

## DECODEUR VIDEO MPEG-2 RESISTANT AUX ERREURS RESIDUELLES DE DIFFUSION PAR SATELLITE

Christophe ALEXANDRE, Han VU THIEN

Laboratoire Electronique et Communications, CNAM, 292 rue saint Martin, 75141 PARIS CEDEX 03

### RESUME

Malgré l'utilisation des codes convolutionnel et Reed-Solomon pour protéger le train binaire MPEG-2 dans une diffusion par satellite, il reste des erreurs résiduelles à l'entrée du décodeur vidéo. L'élaboration d'une stratégie de masquage est donc indispensable pour préserver la qualité de l'image. Nous avons mis au point deux méthodes. Dans la première, le décodeur utilise les informations fournies par la couche système afin de masquer les paquets erronés. Dans la deuxième, on exploite les caractéristiques du train binaire pour réaliser un décodeur "intelligent" capable de détecter et de masquer de manière autonome les erreurs. Nous montrons dans cet article que le décodeur "intelligent" donne des images dont la qualité subjective est la meilleure en cas d'erreurs résiduelles.

### 1. INTRODUCTION

L'évolution technique dans la réalisation des circuits intégrés ainsi que la normalisation dans les domaines de la compression de l'image et du son [1][2][3] et de la diffusion par satellite [4] permettent aujourd'hui la mise en place d'un service commercial de diffusion par satellite de télévision numérique. Dans un tel système, on considère souvent que le démodulateur du récepteur satellite cesse de fonctionner avant les codes correcteurs d'erreurs (CCE) quand le rapport signal sur bruit (C/N) diminue, ce qui implique une coupure complète de la réception. Or on sait faire aujourd'hui des démodulateurs plus robustes et ce sont généralement les CCE qui limitent les performances du système en laissant passer un grand nombre d'erreurs résiduelles. Le but de cette communication est d'exposer différentes stratégies de masquage d'erreurs utilisable dans un décodeur vidéo et de montrer leur influence sur la qualité de l'image. L'intérêt d'un tel masquage est d'améliorer la robustesse du récepteur afin d'obtenir des images de bonne qualité avec un taux d'erreurs élevé. Nous allons d'abord rappeler les différentes étapes composant la chaîne de diffusion.

### ABSTRACT

In spite of the use of convolutional and Reed-Solomon codes in a satellite broadcasting to protect the MPEG-2 bitstream, residual errors will be found at the video decoder's input. Concealment methods must be elaborated to preserve images quality. Two strategies has been elaborated. In the first one, the decoder uses informations provided by the system layer to conceal false packets. In the second one, a "smart" decoder is designed which uses bitstream properties to detect and conceal residual errors. In this paper, we want to show that the "smart" decoder gives better results in terms of images quality than the first decoder when residual errors occur in the bitstream.

### 2. LE SYSTEME DE DIFFUSION

Les différents éléments composant un système complet de diffusion par satellite de télévision numérique sont représentés ci-dessous :

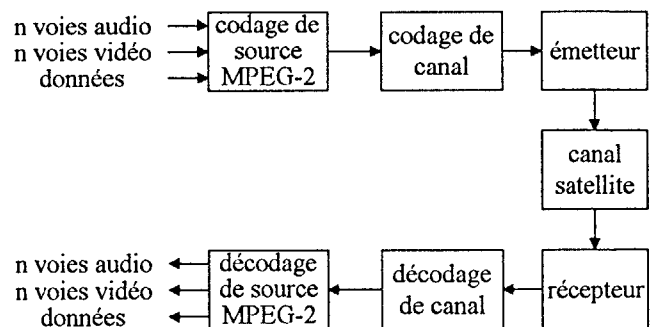


Figure 1 : schéma simplifié d'un système de diffusion par satellite de télévision numérique

Une configuration typique permet de faire passer, dans un canal satellite type « TELECOM 2 » de largeur 36 MHz (soit un débit d'environ 38 Mbit/s), un programme comprenant :



- six voies de données 100 kbit/s permettant la gestion des abonnés (débit : 0,6 Mbit/s).
- deux voies vidéo de qualité studio au format 4/3 ou 16/9 (débit :  $2 \times 6 = 12$  Mbit/s).
- quatre voies vidéo de qualité « PAL » au format 4/3 (débit :  $4 \times 3,5 = 14$  Mbit/s).
- trente voies audio de qualité « Compact-Disc » (débit :  $30 \times 0,192 = 6$  Mbit/s).
- une voie système pour la diffusion du guide de programme (débit : 5,5 Mbit/s).

Le rôle des différents éléments composant le système est le suivant.

Le codage de source MPEG-2 permet de former le train binaire système. Il comprend la compression des signaux audio et vidéo, la mise en paquets élémentaires (PES), la mise en paquets de transport comprenant 184 octets de données et enfin le multiplexage de ces paquets.

Le codage de canal adapte le message à transmettre au canal de diffusion grâce à un embrouilleur rendant le train binaire aléatoire suivi d'un codeur Reed-Solomon(188,204), d'un entrelaceur convolutionnel ayant une profondeur d'entrelacement de 12 paquets et d'un codeur convolutionnel ayant un rendement  $\frac{1}{2}$  suivi d'une perforation  $\frac{2}{3}$ .

L'émetteur, le canal satellite et le récepteur composent la chaîne de diffusion. Nous avons utilisé une modulation de phase 4 états à cause des distorsions non-linéaires apportées par le tube amplificateur à ondes progressives du satellite.

Le décodage de canal et le décodage de source réalisent les fonctions inverses du codage de canal et du codage de source. Le décodeur convolutionnel utilise l'algorithme de Viterbi.

Le système étant décrit, nous allons pouvoir étudier les traitements possibles en cas d'erreurs de diffusion.

### 3. TRAITEMENT DU PAQUET ERRONE

Le train binaire vidéo MPEG-2 est composé de deux types de code, les codes à longueur fixe (**FLC:Fixed Length Codes**) et les codes à longueur variable (**VLC:Variable Length Codes**). L'étude statistique des trains binaires de plusieurs séquences d'images test montre que les VLC représentent environ les  $\frac{2}{3}$  des codes utilisés. L'apparition d'une erreur de transmission dans un de ces codes peut se traduire par deux effets. Le décodeur ne connaît la longueur du VLC qu'il traite

qu'après avoir trouvé sa valeur. Ainsi, s'il se produit une erreur, le décodeur perd le repère du début du code suivant et n'est alors plus capable de trouver aucune valeur de code. Pour limiter cet effet, la norme préconise de découper l'image comprimée en bandes de 16 pixels de hauteur appelées "slices". Chaque slice débute par un mot de 32 bits qui est unique dans le train binaire et qui est aligné sur une frontière d'octet. Ce `start_code` est un point de resynchronisation permettant de reprendre le décodage en cas d'erreur. Ce comportement reste le même quel que soit le VLC traité. Dans le cas d'un FLC, le décodeur connaît a priori la longueur du code qu'il traite. Une erreur provoque une valeur décodée fautive, mais elle n'empêche pas le décodage de la suite du train binaire.

Le code Reed-Solomon(188,204) permet de corriger 8 erreurs dans un paquet et d'en détecter 16. En simulant la chaîne de transmission, nous avons montré qu'un paquet erroné non corrigé par le codage de canal contient toujours moins de 16 octets erronés si le taux d'erreur paquet est inférieur à 1 sur 100. Ainsi, la couche système est à même d'indiquer au décodeur vidéo si un paquet de transport est défectueux ou non. Que faut-il faire de ce paquet ? Nous avons envisagé deux solutions :

- Pour faciliter le démultiplexage des paquets de transport, on peut considérer que tout paquet erroné est un paquet perdu. En sa présence, le décodeur doit sauter au `start_code` suivant et continuer le décodage, puis masquer au mieux les slices qui n'ont pas été décodés. L'inconvénient de cette méthode est que la quantité d'image contenue dans un paquet de 184 octets peut être grande. Or, plus la portion d'image à corriger est grande et plus le défaut dû à l'erreur risque d'être visible.
- Il existe une autre possibilité. On peut essayer de décoder quand même le paquet erroné en se fiant à un mécanisme de détection interne pour sauter au `start_code` suivant en cas d'erreur. Statistiquement, le nombre de slices à reconstituer devrait être plus faible dans ce cas. Par conséquent, la qualité de l'image pourrait être meilleure si la détection s'avère suffisamment fiable.

Nous avons implanté ces solutions dans deux décodeurs dont nous allons maintenant voir le fonctionnement.

### 4. DECODEUR NON-AUTONOME

La détection d'erreurs étant assurée par la couche système, ce décodeur doit seulement masquer au mieux les slices non décodés. La correction doit se faire par compensation de

mouvement avec une image de référence. Afin de faciliter ce masquage, le codeur vidéo doit être configuré en tenant compte des deux aspects suivants.

La structure de slices ; il doit y avoir un grand nombre de slices par largeur d'image pour minimiser la surface d'image à corriger en cas d'erreurs. La longueur d'un en-tête de slice est égale à 38 bits. Le meilleur compromis surcoût-efficacité est de 4 slices par largeur d'image. Ce surcoût est de 1,7 % avec un débit de 6 Mbit/s et de 2,9 % avec un débit de 3,5 Mbit/s par rapport à une image n'ayant qu'une slice par largeur.

Les Concealment Motion Vectors (CMV) ; il est impossible de corriger un slice contenant des macroblocs codés Intra car on n'a pas d'informations de mouvement relatives à une image de référence, ce qui est notamment le cas pour les images I. Or ces images servent elles-mêmes de références pour toutes les images du groupe et il faut les protéger en priorité. La norme préconise donc l'utilisation des CMV. On transmet un vecteur de mouvement en mode "avant-image" pour chaque macrobloc Intra contenu dans une image de type I ou P. Le vecteur transmis avec un macrobloc correspond au déplacement du macrobloc se situant juste en dessous de lui, ce qui permet d'assurer une meilleure protection de l'image en cas de paquet d'erreurs. La longueur moyenne pour coder ce vecteur est de 2 octets par macrobloc. L'utilisation des CMV implique donc un surcoût de 0,9 % à 6 Mbit/s et de 1,5 % à 3,5 Mbit/s.

La correction se fait en deux étapes : la localisation de la partie erronée et la compensation de mouvement du ou des slices manquant. Pour localiser la portion de slice à corriger, il faut mémoriser l'adresse du dernier macrobloc non-sauté décodé correctement ainsi que celle du premier macrobloc décodé après la resynchronisation. La correction proprement dite se fait selon le principe général suivant :

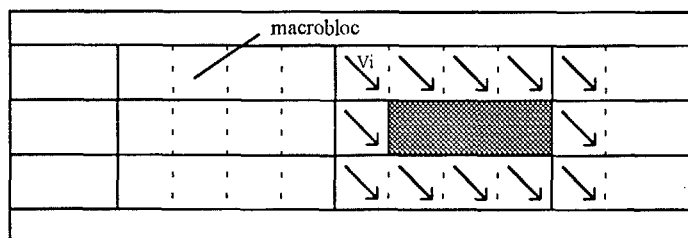


Figure 2 : masquage d'un slice

On compense les macroblocs manquant soit avec les CMV des macroblocs Intra s'il y en a, soit en moyennant les vecteurs environnants pour les autres types de macroblocs.

Le décodeur "intelligent" utilise les mêmes mécanismes de correction que le décodeur non-autonome, mais il incorpore plusieurs méthodes de détection d'erreurs.

## 5. DECODEUR "INTELLIGENT"

En cas d'erreur dans un VLC, le décodage donnera toujours un résultat, même s'il est faux. La seule conséquence visible sera la désynchronisation du décodeur vis à vis du train binaire. Nous devons donc implanter des tests permettant de détecter la perte de synchronisme dans le décodeur, qui soient suffisamment simples pour être réalisables en temps réel :

- détection des VLC inexistantes : Certaines valeurs possibles de VLC ne sont pas utilisées dans MPEG-2. La détection d'une de ces valeurs indique une erreur car ces codes ne doivent pas exister dans le train binaire.
- détection des valeurs de FLC interdites : les variables appartenant aux 6 couches syntaxiques composant la norme ne peuvent pas prendre toutes les valeurs possibles. Par exemple, la variable "quantiser\_scale" ne peut pas valoir 0.
- test de cohérence de l'adresse du macrobloc décodé : l'adresse du macrobloc courant doit être supérieure à l'adresse du dernier macrobloc décodé sans erreur. Il n'y a pas de macrobloc sauté dans une image I.
- test de cohérence sur la reconstruction du macrobloc décodé : Le test de reconstruction du bloc est une détection de valeur limite. Dans les tables de coefficients DCT, chaque VLC correspond à un nombre de zéros suivi d'une valeur non-nulle. Lorsque l'on reconstitue un bloc 8x8 avec ces valeurs successives de plage de zéros, l'index du tableau ne doit pas dépasser 64.
- test sur le bloc reconstruit : on calcule la valeur moyenne du bloc reconstruit, puis on la compare avec la valeur de chaque coefficient du bloc. Des écarts trop élevés entre les deux sont statistiquement très improbables.

Grâce à ces tests, nous avons pu détecter efficacement les paquets d'erreurs se trouvant dans le train binaire. Voyons maintenant comment nous allons tester l'influence des erreurs sur la qualité de l'image.

## 6. TESTS SUBJECTIFS

Il n'existe pas de critère objectif d'évaluation de la qualité d'une image, car on ne sait pas faire de modèle théorique de l'observateur humain moyen. De plus, la situation est compliquée par les techniques numériques non-linéaires et adaptatives utilisées dans la compression MPEG-2. Les erreurs



auront tendances à provoquer des défauts fortement décorréllés avec l'image tels que des déplacements de macroblochs sans lien avec les déplacements dans la séquence d'image. Aussi seul les tests subjectifs permettront d'évaluer correctement la qualité des images. Ces tests consistent dans la mise en oeuvre d'essais subjectifs présentés à un nombre significatif d'observateurs et effectués selon une méthode aussi précise que possible. On procède ensuite à une étude statistique pour calculer les moyennes, les variances, etc. Nous avons utilisé la méthode à double stimulus à échelle à 5 notes de qualité en suivant la recommandation 500 du CCIR [5] pour le déroulement des tests. On présente à l'observateur une séquence de référence et une séquence à tester. La position de la référence n'est pas connue et l'observateur note dans l'absolu les deux séquences. L'échelle de notation donne des repères de qualité d'image (excellente, bonne, moyenne...) et n'est pas expliquée à l'observateur. Cette méthode a été utilisée pour tester les deux types de décodeurs.

## 7. RESULTATS

On a, à taux d'erreurs identiques, environ deux fois plus de paquets erronés par seconde avec un débit de 6 Mbit/s qu'avec un débit de 3,5 Mbit/s. Nous nous sommes placés dans le pire des cas pour les tests avec le débit le plus élevé, la qualité de la voie à 3,5 Mbit/s étant supposée meilleure. Un générateur simulant une diffusion par satellite de bout en bout permet d'injecter les erreurs dans le train binaire. Nous avons fixé arbitrairement à 4 sur 5 le seuil de perception aux erreurs résiduelles, ce qui correspond à un défaut à peine perceptible.

Le tableau suivant nous donne les notes moyennes obtenues par le décodeur "intelligent" et par le décodeur non-autonome en utilisant trois séquences tests de l'UER.

décodeur	non-autonome		"intelligent"	
	1/4000	1/12000	1/4000	1/12000
flower	4,21	4,64	4,34	4,88
tennis	3,46	4,27	3,58	4,47
calendar	4,79	4,88	4,82	4,89
Total	4,15	4,60	4,25	4,75

Tableau 1 : récapitulatif des notes obtenues

On peut constater que la qualité de l'image est meilleure avec le décodeur "intelligent" qu'avec le décodeur non-autonome. Cela montre que les mécanismes de détection d'erreurs sont suffisamment exhaustifs pour que l'on puisse

envisager de décoder un paquet erroné. La simple exploitation du lien avec la couche système ne suffit donc pas pour améliorer la qualité de l'image décodée en cas d'erreurs. Par contre, ce lien apporte une plus grande robustesse au décodeur quand le taux d'erreur devient beaucoup plus élevé. Il faut obligatoirement adjoindre un mécanisme de "chien de garde" au décodeur "intelligent" pour qu'il ait la même résistance. Dans tous les cas, la limite de fonctionnement du décodeur vidéo correspond à un taux d'erreur de un paquet sur vingt.

## 8. CONCLUSION

Dans cet article, nous avons étudié l'influence des erreurs de diffusion sur un décodeur vidéo MPEG-2. Après avoir présenté le système de diffusion complet, nous avons vu deux stratégies possibles de masquage des erreurs. L'une nécessite un lien entre le décodeur vidéo et le système, alors que l'autre rend le décodeur complètement autonome. Grâce aux tests subjectifs, nous avons montré que le décodeur "intelligent" donne des images de meilleure qualité en cas d'erreurs que le décodeur dépendant du système. D'autre part, l'utilisation d'un décodeur vidéo autonome pourrait faciliter la conception du récepteur satellite. On peut aussi l'utiliser avec profit dans le cas d'une diffusion par voie hertzienne.

[1] Organisation Internationale de Normalisation  
ISO/IEC JTC1/SC29/WG11  
DIS 11172-2 audio

[2] Organisation Internationale de Normalisation  
ISO/IEC JTC1/SC29/WG11  
DIS 13818-2 vidéo

[3] Organisation Internationale de Normalisation  
ISO/IEC JTC1/SC29/WG11  
DIS 13818-2 systems

[4] Union Européenne de Radiotélévision  
specification of the « baseline modulation/channel coding system » for digital multi-program television by satellite.  
contribution from V4/MOD-B 7th revised version jan.94

[5] Méthode d'évaluation subjective de la qualité des images de télévision.  
Recommandation 500-4 du C.C.I.R Années 74-78-82-86-90  
Section 11-D. Volume 11-1. page 47.