



Masquage des Erreurs de Transmission pour les Images JPEG.

M. Abdat et Z. Alkachouh

Conservatoire National des Arts et Métiers, 292 rue St-Martin, 75141 Paris Cedex 03, France.

RÉSUMÉ

Cet article présente une méthode pour atténuer les effets des erreurs de transmission dans les images JPEG. La détection et le masquage utilisés ne nécessitent aucune modification du codeur, seul le décodeur est modifié. Les techniques proposées sont causales, seuls les blocs précédant le bloc actuel sont utilisés. La détection des erreurs est basée sur la cohérence du format de transfert et du déroulement du processus de décodage, ainsi que sur les propriétés spatiales et fréquentielles des images naturelles. Le masquage des erreurs détectées est effectué dans le plan fréquentiel, ce qui permet de conserver l'activité fréquentielle de l'image pour le bloc corrigé. Les améliorations apportées en rapport signal à bruit crête pour des taux d'erreur de l'ordre de 10^{-3} sont de l'ordre de 5dB.

ABSTRACT

Digitally transmitted images suffer from quality degradation at high channel noise levels. This paper deals with an approach to counter transmission errors in JPEG coded images. The proposed error detection and concealment techniques are based on prediction, and they do not need modifications in the coder. The error detection technique exploits the transfer format coherence, the smoothness property of natural images at the block boundaries, and the frequency properties of the DCT coefficients. The error concealment of the damaged blocks is carried out in the frequency domain. Thus, the computational complexity is kept reasonably low, the image features are restored and the high frequency blocks are well recovered. The improvements achieved in terms of peak-to-peak signal-to-noise ratios are about 5 dB for bit error rates around 10^{-3} .

I. INTRODUCTION

Avec le développement du domaine du multimédia, des applications liées aux transmissions des images au format JPEG [1] sont envisagées, par exemple la transmission d'images fixes ou faiblement animées sur des réseaux locaux, sur le réseau téléphonique, et même sur des canaux numériques radio. Dans ces conditions, il devient important d'étudier la robustesse aux erreurs de transmission des algorithmes utilisés, afin de pouvoir prendre les mesures de protection appropriées. Une analyse de la robustesse aux erreurs de transmission du format de transfert JPEG a déjà été effectuée. Elle a fait apparaître la nécessité d'introduire des intervalles de reprises pour limiter la propagation des erreurs dans les données entropiques pour des taux d'erreurs de l'ordre de 10^{-4} [2].

Dans cet article, on s'intéresse à la détection et au masquage des erreurs résiduelles dans les images décodées. Le but est de concevoir un décodeur compatible JPEG résistant aux erreurs de transmission sans aucune surcharge de débit supplémentaire. Les techniques de détection et de correction utilisées sont causales, elles n'utilisent que l'information déjà décodée. Elles sont basées sur la prédiction du bloc en cours de décodage. Ainsi, dès sa réception le bloc est testé pour la détection des erreurs puis corrigé s'il en comporte.

La détection directe des blocs erronés se fait par plusieurs techniques successives. La première teste la cohérence de déroulement du processus de décodage, la seconde concerne la

détection des erreurs survenant sur les coefficients correspondant aux hautes fréquences et la troisième est basée sur la comparaison des blocs dans le domaine spatial. Une fois un bloc détecté comme erroné, il est corrigé dans le domaine fréquentiel à partir des blocs de son voisinage.

Dans cette communication, on présente d'abord dans II le processus de base de codage JPEG et le problème de la propagation des erreurs. Ensuite, les techniques de détection et de correction sont présentées dans les sections III et IV. Les résultats objectifs et subjectifs des simulations font l'objet de la section V. Enfin, quelques lignes de travaux complémentaires et orientations actuelles sont soulignées en conclusion.

II. LE PROCESSUS DE CODAGE DE BASE ET LA PROPAGATION DES ERREURS

Le processus de codage de base consiste à effectuer une transformée en cosinus discrète sur des blocs de 8x8 pixels de l'image, à quantifier les coefficients obtenus en utilisant les tables de quantification fournies et enfin à coder avec un codeur entropique les coefficients quantifiés. Comme montré sur la figure 1, avant le codage entropique, les coefficients continus (DC) des blocs sont codés en différentiel et les coefficients fréquentiels (AC) sont réarrangés dans un ordre zig-zag.

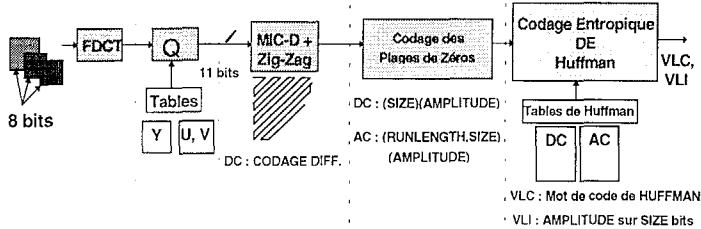


Figure 1. Le processus de codage de base JPEG.

Ensuite, chaque coefficient AC non nul est codé en deux symboles (RUN,SIZE) et (AMPLITUDE). Le premier symbole est constitué de deux parties chacune sur 4 bits, (RUN) représente le nombre de zéros consécutifs précédant le coefficient, et (SIZE) représente le nombre de bits sur lequel le deuxième symbole est codé. Le deuxième symbole (AMPLITUDE) représente la valeur du coefficient codée sur SIZE bits. Seul le symbole (RUN,SIZE) est codé par un codeur de Huffman.

En plus de la spécification des processus de codage et décodage, JPEG spécifie une syntaxe et organisation du flux de données en sortie du codeur et à l'entrée du décodeur. Ce format désigné par format de transfert consiste en un ensemble ordonné de paramètres, de marqueurs et de segments de données à codage entropique.

La présence d'une erreur de transmission dans les données entropiques peut se propager sur plusieurs coefficients DCT et même sur plusieurs blocs. Cette propagation est liée principalement à trois raisons :

- L'utilisation des codes à longueur variable de Huffman et le codage en entier à longueur variable de l'amplitude des coefficients.
- Le codage des plages de zéros.
- Le codage différentiel des coefficients DC.

La présence d'une erreur dans un mot de code de Huffman peut le transformer en un autre de longueur différente, donc conduire à une désynchronisation du décodeur. Afin de minimiser cette propagation, des codes de Huffman autosynchronisés peuvent être utilisés [4,5]. Suivant la définition donnée dans [4], les tables entropiques des coefficients DC données dans le standard JPEG contiennent plusieurs mots de synchronisation.

La compression par codage des plages de zéros propage aussi les erreurs au niveau bloc. En effet, la présence d'une erreur au niveau d'un symbole (RUN, SIZE) modifie la longueur de la plage de zéros lorsque RUN est erroné et peut conduire à la désynchronisation du décodeur dans le cas où SIZE est erroné. L'apparition accidentelle de fin de bloc (EOB : End Of Block) conduit à l'explosion d'un bloc en deux, donc au décalage à droite des blocs suivants. Par contre, une disparition accidentelle d'un EOB conduit à la fusion de deux blocs en un seul, donc au décalage à gauche des blocs suivants.

Enfin, une erreur qui survient au décodage d'une différence DC se propage sur tous les coefficients DC des blocs suivants jusqu'au prochain marqueur de synchronisation. La solution qui consiste à éviter le codage DPCM des coefficients DC est coûteuse, car une perte de l'ordre de 5 à 10% peut être enregistrée au niveau du rapport de compression.

A ces trois types de propagation se rajoutent la suppression et l'apparition accidentelle de marqueurs critiques qui peuvent conduire au blocage du décodeur ou à la perte d'intervalles de

reprise. Le choix des valeurs des marqueurs de synchronisation dans JPEG n'est sans doute pas le meilleur, car leurs valeurs 0xFFD0-0xFFD7 sont très proches de celles de marqueurs critiques pour le processus de décodage.

III. LA DETECTION DES ERREURS DE TRANSMISSION

La majorité des techniques proposées jusqu'à maintenant pour la détection des erreurs de transmission dans les images fixes ont comme principe la continuité de l'évolution des valeurs des pixels pour une région donnée de l'image[5]. D'autres techniques de détection dans le domaine fréquentiel ont été également considérées dans [6] ; elles s'appliquent à des codeurs où la propagation des erreurs ne se produit pas.

III.1. DETECTION PAR TEST D'INCOHERENCE DU FORMAT DE TRANSFERT ET DU PROCESSUS DE DECODAGE

La première incohérence concerne la lecture du flux en entrée du décodeur. Dans certains cas où il y a eu désynchronisation du décodeur, les données entropiques peuvent être épuisées avant la fin de l'intervalle de reprise ou au contraire, il peut rester plusieurs octets avant le marqueur de resynchronisation. Ce type d'erreur est signalé au niveau supérieur et permet de détecter certaines fusions et explosions de blocs.

Le test de cohérence du format de transfert permet de détecter des erreurs qui ont fait apparaître des marqueurs accidentels. Le respect de cette cohérence permet d'éviter des blocages du décodeur et des coupures de décodage.

En ce qui concerne les données entropiques reçues, plusieurs niveaux de cohérence sont testés. Les codes de Huffman utilisables sont connus au décodeur, la détection d'un code inexistant se produit si sa longueur excède la taille maximale permise pour les codes (16bits pour JPEG). L'action entreprise dans ce cas, consiste à ignorer le premier bit du mot lu et à recommencer l'identification des codes à partir du second.

La détection d'erreurs au niveau des plages de zéros repose sur le décodage de valeurs de SIZE et la succession de symboles (RUN,SIZE) incohérents. En effet, dans le cas du mode de base la valeur de SIZE du coefficient DC doit être inférieure à 12 et à 11 dans le cas des coefficients AC. A la détection d'une telle incohérence, le décodeur ne lit pas les prochains bits comme un symbole (AMPLITUDE) mais comme un symbole (RUN, SIZE). Aussi, la succession de deux EOB ou de 4 valeurs (15,0) représentent des cas d'erreurs qui sont rares mais pas impossibles. Le décodage des plages de zéros permet de détecter certains cas de fusion de blocs.

En plus de ces incohérences et en supposant que l'évolution des détails dans l'image est progressive, le nombre de bits utilisés pour le codage de blocs voisins évolue aussi de façon progressive. Deux tests de débits permettent de détecter des fusions et des explosions de blocs. Pour chaque bloc erroné, le débit du bloc D_b est comparé au débit moyen D_m et si

$D_b \leq 0.2 \cdot D_m$, alors le décodeur compense une explosion de bloc et si $D_b \geq 5 \cdot D_m$, le décodeur compense une fusion de deux blocs. Les facteurs 0.2 et 5 sont fixés de façon empirique, après plusieurs tests avec différentes images représentatives.

III.2. DETECTION DES ERREURS DANS LES COEFFICIENTS CORRESPONDANT AUX HAUTES FREQUENCES

Le coefficient DCT du rang courant dans la séquence zig-zag de chaque bloc est comparé avec les coefficients du même rang des blocs voisins montrés sur la figure 2.

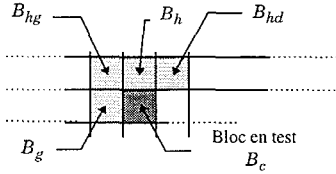


Figure 2 : Voisinage utilisé pour la détection et le masquage des erreurs.

Si k est le rang du coefficient courant, alors la procédure de comparaison est la suivante :

$$\text{si } |B_c(k)| \geq \alpha \cdot \left(\max(|B_g(k)|, |B_h(k)|, |B_{hg}(k)|, |B_{hd}(k)|) + 1 \right)$$

$$B_c(k) = \maxabs(B_g(k), B_h(k), B_{hg}(k), B_{hd}(k))$$

sinon

$$\text{si } |B_c(k)| \leq \beta \cdot \left(\min(|B_g(k)|, |B_h(k)|, |B_{hg}(k)|, |B_{hd}(k)|) + 1 \right)$$

$$B_c(k) = \minabs(B_g(k), B_h(k), B_{hg}(k), B_{hd}(k))$$

Les fonctions \min et \max renvoient respectivement le minimum et le maximum des valeurs à leur entrée. Les fonctions \minabs et \maxabs renvoient le coefficient dont la valeur absolue est respectivement minimale et maximale des coefficients en entrée. Le rajout d'un 1 pour les valeurs de test est fait pour éviter de fausses alertes dans le cas où tous les coefficients des blocs voisins sont nuls. Les coefficients de pondération α et β sont fixés par des tests empiriques, il sont dépendants du rang du coefficient en cours de test :

$$\alpha = (k+1) \text{ et } \beta = \frac{1}{\alpha}$$

III.3. DETECTION DANS LE PLAN SPATIAL

Les erreurs détectées correspondent principalement aux coefficients DC et se traduisent dans l'image par des blocs uniformes très différents de leur voisinage. Compte tenu que les évolutions des valeurs des pixels dans une image naturelle sont lisses, une variation brusque d'intensité entre les frontières de blocs correspond alors à un défaut. Pour détecter un bloc erroné, on utilise une technique statistique de détection de contours, classique en traitement d'images[7]. Dans notre cas, la position du contour est connue, et il reste à vérifier si le contour existe. Soient m_1 et m_2 les moyennes respectives des frontières des blocs et σ_1^2 , σ_2^2 leurs variances. La comparaison de deux blocs s'effectue par l'intermédiaire des valeurs moyennes de leurs frontières en comparant la valeur t à un seuil déterminé.

$$t = 2 \cdot \frac{|m_1 - m_2|}{\sigma_d} \text{ avec } \sigma_d^2 = \frac{\sigma_1^2 + \sigma_2^2}{2}$$

La valeur statistique optimale de t pour détecter un contour naturel est $t_{opt} = 2.92$, pour une probabilité d'erreur de 0.01 et une longueur de 8 pixels de chaque côté du contour[7]. Les frontières utilisées pour le test sont celles en haut et à gauche du bloc. Un bloc est alors déclaré comme erroné lorsque $t_h \geq t_s$ et

$t_g \geq t_s$. Le seuil de comparaison t_s est un seuil dynamique et sa valeur est suffisamment grande pour ne pas détecter les contours naturels de l'image. Il est initialisé par le double de la valeur de t_{opt} ($t_s = 5.82$), ensuite pour chaque bloc non erroné, on calcule la moyenne $t_m = (t_h + t_g)/2$ qui est mémorisée pour être utilisée dans le seuil du bloc suivant. Le seuil pour un bloc donné est alors : $t_s = t_m(\text{dernier bloc non erroné}) + t_{opt}$.

IV. MASQUAGE DES ERREURS DETECTEES

Suivant le type de l'erreur détectée, le décodeur peut être amené à effectuer différentes actions de compensation : des actions de synchronisation, des actions de limitation de la propagation de l'erreur et enfin des actions de masquage et correction du bloc ou coefficient erroné.

La recherche des points de synchronisation est lancée à la suite d'une incohérence au niveau du processus de synchronisation telles que l'absence de marqueur de synchronisation et la rencontre d'un marqueur différent du marqueur attendu.

La limitation de la propagation des erreurs au niveau bloc consiste à repositionner les blocs décalés par une fusion ou explosion, et à répercuter la correction d'un coefficient DC sur les coefficients des blocs suivant le bloc erroné. Dans le cas d'une explosion de bloc, le décodeur corrige les coefficients fréquentiels (AC) et ignore les données entropiques jusqu'au prochain EOB. Dans le cas d'une fusion de deux blocs, en plus des actions précédentes, le bloc suivant est déclaré erroné et le décodeur passe directement à sa compensation.

Le masquage d'un bloc erroné consiste à prédire ses coefficients fréquentiels en utilisant le voisinage montré sur la figure 2. Si une erreur signalée par les tests d'incohérences est détectée au niveau du coefficient DC ou d'un coefficient basse fréquence (rang du coefficient dans l'ordre zig-zag inférieur à 6), alors on suppose que tout le bloc est erroné, car il est fort probable que l'erreur ait commencé plus tôt. Dans le cas contraire, seuls les coefficients suivant celui erroné sont corrigés. Pour une erreur signalée par la détection fréquentielle, la correction porte uniquement sur le coefficient dans lequel l'erreur est détectée. Par contre, pour une erreur signalée par la détection spatiale, tout le bloc est considéré comme faux.

Le coefficient DC est corrigé par la moyenne des coefficients DC des quatre blocs voisins et un coefficient AC de rang k dans l'ordre zig-zag, est corrigé de la façon suivante :

$$B_c(k) = \left(\maxabs(B_g(k), B_h(k)) + \maxabs(B_{hg}(k), B_{hd}(k)) \right) / 2$$

Cette correction est faite de façon à tenir compte de contours horizontaux ou verticaux et de contours diagonaux.

V. RESULTATS

L'image utilisée est "Teenager", contenant 576 lignes et 536 pixels par ligne et une seule composante (luminance). Elle est codée avec des intervalles de reprise de 15 MCU, chaque MCU correspondant à 1 bloc de 8x8 pixels. La diminution du taux de compression par utilisation d'intervalles de reprise de longueur 15 MCU/RST est de 2.76%. L'image décodée sans erreurs a un rapport signal sur bruit crête (PSNR) de 40.05dB, l'erreur quadratique moyenne étant calculée par rapport à l'image



originale. La figure 3 donne l'évolution du rapport signal sur bruit crête de l'image décodée pour différents taux d'erreurs binaires, dans le cas d'un décodeur JPEG conventionnel et dans le cas du décodeur proposé.

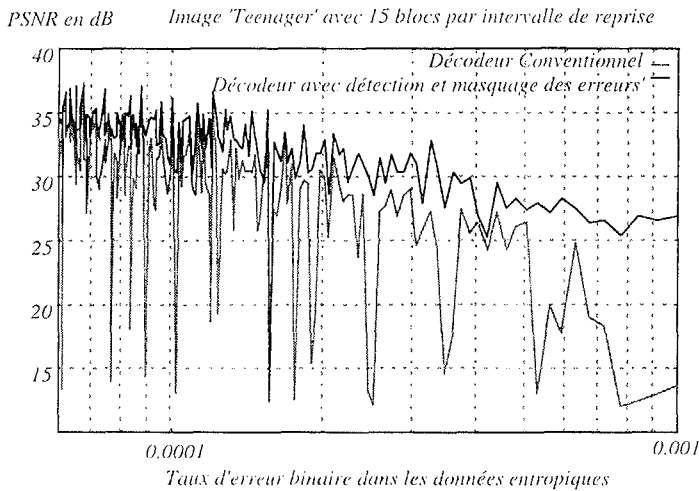


Figure 3 : Evolution du PSNR en fonction du taux d'erreur binaire dans les données entropique de l'image "Teenager" Codée avec 15 MCU par intervalle de reprise.

Elle montre que le décodeur proposé améliore nettement la qualité de l'image décodée. Cette amélioration est particulièrement significative pour les forts taux d'erreur. Les chutes du PSNR dans le cas du décodeur conventionnel correspondent à des coupures au niveau du processus de décodage. A cause de l'incohérence de la syntaxe du format de transfert reçu, seule une partie de l'image est décodée.

En ce qui concerne les résultats subjectifs, Les figures 4, 5 et 6 montrent respectivement l'image décodée sans erreurs, l'image décodée avec le décodeur conventionnel pour un taux d'erreur de 10^{-3} et l'image décodée par le décodeur proposé pour le même taux d'erreur. Toutes les trois ayant été codées avec 15 MCU par intervalle de reprise.

VI. CONCLUSION

Nous avons proposé un décodeur de base JPEG intrinsèquement résistant aux erreurs de transmission. Il utilise des techniques de détection et de masquage d'erreurs causales, et se prête à une mise en oeuvre en temps réel. En plus de sa compatibilité avec JPEG, il améliore nettement le rapport signal sur bruit par rapport au décodeur conventionnel, sans aucune surcharge du débit en sortie du codeur. Il élimine quasiment toutes les coupures du processus de décodage par le test de cohérence du format de transfert reçu.

Une adaptation des techniques utilisées au cas du mode étendu de JPEG avec le codeur arithmétique est en cours. Ensuite, il est envisagé d'examiner les techniques de détection et de masquage basées sur l'interpolation à la place de la prédiction.

Références :

[1] ISO/CEI JTC1, IS 10918, "Technologies de l'information: compression numérique et codage des images fixes de nature photographique", Partie 1 : Spécifications et Directives, 1991.

- [2] M. Abdal and M. G. Bellanger, " Three JPEG coding Schemes and their Robustness to Transmission Errors", EUSIPCO '94, Edinburg (Scotland), September 1994.
- [3] T. J. Ferguson and J. H. Rabinowitz, " Self-Synchronising Huffman Codes", IEEE Transactions on Information Theory, Vol. IT-30, No. 4, pp. 687-693, July 1984.
- [4] W. M. Lam and A. R. Reibman, "Self-Synchronizing Variable-Length Codes for Images Transmission", ICASSP'1992, Vol. III, pp. 477-480, 1992.
- [5] J. W. Park, D. S. Kim and S. U. Lee, " On the Error Concealment Technique for DCT Based Image Coding", ICASSP'94, pp. III-293-III-296, 1994.
- [6] O. R. Mitchell and A. J. Tabatabai, "Channel Error Recovery for Transform Image Coding", IEEE Trans. on Comm., Vol.29, No.12, December 1981, pp.1741-1754.
- [7] M. Kunt, G. Granlund et M. Kocher, "Traitement Numérique des Images", Traitement de l'information: volume 2. Presses polytechniques et universitaires romandes, Juillet 1993.



Figure 4 : Image "Teenager" décodée sans erreurs (PSNR=40.05dB)



Figure 5 : Image "Teenager" décodée avec le décodeur conventionnel ($BER=10^{-3}$, PSNR=21.19dB).



Figure 6 : Image "Teenager" décodée avec le décodeur proposé ($BER=10^{-3}$, PSNR=26.51dB).