

Résistance aux erreurs de l'algorithme JPEG pour la compression des images.

M. Abdat, M. Bellanger

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

RÉSUMÉ

Cet article présente une analyse du format de transfert de la norme de compression d'images fixes de nature photographique (JPEG). La résistance aux erreurs de transmission a été analysée et des tests ont été effectués. Les résultats de ces tests nous ont amenés à proposer une classification des différents champs de bits constituant le format de transfert, suivant leur importance dans le processus de décodage.

Pour combattre les erreurs, deux techniques peuvent être envisagées. La première consiste à introduire une redondance adaptée pour chaque classe au niveau du codeur. La deuxième technique, dite de masquage, consiste à estimer les erreurs qui ont entaché l'image, et à les soustraire.

ABSTRACT

An analysis of the compressed data format for the JPEG (Joint Photographic Experts Group) still image compression standard is presented and the effect of bit errors is investigated. A classification with three classes is proposed for the data.

To combat the error effects, two methods can be envisaged. The first one consists of introducing an adapted redundancy for each class. The second method is based on error estimation and subtraction in the decoded image.

I- INTRODUCTION

Depuis plusieurs années, l'ISO (International Standardization Organization) mène des travaux sur la normalisation des techniques de compression des images, et une norme portant sur la compression des images fixes vient d'être approuvée, désignée par IS 10918 ou plus communément JPEG (Joint Photographic Experts Group)[1][2]. Bien que le domaine moteur pour ces travaux ait été la microinformatique, des applications liées aux transmissions sont envisagées, par exemple la transmission d'images fixes sur des réseaux locaux ou sur le réseau téléphonique, 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.

La norme JPEG porte sur le codeur, le décodeur et le format de transfert[1]. La spécification du format de transfert (format de données comprimées) est très importante, car elle assure la compatibilité entre le codeur et le décodeur et entre différentes applications. Le format de transfert est un ensemble ordonné de paramètres, de marqueurs et de segments à codage entropique. Syntaxiquement, le format de données comprimées est constitué de plusieurs niveaux : trame, balayage, codage entropique et unité codée minimale. Chaque niveau possède un en-tête suivi du contenu spécifié par celui-ci.

Dans la présente communication, on effectue d'abord une classification des différents groupes de bits suivant leur importance au décodage. Trois classes sont proposées : la première regroupe les paramètres indispensables au décodage. Une erreur sur l'un de ces paramètres rend le décodage impossible ou catastrophique. La deuxième catégorie est constituée d'un ensemble de paramètres qui sont importants pour le décodage. En cas d'erreur, le décodage peut se faire mais la qualité de l'image décodée est mauvaise. Enfin, la troisième classe regroupe les paramètres pour lesquels, même s'ils sont entachés d'erreurs la qualité de l'image n'est pas très affectée.

Ensuite, on étudie les types de protection du format de transfert. Une première approche de protection contre les erreurs consiste en une protection séparée et graduée pour chacune des trois classes. Dans le souci de préserver la compatibilité du format de transfert, il est souhaitable que le bloc protecteur vienne en fin de trame, ou occupe les champs réservés aux commentaires dans le format de transfert.

II- RAPPEL SUR LE FORMAT DE TRANSFERT

Dans la norme JPEG, afin de couvrir une large gamme d'applications, plusieurs modes de codage ont été spécifiés dont un codage de base où l'image comprimée puis décompressée n'est plus identique (au sens du bit) à l'image originale. Ce processus de codage de base utilise la transformée en cosinus discrète (DCT : Discrete Cosine Transform) suivie d'une quantification des coefficients et d'un codage entropique de Huffman (figure 1)[2].

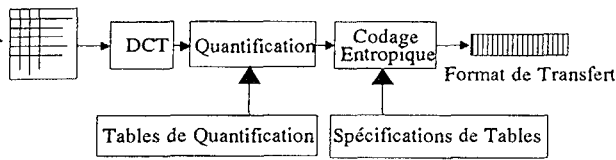


Figure 1 : Principe du codeur de Base JPEG.

Les autres types diffèrent du codage de base principalement par le codage entropique, en utilisant un codage arithmétique, ou par la restitution progressive de l'image. Le codage sans pertes est un codage non basé DCT.

Structurellement, le format de données comprimées consiste en un ensemble ordonné de paramètres, de marqueurs et de segments de données à codage entropique. Les paramètres ont des valeurs spécifiques aux processus de codage, aux caractéristiques de l'image source et à d'autres fonctions relevant de l'application. Les paramètres codent des informations critiques sans lesquelles le processus de décodage ne pourrait pas correctement reconstruire l'image. Les marqueurs servent à identifier les différentes parties structurelles des formats de données comprimées. La plupart des marqueurs servent comme en-tête pour un groupe de paramètres.

La figure 2 présente la syntaxe et l'organisation haut niveau du format de transfert (cas du mode hiérarchique exclu)[1]. Le format de transfert doit débuter par un marqueur SOI (Start Of Image), contenir une trame (dans le cas du mode hiérarchique, le format de données comprimées peut contenir plusieurs trames) et se terminer par un marqueur EOI (End Of Image). Une trame est constituée d'un ou plusieurs segments de spécification de tables et/ou de segments optionnels. Elle doit obligatoirement contenir un entête de trame et un (ou plusieurs) balayage(s) de l'image. Le troisième niveau de la figure 2 spécifie qu'un balayage doit commencer par un entête de balayage et doit contenir un ou plusieurs segments de données à codage entropique. L'entête peut être précédé d'un ou plusieurs segments de spécification de tables et/ou optionnels. Les marqueurs RSTi (ReStart marker) sont des marqueurs de définition des intervalles de reprise pour les données entropiques. A la rencontre de ce marqueur, le décodeur doit réinitialiser tous les paramètres de décodage entropique (Huffman ou arithmétique) pour continuer le décodage. Le quatrième niveau spécifie que chaque segment entropique (ECSi) doit contenir une séquence de MCU (Minimum Coded Unit) à codage entropique.

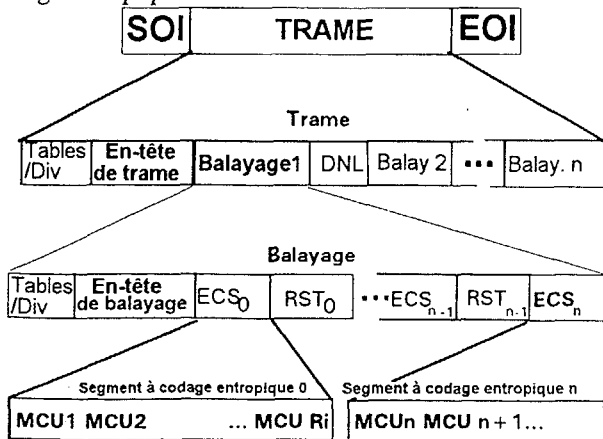


Figure 2 : Syntaxe de haut niveau du format de transfert (mode hiérarchique exclu).

III-ANALYSE DES MARQUEURS ET PARAMETRES DU FORMAT DE TRANSFERT

Cette analyse consiste à déterminer l'importance de chaque paramètre et segment du format de transfert dans le processus de décodage dans les cas des modes de codage de base et étendu. La méthode d'analyse consiste à altérer chaque paramètre, décodé l'image correspondante et à évaluer la qualité de celle-ci par rapport à l'image décodée sans erreurs. Pour chaque altération et composante de l'image, l'erreur quadratique moyenne (EQM) et le rapport signal sur bruit (S/B) sont calculés respectivement suivant les formules (1) et (2).

$$EQM = 10 \cdot \log_2 \left\{ \frac{1}{X \cdot Y} \sum_{i=0}^{X-1} \sum_{j=0}^{Y-1} (p_{ij} - \hat{p}_{ij})^2 \right\} \quad \dots(1)$$

$$S/B = \frac{10 \cdot \log_2 \left\{ \frac{1}{X \cdot Y} \sum_{i=0}^{X-1} \sum_{j=0}^{Y-1} (p_{ij})^2 \right\}}{EQM} \quad \dots(2)$$

où :

p_{ij} et \hat{p}_{ij} désignent respectivement, la valeur du pixel j de la ligne i de l'image décodée sans erreurs et de l'image décodée avec erreurs. X et Y désignent respectivement le nombre de lignes et de pixels par ligne de l'image.

Les erreurs quadratiques moyennes et les rapports signal sur bruit totaux (pour toute l'image) sont respectivement la somme de ceux des composantes constituant l'image. Les valeurs obtenues dans cette analyse correspondent à des dégradations très visibles, le seuil de perception correspondant à des valeurs de S/B de l'ordre de 30dB.

III-1-MARQUEURS ET PARAMETRES DE L'EN-TETE DE TRAME ET DE BALAYAGE

La majorité des marqueurs des en-têtes de trame et de balayage sont indispensables pour le décodeur. L'altération de l'un de ces paramètres rend le décodage impossible ou catastrophique. En effet, les en-têtes de trame et de balayage contiennent les informations essentielles de l'image comprimée telles que le processus de codage utilisé, le nombre de composantes de l'image, le nombre de lignes, le nombre de colonnes, les identificateurs des tables de quantification etc. Les deux en-têtes font une taille de l'ordre de 49 octets (cas d'image avec trois composantes non entrelacées (3 balayages)).

III-2-PARAMETRES DES SEGMENTS DE DEFINITION DES TABLES DE QUANTIFICATION

Ces segments contiennent les pas de quantification des coefficients DCT. L'altération des pas de quantification du coefficient DC et des coefficients associés aux basses fréquences ont des conséquences nettement plus néfastes que celles occasionnées par l'altération des pas de quantification des coefficients associés aux hautes fréquences, dont la plupart sont nuls. La table 1 illustre cette sensibilité.

Pas de quantification altéré	EQM (dB)	S/B(dB)
Pas de quantification DC (Q0)	38.10	8.28
Pas de quantification basses freq.	29.28	17.10

Table 1 : Erreurs quadratiques moyennes (EQM) et rapports signal sur bruit (S/B) des images décodées avec altération des pas de quantification. Pour le cas testé, l'image est insensible aux modifications des pas de quantification très hautes fréquences (image "teenager" avec 1.05bits/pixel).

III-3-PARAMETRES DE DEFINITION DES TABLES DE HUFFMAN

Les tables de Huffman sont utilisées pour le codage et décodage entropiques. L'altération des tables de codage entraîne l'altération de plusieurs mots décodés. Le tableau 2 illustre l'influence de ces altérations sur l'image décodée.

Type de l'erreur	EQM(dB)	S/B(dB)
Modification d'un mot de code dans la table DC de Huffman	46.07	0.32
Modification d'un mot de code dans la table AC de Huffman	40.36	6.03
Modification du nombre de mots de codes pour une longueur donnée.	44.89	1.49

Table 2 : Erreurs Quadratiques moyennes (EQM) et rapports signal sur bruit (S/B) induits par altération des tables de Huffman (image "teenager" avec 1.05bits/pixel).

III-4-LES DONNEES ENTROPIQUES

Elles constituent la partie la plus importante en taille de l'image comprimée. Les résultats de tests obtenus sont différents pour les deux codages, Huffman et arithmétique. Les données entropiques de Huffman sont constituées d'une suite de mots de codes à longueurs variables. La propriété principale des codes de Huffman est le respect de la condition préfixe. Si un mot de code dans les données entropiques est altéré, plusieurs conséquences peuvent se produire. Le mot de code altéré donne naissance à un autre mot de code de même longueur. Ce cas est le plus favorable ; en effet, seul le mot de code altéré est mal décodé. On aura une erreur de décodage sur un coefficient (au plus sur quelques coefficients si des plages de zéros sont codées dans le mot altéré). L'image décodée reste toujours de bonne qualité par rapport à celle décodée sans erreurs. Le deuxième cas correspond à la situation où le mot de code altéré donne naissance à un autre mot de code de longueur différente. Les résultats peuvent être catastrophiques comme ils peuvent être acceptables, tout dépend des mots de codes qui suivent le mot altéré. En effet, supposons que le mot de code altéré ait donné naissance à un autre mot de code plus court. Les bits restants du mot de code peuvent désynchroniser complètement le décodeur d'où une avalanche d'erreurs. Le décodage n'est rétabli que lorsque le décodeur est recalé sur un mot de code correct. Dans un tel cas, une resynchronisation rapide du décodeur est nécessaire pour un décodage correct. Le dernier cas correspond au cas où le mot de code altéré ne donne naissance à aucun autre mot de code existant. Ce cas est intermédiaire entre les deux premiers. Les résultats peuvent être similaires à ceux du premier cas, lorsque le décodeur est tout de suite resynchronisé sur le mot de code suivant. Autrement, les résultats sont similaires à ceux du deuxième cas. Pour tous les cas, la désynchronisation du décodeur peut être catastrophique pour tout le décodage de l'image. L'utilisation de plusieurs tables de Huffman (une par composante) peut s'avérer bénéfique pour ce type d'erreurs. Les erreurs quadratiques moyennes et les rapports signal sur bruit obtenus dépendent du type de l'erreur. La table 3 présente certains résultats obtenus :

Type de l'erreur	EQM (dB)	S/B (dB)
Modification d'un mot de code au début des données entropiques	28.66	17.72
Modification d'un mot de code au milieu des données entropiques	41.30	5.08
Modification d'un mot de code en fin de données entropiques	18.30	28.08

Table 3 : Erreurs Quadratiques moyennes (EQM) et rapports signal sur bruit (S/B) induits par altération de mots de codes dans les données entropiques de Huffman.

Avec les données entropiques arithmétiques, à partir du moment où le décodeur rencontre l'erreur, la suite du décodage est catastrophique. Les données altérées le sont d'autant plus qu'on accumule les erreurs. Cette grande altération des données décodées est due d'abord à la propriété principale du codage arithmétique qui est la double récursion[5][6]. La subdivision des intervalles étant récursive, une erreur sur un intervalle particulier entraîne des erreurs sur les intervalles suivants. La deuxième cause est le calcul adaptatif des contextes[6]. Une erreur sur les symboles déjà décodés induit des erreurs sur les contextes suivants. Les erreurs quadratiques moyennes et les rapports signal sur bruit d'image varient suivant la localisation de l'erreur dans les données entropiques, comme indiqué sur le tableau 4.

Type d'erreur	EQM (dB)	S/B (dB)
Inversion d'un bit au début des données entropiques.	44.48	1.90
Inversion d'un bit au milieu des données entropiques	42.63	3.75
Inversion d'un bit en fin de données entropiques	22.95	23.43

Table 4 : Erreurs Quadratiques moyennes (EQM) et rapports signal sur bruit (S/B) induits par altération des données entropiques arithmétiques.

La comparaison avec le tableau 3 montre bien que les données entropiques arithmétiques sont beaucoup plus vulnérables que les données entropiques de Huffman. Dans les cas testés, il n'y a qu'un seul segment entropique, ce qui est une situation défavorable. L'utilisation des intervalles de reprises peut s'avérer bénéfique. En effet, dans ce cas, l'image est codée sur plusieurs segments entropiques, et à chaque début de segment, on réinitialise le codeur et le décodeur, en oubliant ce qui a été codé précédemment. Par contre en réinitialisant le codeur et le décodeur, on perd en taux de compression.

IV-CLASSIFICATION DES MARQUEURS ET PARAMETRES DU FORMAT DE TRANSFERT

Nous proposons trois classes. La première classe nécessite une protection maximale. Elle correspond à tous les marqueurs et paramètres importants du format de transfert : les en-têtes de trame, les en-têtes de balayage, les pas de quantification basses fréquences, les tables de codage de Huffman, les tables décisionnelles pour le codage



arithmétique et les segments de définition du nombre de lignes. La taille de cette classe varie suivant le type du codage utilisé, et les caractéristiques de l'image à coder, d'une cinquantaine d'octets à 130 octets (tables de Huffman exclues). La deuxième classe, correspond aux données entropiques. C'est la plus importante en taille, plusieurs dizaines de kilooctets. Son importance varie selon la position des bits concernés dans le train binaire et le processus utilisé. La troisième classe correspond aux paramètres et marqueurs, dont dépend peu l'image, tels les segments application (commentaires), ou certains paramètres conditionnels ainsi que les pas de quantification des coefficients hautes fréquences (11-63) des blocs. Sa protection n'apparaît pas nécessaire.

A titre d'exemple, pour une des images utilisées avec une compression à 1bit/pixel, la répartition s'effectue comme suit:

Classe 1 : 3992 bits
 Classe 2 : 318912 bits
 Classe 3 : 992 bits.

V- PROTECTION ET MASQUAGE DES ERREURS

On peut envisager deux approches pour combattre les erreurs. La première consiste à protéger le format de transfert d'une manière graduée. Chacune des trois classes est protégée suivant son importance dans le processus de décodage. Pour la classe des en-têtes (première classe), des codes correcteurs d'erreurs sont recommandés. La deuxième classe (données entropiques) est très importante en taille, et il n'est pas réaliste de la protéger de la même façon que la première. On peut réduire l'influence des erreurs en utilisant plusieurs intervalles de reprises, ce qui est surtout recommandé pour le cas du codage entropique arithmétique. La deuxième approche celle du masquage, consiste à estimer les erreurs ayant entaché l'image décodée et à les soustraire. Elle est difficile à mettre en oeuvre dans le cas de JPEG à cause de la présence des codeurs entropiques. En effet, les techniques de masquage sont fondées sur les comparaisons de blocs de pixels voisins pour les processus de codage comprenant uniquement une DCT et une quantification, et, si plusieurs blocs voisins sont altérés, elles ne peuvent pas donner de bons résultats[4]. Les techniques de masquages qui peuvent être envisagées, devraient exploiter les principes des codages entropiques (Huffman et Arithmétique).

Le masquage des erreurs n'introduit aucun traitement au codeur ; par contre, les calculs sont importants dans le processus de décodage. Il présente un grand intérêt par le fait que ni le format de transfert ni le taux de compression ne sont modifiés.

V-CONCLUSION

Nous avons présenté une analyse de la sensibilité aux erreurs du format de transfert de la norme JPEG. Les tests effectués sont estimés représentatifs pour des taux d'erreurs de l'ordre de 10^{-5} . Cette analyse a montré que les différents champs constituant le format de transfert ont une importance différente pour le processus de décodage. La classification présentée illustre l'importance de la classe des en-têtes (première classe). En ce qui concerne la deuxième classe, la

sensibilité dépend du processus de codage utilisé, et une seule erreur peut engendrer une avalanche d'erreurs de décodage. Nous avons pu vérifier que les données entropiques arithmétiques sont plus vulnérables que celles de Huffman, principalement en raison du caractère récursif et adaptatif de la technique de codage arithmétique.

Le masquage des erreurs au décodage est une approche attrayante, mais difficile à mettre en oeuvre avec les codeurs entropiques. Des travaux complémentaires sont en cours sur ce point.

L'analyse faite pour la norme JPEG est utile également pour la compression d'images animées. En effet, la norme MPEG (Moving Pictures Experts Group) utilise une compression JPEG toutes les 12 trames (0.5 secondes). Une extension de l'étude pour MPEG est envisagée.

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] Gregory K. Wallace, "The JPEG still picture compression standard", Communications of the ACM, Vol.34, No.4, April 1991, pp. 31-44.
- [3] K. Joseph, S. Ng, D. Raychaudhuri, R. Siracusa, J. Zdepski, R. Saint Grons and T. Savatier, "MPEG++ : A robust compression and transport system for digital HDTV", Signal processing : Image communication 4, Vol.4, Nos. 4-5, August 1992, Elsevier Science Publishers, pp. 307-323.
- [4] 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.
- [5] W. B. Pennebaker, J. L. Mitchell, G. G. Longdon Jr, R. B. Arps, "An overview of the basic principles of the Q-coder", IBM Journal of Research and Development, Vol. 32, No. 6, Nov. 1988, pp. 717-726.
- [6] G. G. Longdon Jr. , "An introduction to arithmetic coding", IBM Journal of Research and Development, Vol. 28, No. 2, Nov. 1988, pp. 717-726.