazine*, vol. 1, pp. 4-29, Apr. 1984.
- [8] T. Kohonen, "Self organization and associative memory," *Springer-Verlag*, 1984, New York.
- [9] D. Chase, "A class of algorithms for decoding block codes with channel measurement information," *IEEE Trans. Inform. Theory*, vol IT-18, Jan. 1972, pp. 170-182.
- [10] O. Aitsab, R. Pyndiah and B. Solaiman, "Joint optimization of multi-dimensional SOFM codebooks with QAM modulations for vector quantized image transmission," *3rd Int. Workshop in Signal/Image processing*, 4-7 Nov. 1996, Manchester.
- [11] E. R. Berlekamp, "Algebraic coding theory," Mac Graw-Hill, 1968.
- [12] B.Solaiman, R.Pyndiah, O.Aitsab, G.Cazuguel and C.Roux, "Mesures de similarité floue pour la détection d'erreurs de transmission d'images avec quantification vectorielle", *CORESA '97*, 26-27 Mars 1997, Paris.