



# Technique d'allocation de ressources binaires pour codeur vidéo en sous-bandes adapté à l'ATM

Jean Menez\*, Michèle Rosso\*\*, Paolo Scotton\*\*

\* Université de Nice - Sophia Antipolis  
URA - CNRS 1376  
41, Bd Napoléon III  
F-06041 Nice Cedex

\*\* Compagnie IBM France  
Centre d'Etudes et Recherches  
Le plan du Bois  
F-06610 La Gaude  
ABSTRACT

## RÉSUMÉ

Dans cette communication nous présentons une méthode de compression vidéo de faible complexité adaptée à la transmission sur des réseaux ayant pour mode de transfert l'ATM. Une première partie est consacrée à la description des techniques de compression utilisées pour réaliser ce codeur qui permet d'obtenir une très bonne qualité d'image pour une complexité d'environ 17 Mips et un taux de compression d'environ 10. Dans une deuxième partie, après un bref rappel de la technique ATM et de ses fondements, nous explicitons le caractère très adapté du codeur proposé pour ce type de réseaux et nous présentons un protocole de transport permettant de pallier les difficultés de mise en œuvre.

In this paper a low complexity video coder fully compatible for transmission over ATM based networks is presented. The first section shows how this coder is designed to compress an incoming video signal by a factor of about 10 with a very good visual quality, such results being obtainable with 17 Mips calculation power. The second section begins with a short description of ATM. It is then shown how the proposed coder is well suited for ATM networks. Finally, a transport protocol that solves the ATM video transmission problems is briefly presented.

## 1 Introduction

Avec le mariage de l'ordinateur, du son et de l'image on assiste à l'émergence d'un nouveau type de communication: la communication *multimédia*. Outre le développement de stations multimédia, elle implique la mise en œuvre de nouveaux réseaux de communication permettant de véhiculer sur des mêmes liens des connexions de nature et d'exigences très différentes.

Le mode de transfert ATM (*Asynchronous Transfer Mode*) sera sans doute, dans les années à venir, la technique de transport pour ces réseaux à haut débit à l'échelle mondiale.

Malgré les largeurs de bande très élevées que ces réseaux offriront, il restera quand même nécessaire de comprimer l'information émanant de sources, telles que la vidéo, pour des problèmes de rentabilité.

Dans cet article nous présentons une méthode de compression vidéo bien adaptée à l'ATM.

## 2 Codeur vidéo avec quantification multirésolution

Le synoptique du codeur est présenté à la figure 2.

### 2.1 Réduction de la redondance temporelle

Nous avons choisi d'utiliser, pour réduire la redondance temporelle, un codage MICD (*Modulation par Impulsions et Codage*). Cette technique présente l'avantage d'être simple et permet d'obtenir de bonnes performances [Raim91]. En contrepartie, l'inconvénient du codage MICD réside dans sa faible résistance aux erreurs de transmission. Nous présentons dans la section 3 un protocole de transport permettant de pallier cet inconvénient.

### 2.2 Réduction de la redondance spatiale

Dans le codeur proposé, la réduction de la redondance spatiale est obtenue par décomposition/recomposition fréquentielle en sous-bandes. Le signal est tout d'abord décomposé en sous-bandes à l'aide d'un banc de filtres séparables. Le signal de chaque sous-bande est ensuite décimé à sa nouvelle fréquence de Nyquist avant codage et transmission. Au récepteur, on effectue les opérations "inverses". Dans le cas monodimensionnel, D. Esteban et C. Galand [Este77] ont montré que pour un signal numérisé  $\{x\}$  si on applique une décomposition/recomposition en deux sous-bandes



(voir figure 1), pour obtenir une reconstitution parfaite les fonctions de transfert des différents filtres doivent satisfaire les conditions suivantes:

$$H_0(-Z)G_0(Z) + H_1(-Z)G_1(Z) = 0 \quad (1)$$

$$H_0(Z)G_0(Z) + H_1(Z)G_1(Z) = 2Z^{-k} \quad (2)$$

où  $k \in \mathcal{N}$ . La condition (1) est une condition d'anti-repliement (ou *anti-aliasing*), la condition (2) est la condition de gain. Plusieurs familles de filtres remplissant de manière approchée ou exacte à ces conditions ont été proposées: QMF (*Quadrature Mirror Filters*) [Croi76], CQF (*Conjugate Quadrature Filters*) [Smit86], filtres polynômiaux [Vett86] ... Plus récemment, la théorie des ondelettes [Mey92] a permis d'aboutir à des conditions similaires, s'agissant de la transformation en ondelette orthogonale ou bi-orthogonale.

Pour limiter la charge de calcul extrêmement importante que requiert le codage en sous-bandes d'un signal vidéo, notre choix s'est porté sur un filtre ayant pour réponse impulsionnelle l'ondelette de Haar: c'est un filtre QMF.

### 2.3 Codage des sous-bandes

L'originalité du codeur que nous proposons réside dans l'utilisation d'une technique de quantification "multirésolution". On a constaté que les images constituant l'erreur de prédiction décomposée en sous-bandes sont des images de "contours". Ceci peut s'expliquer pour les raisons suivantes:

- L'erreur de prédiction est issue d'un filtrage passe-haut temporel (MICD).
- Toutes les sous-bandes (sauf la bande de base) ont subi au moins un filtrage passe-haut.

La technique proposée consiste à décomposer les images de chaque sous-bande en blocs de dimension carrée. Un algorithme de détection d'activité permet de distinguer les blocs "actifs" des blocs de "fond". Le principe de cet algorithme, s'appliquant à chaque bloc, est le suivant

- Si la valeur d'un *pixel* est supérieure à un certain seuil  $T_1$ , alors le *pixel* est considéré comme actif.
- Si un bloc contient au moins  $T_2$  *pixels* actifs alors le bloc sera décrété actif.

Les blocs actifs sont quantifiés avec un nombre d'éléments binaires plus important que les blocs de fond. Ces différents blocs présentent une densité de probabilité proche d'une densité Laplacienne [Scot93b].

Afin de réduire la complexité de l'algorithme de détection d'activité, on exploite la corrélation

énergétique existant entre les sous-bandes. Comme le codeur fait appel à une décomposition arborescente en 7 sous-bandes, on effectue la détection sur les sous-bandes 1 à 4 et les résultats sont ensuite reportés aux sous-bandes "haute fréquence".

En réalisant en tête la décomposition en sous-bandes du signal vidéo avant que de coder chaque sous-bande par un codeur prédictif du type MICD, on obtient une réduction sensible de la complexité de mise en œuvre de l'algorithme sans altérer la qualité des séquences reconstruites. Ceci s'explique par le fait qu'il n'y a plus lieu de réaliser la recombinaison du signal au niveau de l'émetteur.

Dans le cadre de notre réalisation, nous n'avons pas appliqué de codage entropique sur les données issues du quantificateur Laplacien. Il est donc possible d'augmenter ultérieurement le taux de compression du codeur en utilisant une technique de codage sans perte d'information, telle que l'algorithme du Contexte [Riss83] en combinaison avec un codage arithmétique [Lang84].

Nous avons testé ce codeur avec des séquences de type "Head & Shoulders" et de type "Full motion". Les résultats obtenus sont consignés dans le tableau ci-dessous:

Séquence	Nombre de bits/pixel moyen	Peak SNR (dB)
Miss America	0.83	37.42
Table Tennis	0.91	32.15

Tableau 1: Résultats du codeur proposé

Une image originelle extraite de la séquence "Table Tennis" et son homologue reconstitué après codage/décodage sont présentés sur la figure 4.

## 3 Interfaçage avec un réseau de type ATM

### 3.1 La technique de transfert asynchrone

Le choix du mode de transfert ATM pour les futurs réseaux intégrés large-bande est essentiellement motivé par la nécessité de faire cohabiter différents types de trafics. En particulier, cette technique tend à réunir les avantages de la commutation de paquets et de la commutation de circuits.

Son principe revient à transformer le trafic entrant, quel que soit le type de source l'ayant engendré, en cellules de 53 octets: 48 octets d'information et 5 octets d'en-tête pour le routage. Les cellules sont ensuite routées à haute vitesse à travers le réseau: les fonctions de contrôle d'erreurs et de contrôle de flux sont assurées uniquement dans les nœuds terminaux.

Afin d'optimiser le fonctionnement de ce type de réseau, on essaye de tirer profit du multiplexage statistique: on alloue dynamiquement la bande-passante pour chacune des différentes sources qui ont à se partager une ressource globale fixe. Il est clair que cette technique n'a lieu d'être que si les sources multiplexées sur un même lien présentent un trafic à débit variable.

La mise en œuvre du multiplexage statistique est étroitement liée à la connaissance des sources. En effet, si on multiplexe  $N$  sources de débit maximal  $R_i$ ,  $i = 1..N$ , pour obtenir un *gain statistique* il est indispensable que  $\sum_{i=1}^N R_i \geq C$  où  $C$  est la capacité de la ligne de transmission. En fait, pour chaque source, on a la possibilité de réserver une bande comprise entre son débit moyen et son débit maximal. Dans le premier cas on aboutit à un multiplexage statistique "poussé" alors que dans le deuxième cas on obtient un multiplexage déterministe. Il est évident que pour se placer dans une situation de multiplexage statistique tout en garantissant une certaine *qualité de service* il est indispensable de disposer d'un bon modèle de la source. Toutefois, même si on ne dispose pas d'un modèle très précis, il est malgré tout possible de garantir une certaine qualité de service grâce à l'utilisation de *protocoles de transport* adéquats [Scot93b]. On peut, par exemple, mettre œuvre des techniques de contrôle de flux adaptatives permettant d'estimer et de poursuivre les caractéristiques de trafic d'une source.

D'autre part, par des mécanismes de priorité, on peut éliminer certaines cellules en cas de congestion. Ainsi, on peut distinguer des cellules "prioritaires" qui doivent parvenir au récepteur et des cellules moins prioritaires qui peuvent être perdues en cas de congestion.

### 3.2 Codeur vidéo et réseau ATM

Le codeur proposé est bien adapté à la transmission sur réseau ATM car il est à débit variable et hiérarchisé.

Malgré le fait que l'on utilise une allocation de ressources binaires fixe, la quantité d'éléments binaires à transmettre pour chaque image différence est variable. En effet, si l'activité de la séquence est importante, le nombre de blocs actifs augmente considérablement, entraînant une augmentation du débit binaire. En revanche, dans le cas de séquences à faible activité le nombre de blocs actifs est faible. La figure 3 montre un exemple des variations du débit du codeur dans le d'une séquence comportant différents types de scènes.

Le fait de distinguer les blocs actifs des blocs de fond conduit à un codage hiérarchisé. D'après des résultats expérimentaux, on a pu constater que le fait de ne pas transmettre les blocs de fond permet d'obtenir une qualité acceptable tout en réduisant considérablement

le débit binaire. On peut donc définir un mode de fonctionnement "dégradé" qui est activé en cas de congestion du réseau. C'est là une des possibilités qui sont offertes par le protocole de transport que nous proposons.

De plus, grâce au mécanisme de priorité que nous avons évoqué plus haut, les cellules constituant les blocs de fond peuvent être transmises comme non-prioritaires. Ainsi, en cas de congestion, certaines d'entre elles peuvent être perdues sans que la qualité soit véritablement dégradée.

### 3.3 Protocole de transport

Dans les réseaux haute vitesse, pour pouvoir atteindre des bandes-passantes importantes, il est indispensable d'alléger autant que possible les protocoles de transport et de les rendre compatibles avec les contraintes imposées par le réseau. C'est pour cela que les nouveaux protocoles sont de plus en plus simples et basés sur une "*philosophie optimiste*". De plus, pour s'affranchir de l'augmentation du produit bande-passante - temps de propagation, ils ont tendance à limiter l'utilisation des techniques par acquittement. D'autre part, dans les modèles de référence des réseaux on peut voir qu'ils sont de plus en plus répartis entre les différentes couches.

Dans le cas d'une application vidéo, nous proposons un protocole de transport au niveau codeur - décodeur (les parties contrôle de flux et d'erreurs sont implémentés dans d'autres couches) qui permet de gérer les problèmes de mise à jour des boucles de prédiction et de passage en mode dégradé.

La boucle de prédiction du décodeur doit être mise à jour en cas d'erreurs de transmission non corrigibles. Etant donné que c'est le décodeur qui peut se rendre compte d'une telle situation, c'est ce dernier qui demande au codeur la remise à jour. Comme le temps de propagation a tendance, dans les réseaux haute vitesse, à être important vis à vis du débit des sources, le décodeur devra effectuer cette demande en "avance".

Le choix de passage en mode dégradé peut être demandé par le gestionnaire de réseau ou par le décodeur. Bien entendu ce mode est maintenu tant que le composant qui l'a décrété ne décide pas le passage en mode normal.

Une description complète de ce protocole est donnée dans [Scot93b].

## 4 Conclusion

L'effervescence que l'on connaît aujourd'hui dans le monde des télécommunications pour les réseaux haute vitesse de type ATM permettra de répondre dans un



futur proche aux nouveaux besoins de communication multimédia.

Dans cet objectif, les techniques de compression vidéo, toujours nécessaires malgré les fortes capacités de ces réseaux pour des raisons de coût, seront d'autant plus appréciées qu'elles présenteront un mode de fonctionnement bien adapté à ces réseaux. Dans cet article nous avons présenté un codeur vidéo qui en plus de ses performances qualité/débit/complexité, répond parfaitement aux notions de multiplexage statistique et de priorité des réseaux ATM par un débit variable et un codage hiérarchisé.

## Références

- [Croi76] A. Croisier, D. Esteban and C. Galand, *Perfect channel splitting by use of interpolation/decimation tree decomposition technique*, Int. Conf. on Information Sciences / Systems, Patras, August 1976
- [Este77] D. Esteban and C. Galand, *The application of quadrature mirror filters to split band voice coding schemes*, Proc. of the 1977 ICASSP, Hartford, Conn., pp 191-195, May 1977
- [Lang84] G.G. Langdon, *An introduction to arithmetic coding*, IBM Journal of Research and Development, vol.28, n.2, March 1984
- [Meye92] Y. Meyer, *Les ondelettes: algorithmes et applications*, Collection Acquis Avancés de l'Informatique, Armand Collin 1992
- [Raim91] C. Raimondo *et alii*, *Codeurs vidéo de faible complexité à base de codage en sous-bandes, de compensation de mouvement hiérarchisée et de quantification vectorielle*, Trezième Colloque du GRETSI, septembre 1983
- [Riss83] J. Rissanen, *A universal data compression algorithm*, IEEE Transaction on Information Theory, vol.IT-29, n.5, Sept. 1983
- [Scot93a] P. Scotton, J. Menez, *A low complexity sub-band video coder for ATM*, to appear in Signal Processing, Image Communication
- [Scot93b] P. Scotton, *Compression et transmission de signaux vidéo sur des réseaux haut débit avec contrôle de congestion du flux de données*, Thèse de Doctorat qui sera présentée à l'Université de Nice - Sophia Antipolis en octobre 1993
- [Smit86] M.J.T. Smith, T.P. Barnwell, *Exact reconstruction techniques for tree-structured sub-band coders*, IEEE Transactions on Acoustics, Speech and Signal Processing, vol. ASSP-34, n. 3, June 1986
- [Vett86] M. Vetterli, *Filter banks allowing perfect reconstruction* Signal Processing, vol. 10, pp. 219-244, 1986

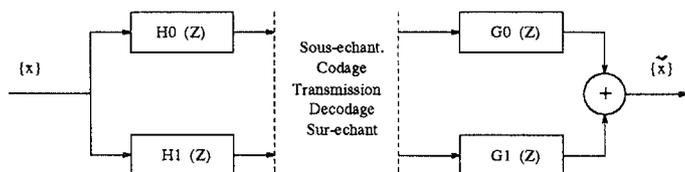


Figure 1: Principe du codage en sous-bandes

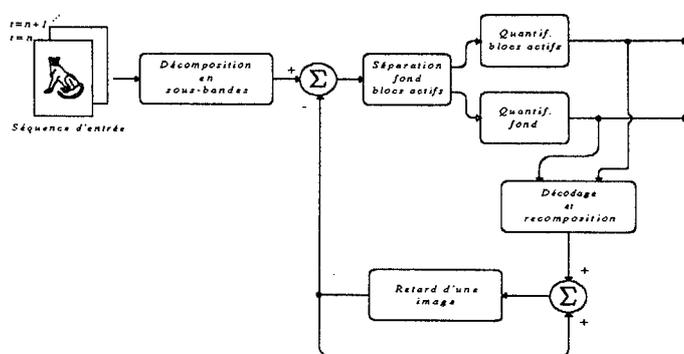


Figure 2: Synoptique du codeur hybride

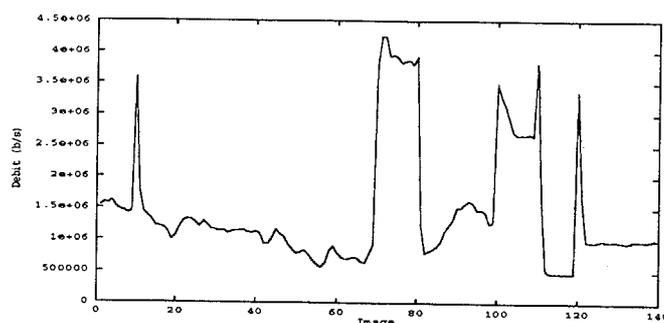


Figure 3: Exemple de débit d'émission

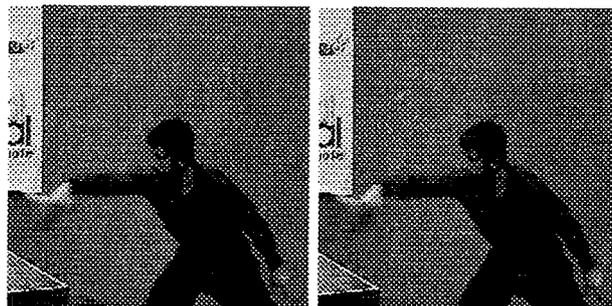


Figure 4: Images d'origine et reconstruite