

LE CODAGE D'IMAGES VU PAR UN INDUSTRIEL

Jean-François VIAL et Philippe TOURTIER

THOMSON-CSF / Laboratoires Électroniques de Rennes
Avenue de Belle Fontaine, 35510 Cesson-Sévigné, France

Tél. : +33 99.25.42.00 Fax : +33 99.25.43.34 e-mail : tsi@ler.thomson.fr

RÉSUMÉ

Après un rappel du contexte du codage d'images au niveau Européen et international, le point est fait sur la réduction de débit d'images numériques à THOMSON-CSF/LER : dans un premier temps sont présentées les études achevées, ayant débouché sur des normes internationales et des produits déjà disponibles. Puis sont détaillées les études menées actuellement, principalement dans le domaine de la TVHD. Une contrainte importante est le codage compatible multi-résolution ; le problème ainsi que les solutions à l'étude sont exposés.

ABSTRACT

After recalling the image coding context at the European as well as international level, digital image bit rate reduction in THOMSON-CSF/LER is discussed : in a first part completed studies, having resulted in international standards and already available products, are presented. Then the studies currently carried out, mainly in the field of HDTV, are detailed. An important constraint is multi-resolution compatible coding ; the problem and solutions under study are exposed.

1. INTRODUCTION

Le codage d'images et particulièrement de séquences d'images est un domaine de recherche et de développement important dans un groupe comme THOMSON. Les évolutions, dans ce domaine, du contexte outre-Atlantique (décision de la FCC) et en Europe (projets RACE et Eureka) ont accéléré les travaux. Conjointement des groupes de travail liés aux comités internationaux tels que le CCIR, le CCITT et l'ISO ont contribué à unifier ces développements et à passer du stade de la recherche à celui des spécifications détaillées de systèmes.

2. LE CONTEXTE

L'Europe a choisi il y a près d'une décennie le HD-MAC pour la diffusion de TVHD. Cette technique partiellement numérique a pour avantage la compatibilité avec la TV diffusée en D2-MAC : un récepteur TV est capable d'extraire le signal correspondant à sa propre résolution dans une émission en TVHD. Seule une telle approche autorise une introduction progressive du service TVHD, en permettant aux téléspectateurs de recevoir sur leurs postes TV actuels des programmes diffusés en TVHD et d'accéder à la réception en haute définition lorsqu'ils le désirent.

Aux États-Unis, la FCC (Federal Communications Commission) a récemment opté pour le tout numérique avec un

système non encore totalement défini mais qui sera basé sur la future norme internationale ISO/MPEG-2.

La médiatisation de cette décision a fait que de nombreuses voix se sont élevées en Europe pour dénoncer les pertes d'argent et d'avance technologique causées par le choix du HD-MAC. Or il n'en est rien : si l'avenir appartient incontestablement aux systèmes numériques, cela était prévu dès le début (bien qu'originellement à plus long terme) et le HD-MAC était surtout destiné à accompagner la migration du parc de récepteurs vers la TVHD. Les études de systèmes de diffusion de TV et TVHD numérique se sont donc poursuivies en Europe, sans aucun retard sur les USA ni le Japon. De plus, dans le domaine de la contribution (transmissions point à point studios TV) des systèmes numériques ont été développés en Europe et sont déjà utilisés, aussi bien en TV qu'en TVHD. Enfin une grande majorité des développements réalisés pour le HD-MAC sont en numérique ou sont indépendants du standard de diffusion comme les caméras, les équipements de studio et les tubes cathodiques, dont le format 16/9 et la résolution sont communs aux différents systèmes tant analogiques que numériques.

3. LE CODAGE DE QUALITÉ CONTRIBUTION

Le projet Européen RACE 1018 HIVITS (HIGH quality Videotelephone and Television Systems) a débuté en 1988 et s'est achevé en 1992. Dirigé par THOMSON-CSF/LER, HIVITS a été



durant cette période le principal projet Européen d'études et de développement dans le domaine du codage numérique, et le seul pour les applications de contribution ; le champ d'application allait du vidéotéléphone à 64 Kbit/s à la TVHD en qualité contribution à 140 Mbit/s.

3.1. La télévision

Au cours du projet HIVITS, le CCIR publiait sa recommandation 723 pour le codage de télévision numérique à la qualité contribution à 34 Mbit/s (45 Mbit/s aux États-Unis) [1]. L'algorithme de compression est du type hybride : transformation cosinus discrète (DCT) et prédiction temporelle compensée en mouvement. Contrairement au désormais célèbre algorithme MPEG-2 le format des images est fixe et correspond à la recommandation CCIR 601 (720x288 pixels utiles par trames entrelacées à 50 Hz), les trames paires et impaires sont traitées séparément, et les codes à longueur variable (VLC) sont différents de ceux de la recommandation CCITT H.261 pour le vidéotéléphone. THOMSON-CSF/LER, qui avait pris une part très active dans l'élaboration de cette norme, développait et réalisait dans le cadre du projet HIVITS un codec répondant à cette spécification.

Une gamme de produits a été dérivée de ce prototype, le dernier en date étant reconfigurable pour fonctionner à 34 Mbit/s, à 17 Mbit/s (permettant deux programmes TV dans un canal à 34 Mbit/s, par exemple pour la distribution primaire) et enfin jusqu'à 8 Mbit/s (par exemple pour des applications du type "Satellite News Gathering", c'est-à-dire de reportage léger). Ces produits sont déjà utilisés par divers opérateurs "télécom" et diffuseurs TV en Europe et de par le monde.

3.2. La télévision haute définition

L'architecture de ces codecs TV et le jeu de circuits intégrés ASICs développés pour cela ont été conçus suffisamment flexibles pour être utilisés dans d'autres applications. En effet, toujours dans le cadre du projet HIVITS, un codec TVHD à la qualité contribution à 140 Mbit/s a aussi été réalisé. Les images TVHD ayant 2 fois plus de lignes que les images TV et 2,67 fois plus de pixels par ligne en raison du rapport 16/9, le principe retenu pour traiter une fréquence pixel environ cinq fois supérieure à celle de la TV consiste à diviser les images TVHD en bandes horizontales de huit lignes, chaque bande étant traitée par un parmi cinq processeurs travaillant en parallèle. Ces processeurs sont en fait des codecs TV modifiés pour communiquer entre eux car la compensation de mouvement nécessite des échanges de données entre les processeurs traitant des bandes adjacentes.

La syntaxe du train binaire est une simple extension de celle

de la recommandation CCIR 723 : plus de pixels, plus de lignes, plus de bits. Cette approche ne permet toutefois pas la compatibilité, car un décodeur TV ne peut pas extraire d'un tel train binaire de quoi obtenir une image de résolution inférieure, comme expliqué plus loin. La compatibilité multi-résolution n'est de toute manière pas nécessaire en contribution car contrairement à la distribution seuls un codeur et un décodeur communiquent, et cela à un seul format d'image et un seul débit (l'idée est la même en vidéotéléphonie avec les formats CIF et QCIF).

Le produit correspondant à ce prototype a lui aussi été mis sur le marché par THOMSON-CSF/LER. De la même façon que le codec TV, ce codec TVHD peut fonctionner à 70 Mbit/s et à 34 Mbit/s pour des applications ne nécessitant pas la haute qualité contribution.

4. LE CODAGE DE QUALITÉ DISTRIBUTION

Le codage pour diffusion de TV ou TVHD numérique par câble, satellite ou voie hertzienne terrestre, a aussi été étudiée dans le projet HIVITS. L'approche retenue était que même si un système de TVHD numérique ne sera utilisé qu'une fois le parc de récepteurs suffisamment développé, la compatibilité TV / TVHD sera toujours nécessaire (comme illustré plus loin). Il est alors apparu que pour cela les algorithmes de codage TV et TVHD doivent être définis ensemble, ce qui a été fait dans ce projet (les études réalisées dans HIVITS sur la TV seront donc décrites dans le chapitre sur la TVHD).

4.1. La télévision

Concernant la distribution de TV numérique seule, le groupe de travail ISO/MPEG-2 vient de figer la spécification qui deviendra prochainement une norme internationale. L'extension vers la TVHD - compatible ou non avec la TV - est actuellement à l'étude ; mais comme mentionné plus haut, le fait d'avoir commencé par la TV rend la compatibilité plus difficile. THOMSON-CSF/LER participe activement à ces travaux, et a été parmi les premiers à y proposer des solutions pour le codage compatible TV / TVHD. Le projet Européen Eureka 625 VADIS (Video-Audio Digital Interactive System), auquel participe aussi THOMSON-CSF/LER, coordonne les activités d'étude et de développement liées à MPEG-2 en Europe. De nombreux résultats de HIVITS ont été ré-utilisés dans VADIS, qui a débuté en 1991.

4.2. La télévision haute définition

4.2.1. La compatibilité multi-résolution

Une fonctionnalité importante en codage numérique d'images animées est la compatibilité multi-résolution. Le principe est de pouvoir ne décoder qu'une partie d'un flux binaire, les images

obtenues étant bien-sûr d'une qualité et/ou d'une résolution spatiale et/ou temporelle inférieure à celles des images obtenues en décodant la totalité du flux binaire.

Les avantages résident dans l'existence de décodeurs de plus faible coût que d'autres car nécessitant moins de mémoire et une résolution d'écran inférieure (par exemple des récepteurs TV en cas de diffusion TVHD, comme c'est le cas avec le système D2-MAC / HD-MAC), ainsi que dans la possibilité de connexion à un réseau de moindre coût ou de réception hertzienne plus robuste aux erreurs car correspondant à un débit plus faible. Un autre avantage du codage multi-résolution est d'offrir une dégradation progressive des images lorsque le taux d'erreurs dans le flux binaire augmente, ce que ne permet pas un algorithme de codage numérique classique.

4.2.2. La transformation sous-bande

Le codage multi-résolution repose sur la transformation sous-bande des images, permettant d'obtenir des images de résolution spatiale inférieure en ne décodant que des coefficients de basse fréquence. En effet le principe de la transformation sous-bande consiste, dans l'exemple mono-dimensionnel, à filtrer le signal d'une part avec un filtre passe-bas et d'autre part avec un filtre passe-haut. Les deux signaux résultant, ayant chacun une bande passante moitié de celle du signal original, peuvent être sous-échantillonnés par un facteur deux sans perte d'information. Le nombre total d'échantillons est donc inchangé, mais les propriétés statistiques et psycho-visuelles des signaux sous-bandes les rendent plus facile à coder efficacement. De plus chacun de ces signaux peut éventuellement à son tour être décomposé en deux sous-bandes, et le processus peut se répéter aussi longtemps que nécessaire pour le codage.

La première décomposition bidimensionnelle retenue comme optimale par THOMSON-CSF/LER comprenait 16 sous-bandes (figure 1) et utilisait des filtres QMF (Quadrature Mirror Filters) à 16 coefficients appliqués hiérarchiquement comme expliqué précédemment. Finalement une décomposition parallèle en 64 sous-bandes (figure 2) a été préférée pour des raisons de réalisation matérielle et à cause des bonnes propriétés pour le codage d'images des filtres PRMF (Perfect Reconstruction Modulated Filters, [2][3]) à 15 coefficients.

En ce qui concerne la compatibilité multi-résolution, chacun de ces deux types de décomposition sous-bande permet de reconstruire des images ayant un quart de la résolution originale, correspondant en fait au quart basse fréquence. Des résolutions inférieures sont même possibles, ce qui peut être intéressant pour certaines applications : par exemple en diffusion de TVHD, outre des récepteurs TVHD "de salon" (reliés à une antenne extérieure ou au câble), on peut utiliser des récepteurs TV portables (avec une antenne intégrée, mais reliés au secteur), et des récepteurs TV portatifs (sur batterie, avec un petit écran à

cristaux liquides).

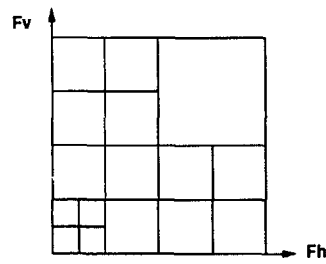


Fig.1 : Découpe hiérarchique

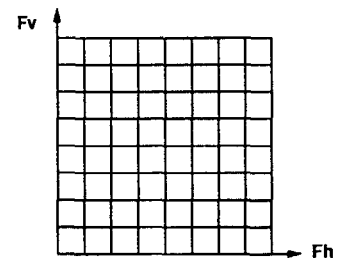


Fig.2 : Découpe parallèle

4.2.3. L'organisation du train binaire

Un décodeur TV doit pouvoir extraire les données relatives à sa résolution sans avoir à décoder la totalité d'un train binaire TVHD. Comme la réduction d'entropie repose sur l'utilisation de codes à longueur variable (VLC : Variable Length Codes), il faut que les données relatives aux basses fréquences soient séparées de celles relatives aux hautes fréquences par des codes facilement repérables dans le train binaire ("mots de synchronisation"). Le principe retenu est de coder ensemble les coefficients d'une même sous-bande et à ordonner les sous-bandes par fréquences croissantes en commençant par un mot de synchronisation, et cela pour chaque groupe de 8 ou 16 lignes horizontales de l'image.

4.2.4. La prédiction compensée en mouvement

Les grandes lignes d'un codec mono-résolution sont illustrées figure 3. La principale différence avec un schéma à base de transformation par blocs comme la DCT (par exemple H.261, MPEG, CCIR 723, etc.) est qu'il est préférable de procéder à la décomposition en sous-bandes avant la soustraction de la prédiction, car après cette dernière le signal comprend des blocs en mode *inter* (avec soustraction) ou *intra* (sans) ; or le filtrage sous-bande n'étant pas limité aux frontières de blocs, cela créerait des hautes fréquences artificielles et difficiles à coder.

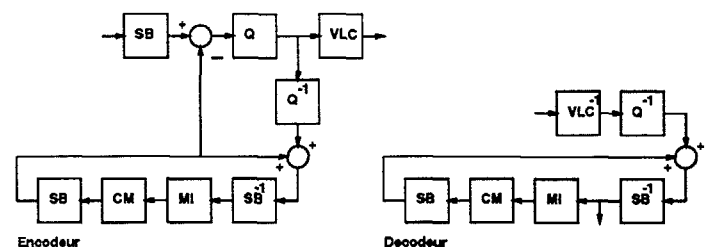


Fig.3 : Codec "mono-résolution"

Toutefois le schéma de la figure 3 ne permet pas la compatibilité entre des résolutions différentes. En effet la compensation de mouvement et la transformation en sous-bandes ne sont pas commutatives, ainsi la prédiction ajoutée au signal



dans un décodeur TV, obtenue après compensation en mouvement de l'image TV précédemment décodée, ne sera pas égale à la partie basse fréquence de la prédiction soustraite dans un codeur TVHD, obtenue, elle, après compensation en mouvement de l'image de pleine définition.

Une solution est qu'un codec TV soit inclus dans le codec TVHD, de façon à ce que le codage de la partie basse fréquence des images de pleine définition ne fasse appel qu'à des données disponibles dans un décodeur TV (figure 4). Cette solution présente le désavantage de rendre plus complexe le décodeur TVHD (le cas du codeur est moins important en distribution). Cela est d'autant plus vrai que plusieurs niveaux de résolution sont utilisés.

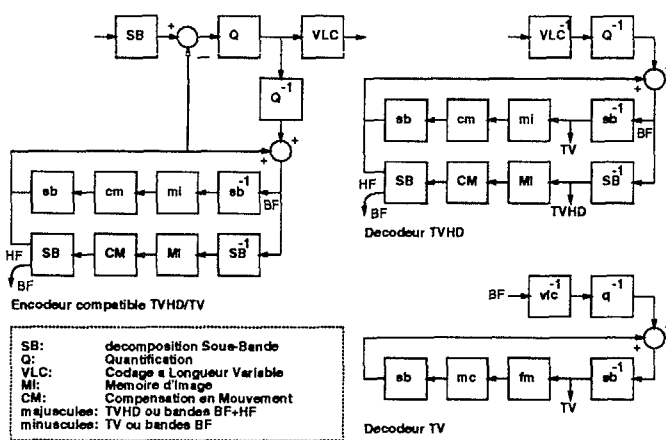


Fig.4 : Codec à deux boucles de prédiction

Une solution plus simple à mettre en oeuvre consiste à réaliser la compensation en mouvement dans le domaine fréquentiel (figure 5). Dans le cas d'une découpe parallèle 8×8 , chaque sous-bande est sous-échantillonnée 8 fois dans chaque direction. Donc la même qualité de prédiction qu'avec une compensation de mouvement au demi pixel dans l'image, ce qui est couramment le cas, nécessite une compensation de mouvement au seizième de pixel dans les sous-bandes. Pour cela une interpolation bilinéaire n'est pas suffisante, d'autant plus que du repliement de spectre ("aliasing") est présent dans les sous-bandes car les filtres ne sont pas des passe-bandes parfaits. Une analyse mathématique du problème a permis de définir les filtres d'interpolation idéaux, dérivés des filtres de décomposition en sous-bandes [4]. Avec cette technique, l'interpolation dans une sous-bande doit faire appel aux sous-bandes adjacentes : la compatibilité multi-résolution apporte comme contrainte que les sous-bandes à la limite supérieure des basses fréquences ne peuvent pas utiliser les sous-bandes voisines de fréquence supérieure car celles-ci ne sont pas accessibles à un décodeur TV. La faible perte en qualité d'image par rapport à un schéma

mono-résolution est la même qu'avec le schéma à plusieurs boucles.

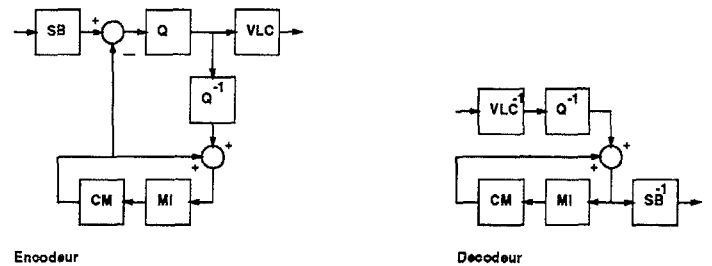


Fig.5 : Codec à compensation en mouvement "dans les bandes"

4.2.5. La régulation de débit

Pour de nombreuses applications, il est important de pouvoir réguler le débit correspondant à la partie TV indépendamment du débit total (Une régulation de débit est nécessaire à cause des VLC). Or avec les schémas présentés ici les sous-bandes basses fréquences ne sont codées qu'une seule fois et servent pour la TV et la TVHD, donc elles doivent être à un niveau de qualité suffisant pour la TVHD. Cela impose de n'avoir qu'une seule régulation, donc une seule quantification pour toutes les sous-bandes, mis à part des facteurs de pondération psycho-visuels. Pour le type d'image et les débits utilisés couramment en télévision, la partie TV représente en moyenne de 50 à 75 % du débit total. Afin de réduire ce rapport, il faut quantifier les coefficients basses fréquences en deux étapes : une première fois selon le débit fixé pour la TV, puis une deuxième fois avec les coefficients hautes fréquences pour la TVHD. Ce sujet est en cours d'étude.

5. CONCLUSION

Cet article a présenté les travaux (études et produits) réalisés par THOMSON-CSF/LER dans le domaine du codage d'images. Ceux-ci se poursuivent actuellement vers des débits de plus en plus faibles à l'aide de nouvelles techniques comme le codage par segmentation ou à base de modèles.

6. REFERENCES

- [1] "Transmission of component coded digital television signals for contribution quality applications at the third hierarchical level of CCITT Recommendation G702", CCIR Recommendation 723, 1990, volume XII.
- [2] J. Mau, "Perfect Reconstruction Modulated Filter banks", IEEE Conference on ASSP, 1992.
- [3] J. Mau, "PRMF : fast algorithms and attractive new properties", IEEE Conference on ASSP, 1993.
- [4] P. Tourtier, M. Pécot and J.-F. Vial, "Motion compensated subband coding schemes for compatible HDTV coding", Signal Processing : Image Communication, vol. 4, n°4-5, August 1992.