

**SYSTEME DE REDUCTION DU DEBIT D'INFORMATIONS
ASSOCIE A LA TRANSMISSION D'IMAGES A HAUTE DEFINITION**

M. HAGHIRI, P. GUICHARD

LEP : Laboratoires d'Electronique et de Physique appliquée
3, avenue Descartes - 94451 Limeil-Brévannes Cédex (France)
*A Member of the Philips Research Organization

RESUME

Cet article traite du problème de la transmission d'informations numériques associées au signal de télévision à haute définition lorsqu'il est transmis dans un canal à bande réduite suivant la norme MAC. Ces informations d'assistance numérique (AN) sont composées pour l'essentiel de vecteurs mouvement qui permettent de reconstruire dans le décodeur les trames non transmises à l'aide d'une interpolation temporelle compensée en mouvement. La capacité du canal numérique disponible dans le retour trame du MAC étant insuffisante, il est nécessaire d'appliquer un algorithme de réduction de débit. Nous présentons une méthode qui se base sur deux postulats :

- le nombre de vecteurs mouvement différents lorsque la mesure est faite entre 2 trames espacées de 40 ms est faible.
- les valeurs de ces vecteurs présentent une importante corrélation dans la direction temporelle.

Ces hypothèses ont été validées par simulations et ont conduit à la définition d'un système de réduction de débit qui limite la taille de la source à coder. Dans le même esprit, nous avons déterminé un nouvel algorithme d'estimation de mouvement qui permet d'obtenir le débit requis sans introduire de nouvelle distorsion sur l'image.

SUMMARY

This paper deals with the problem of transmission of digital information required to enable a high definition television signal to be conveyed via narrow bandwidth channels into a MAC compatible form. These Digital Assistance informations (DA) mainly consist of motion vectors used to achieve a motion compensated interpolation of the non transmitted fields in the decoder. The digital channel capacity available in a MAC/ Packet standard is limited and therefore a bit rate reduction technique is necessary. We present a technique that takes benefit of two main postulates :

- the number of different motion vectors between two fields spaced of 40 ms is not large,
- these vectors are strongly correlated in the temporal direction.

These assumptions have been checked by simulations and have led to a bit rate reduction scheme that reduces the amount of information to be transmitted. In the same way we have developed a new motion estimation algorithm that fulfil the bit rate requirement while keeping the same picture quality.

INTRODUCTION

Dans le cadre du projet Eurêka EU95 sur la télévision à haute définition (HDTV), tous les problèmes liés à la production, à la transmission, aux moyens de restitution et de stockage de ces images sont abordés. Pour permettre une introduction progressive de la HDTV, un élément essentiel est la compatibilité avec les standards de transmission existants. Le support de transmission retenu est un canal au standard MAC qui permet la diffusion d'images de définition normale (625 lignes de 720 points). La transmission d'images haute définition (HD) de 1250 lignes de 1440 points, à travers ce canal, nécessite donc une compression de la bande du signal par un facteur 4.

Le principe de compression de bande retenu (HDMAC) consiste à adapter la structure d'échantillonnage spatio-temporelle à la vitesse mesurée sur l'image. Ainsi, dans le cas de mouvements de faible amplitude, la transmission d'une image se fait sur 4 trames. On obtient alors une définition spatiale 4 fois supérieure à celle obtenue dans le cas de mouvements importants où la transmission d'une image s'effectue en deux trames entrelacées. Dans ce dernier cas, toute la résolution temporelle est conservée au détriment de la résolution spatiale. Entre ces deux modes, un mode intermédiaire emploie le principe de la compensation de mouvement afin de maintenir une définition spatiale importante jusqu'à des vitesses élevées. Un synoptique de ce codeur est donné figure 1.



Dans ce système de transmission à 3 modes [1], il est nécessaire de transmettre une information relative au mode choisi ainsi que des informations de mouvements qui permettent la reconstruction des images HD à la réception. Dans le standard MAC, la période de retour trame représente un canal disponible d'une capacité d'environ 1 Mbit/s et permet, moyennant un système de réduction de débit, la transmission de ces informations dites d'assistance numérique (AN). Cet article est consacré à la description de la technique de réduction de débit utilisée pour le système HDMAC retenu par les participants au projet Eurêka EU95.

Dans la première partie de l'article, nous allons détailler les caractéristiques du signal AN à coder. Une deuxième partie sera consacrée à l'analyse des propriétés de ce signal qui nous ont amené à envisager 2 techniques de réduction de débit. Les performances obtenues par ces 2 techniques seront estimées. Dans une troisième partie, nous décrivons tout d'abord, un nouvel algorithme d'estimation de mouvement qui permet encore d'améliorer les performances du système de réduction de débit. La solution finale adoptée (dans le cadre du projet Eurêka) sera ensuite exposée. Nous concluons enfin par les points forts de cette dernière technique de réduction de débit.

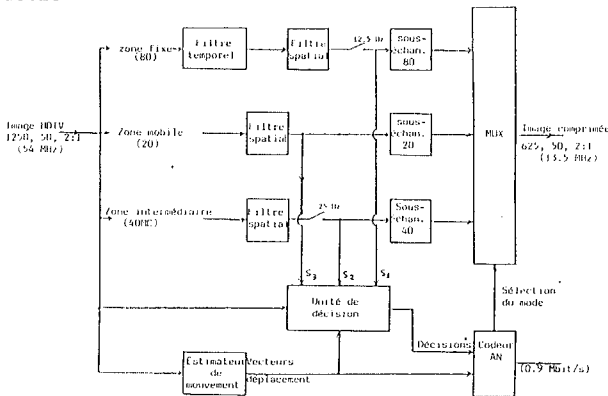


Figure 1 : synoptique codeur HDMAC

1. Caractéristiques du signal d'Assistance Numérique

Dans le système HD-MAC retenu par Eurêka, la commutation entre les trois différents modes s'effectue sur la base de blocs de 16 lignes de 16 échantillons pris dans l'image HD de 1250 lignes de 1440 pixels. Le mode choisi est celui qui permet de minimiser l'énergie de la différence entre l'image originale et l'image reconstruite. Pour chacun de ces blocs, il faut transmettre le type de filtrage spatio-temporel utilisé. Dans le cas du mode intermédiaire qui utilise le principe de l'interpolation temporelle compensée en mouvement [2], un vecteur mouvement associé au bloc est en plus nécessaire.

Dans la suite de l'article nous appellerons le mode fixe (l'échantillonnage temporel est effectué à 12.5 Hz et la transmission d'une image s'effectue en 4 trames de 20 ms) mode 80 ; le mode mobile (l'échantillonnage temporel est effectué à 50 Hz, la transmission d'une image étant alors réalisée en 2 trames entrelacées) mode 20 ; le mode intermédiaire (compensé en mouvement) (l'échantillonnage temporel est effectué à 25 Hz et la transmission d'une image est faite en 2 trames de 20 ms) mode 40CM.

1.1. Commutations entre modes

Afin d'éviter de trop fréquentes ruptures de phases d'échantillonnage, une cohérence temporelle sur 4 trames est introduite dans la prise de décisions. Cette cohérence limite à 5 les différents cas possibles de commutations sur un intervalle de 80 ms, dans le codeur HDMAC. Dans le mode 40CM, en plus de cette information il faut transmettre le vecteur mouvement associé.

1.2. Vecteurs mouvement

L'algorithme d'estimation de mouvement est un algorithme de corrélation de blocs (en anglais BMA) à recherche en trois étapes [3]. A chaque étape, le vecteur retenu est celui qui minimise l'énergie des DFD (Displaced Field Difference) sur un bloc 16 x 16.

$$DFD(x,y,\vec{V}) = I_0(x,y) - \left(\frac{I_{-1}(x-V_x, y-V_y) + I_1(x+V_x, y+V_y)}{2} \right)$$

$$E = \sum_{\text{bloc}} (DFD)^2$$

où $I_0(x,y)$ représente l'intensité du pixel de coordonnées x,y dans la trame 0, V_x, V_y les composantes en x et en y du vecteur déplacement, les trames I_{-1}, I_0, I_1 sont espacées de 20 ms chacune.

Cet algorithme permet d'obtenir les vecteurs dans la gamme de vitesse choisie (± 12 pixels/40 ms). Pour chaque bloc traité dans le mode 40CM, il existe donc 169 vecteurs possibles.

Il est à noter que l'on procède à 2 estimations indépendantes sur deux images successives.

1.3. Source à coder

Les informations de commutation entre modes et les vecteurs mouvement ont été regroupés en une seule source à coder. Le tableau I représente pour chacun des 5 cas, le nombre de possibilités de séquences de codage sur cet intervalle de 80 ms. Ceci permet de donner une borne supérieure au débit à transmettre. Pour décrire le mode choisi pour un bloc ainsi que le vecteur mouvement éventuellement associé, il suffit de $\log_2(1+169+169*169+1)$ bits, c'est-à-dire de 14,4 bits. Compte tenu du nombre de blocs à transmettre pour décrire l'image HD (81*10³/sec), un maximum de 1,16 Mbit/sec est suffisant. La capacité allouée, d'environ 900 Kbits/sec, requiert une simple réduction de 25 % du débit. Pour ce faire, nous avons envisagé dans un premier temps, un codage statistique de la source (utilisation de codes à longueur variable par exemple). Cependant, ce type de codage ne permet pas d'obtenir un débit fixe par image et pose des problèmes bien connus de régulation de débit et de sensibilité aux erreurs. Dans le cadre du projet Eurêka HDTV, de telles solutions ont été étudiées mais compte tenu des impératifs d'exploitation de la norme HDMAC (magnétoscope grand public par exemple) nous avons préféré un axe de recherche favorisant la simplicité de réalisation et une meilleure robustesse aux erreurs. En particulier, seuls des codes à longueur fixe et de longueur maximale 11 bits ont été envisagés.

T1	T2	T3	T4	Numéro de la voie	Nombre de cas possibles par voie
20	20	20	20	1	1
20	20	40CM	40CM	2	169
40CM	40CM	20	20	3	169
40CM	40CM	40CM	40CM	4	169*169
80	80	80	80	5	1

TABEAU 1

2. Système de réduction de débit par extraction de vecteurs dominants

2.1. Première technique

La première technique pour réduire le débit consiste à limiter le nombre de vecteurs transmissibles à un nombre N (Tableau II). Pour cela, on détermine image par image, l'histogramme des

vecteurs pour les voies dans lesquelles la compensation de mouvement est appliquée. De cet histogramme fenêtré par les décisions issues du codeur HDMAC, on extrait les N vecteurs ayant les occurrences les plus grandes (nous appellerons dans la suite ce jeu de N vecteurs, "les vecteurs dominants"). La transmission de ce dictionnaire s'effectue toutes les 80 ms par exemple. L'algorithme de réduction de débit consiste à indiquer au codeur HDMAC que pour les blocs initialement traités dans le mode 40 CM mais pour lesquels le vecteur ne fait pas partie des vecteurs dominants, il faut transmettre le mode 20 pour ce bloc (ce mode n'utilise pas de vecteurs mouvement).

La figure 2 représente le pourcentage de vecteurs transmissibles en fonction du nombre de vecteurs dominants conservés. Il apparaît que sur 4 des 5 séquences tests (un bref descriptif des séquences est donné en annexe) en choisissant N = 44 (1 mot code sur 11 bits) plus de 90% des vecteurs sont transmis.

T1	T2	T3	T4	Numero de la voie	Nombre de cas possibles par voie
20	20	20	20	1	1
20	20	40CM	40CM	2	N
40CM	40CM	20	20	3	N
40CM	40CM	40CM	40CM	4	N*N
80	80	80	80	5	1

TABLEAU 2

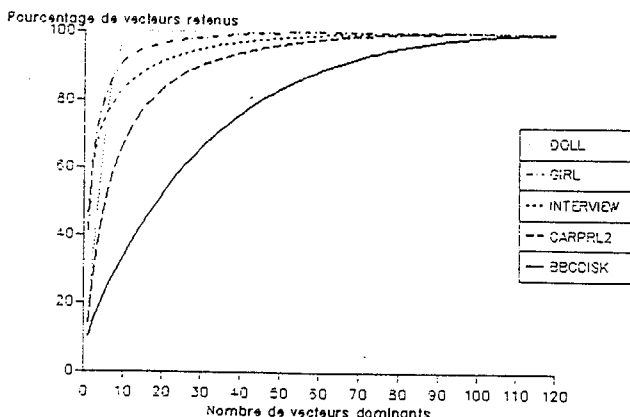


Figure 2

Les défauts visuels générés par les commutations dans le mode 20 provoqués par la réduction de débit sont négligeables car, c'est principalement sur des zones où la définition spatiale du mode 40CM apporte peu par rapport à celle du mode 20 qu'ils apparaissent (zone uniforme, texture fine peu contrastée par exemple). Cette faible visibilité des défauts s'explique par le fait que l'estimation de mouvement sur une zone uniforme ou peu contrastée est sensible au bruit et n'est pas toujours cohérente vis-à-vis du mouvement global mesuré sur des zones plus contrastées.

Cependant, sur la séquence BBC DISK, les commutations vers le mode 20 détériorent de façon importante la qualité des images. En effet, un mouvement de rotation rapide peut générer tous les vecteurs possibles simultanément. Un zoom rapide provoquerait des résultats similaires. Ce principe de réduction peut donc introduire des défauts sur certains types de séquences. Nous avons alors étudié une seconde technique qui permet de ne pas réduire le nombre de vecteurs transmissibles sur la première image. Ce second principe s'applique dans le cas où la voie 40CM est choisie sur les 2 images de l'intervalle de 80 ms.

2.2. Seconde technique

Cette seconde technique consiste à réaliser les mêmes opérations que celles décrites précédemment mais en utilisant comme source à réduire les variations temporelles des composantes du vecteur mouvement mesurées entre la première et la seconde image. Le tableau III donne le nombre de cas possibles et la figure 3 représente le pourcentage de ces ΔV transmis pour la deuxième image en fonction du nombre de variations dominantes gardées. On peut constater sur cette figure que dans le cas des 5 séquences, si l'on choisit $\Delta N = 10$ (qui permet de définir un mot code sur 11 bits) près de 100% des vecteurs sont transmissibles.

Une analyse de la variation temporelle des composantes des vecteurs a montré que celle-ci était très souvent égale à (0,0) (plus de 40% des occurrences sur chaque séquence) et que l'histogramme bidimensionnel de ces variations était centré en (0,0) et de forme losangée. Ceci démontre bien l'importante corrélation temporelle du mouvement sur un intervalle de 80 ms.

Les deux techniques décrites ci-dessus pourraient être combinées, ceci n'était pas utile puisque l'application de la deuxième technique seule répondait bien au problème posé. Cependant la transmission du dictionnaire des variations temporelles même bien protégé est un point faible de ce système vis-à-vis des erreurs introduites par le canal (satellite ou magnétique). Ces constatations nous ont amené à aller plus loin dans la technique de réduction de débit en envisageant un estimateur de mouvement à cohérence spatio-temporelle améliorée.

T1	T2	T3	T4	Numero de la voie	Nombre de cas possibles par voie
20	20	20	20	1	1
20	20	40CM	40CM	2	169
40CM	40CM	20	20	3	169
40CM	40CM	40CM	40CM	4	169 * ΔN
80	80	80	80	5	1

TABLEAU 3

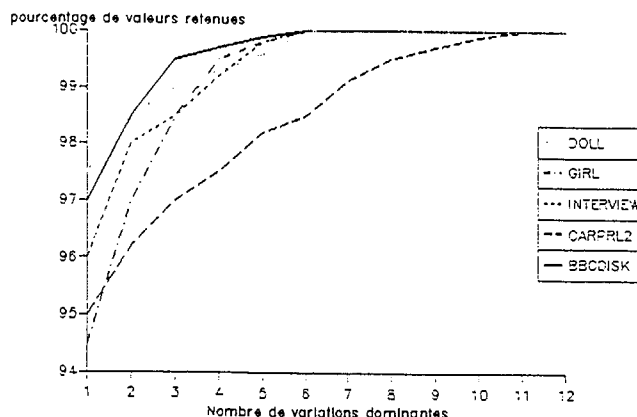


Figure 3

3. Système de réduction de débit à estimateur de mouvement modifié

Comme il a été démontré dans ce qui précède, les vecteurs mouvement sont fortement corrélés dans la direction temporelle. Cette propriété a été utilisée pour définir un nouvel estimateur de mouvement. Cet estimateur opère sur deux images successives (un intervalle de 80 ms) en deux phases temporelles distinctes.



La première phase utilise l'estimateur de mouvement décrit dans le paragraphe I pour la première image, la seconde phase détermine des vecteurs pour la seconde image dans l'intervalle de 80 ms de la façon suivante : on choisit comme vecteurs candidats, le vecteur déjà obtenu sur l'image précédente à position spatiale équivalente $V-1(x,y)$ ainsi que ces 8 voisins spatiaux (fig. 4). Le vecteur retenu $V0(x,y)$ parmi les 9 est celui qui minimise l'énergie de la DFD sur le bloc 16×16 .

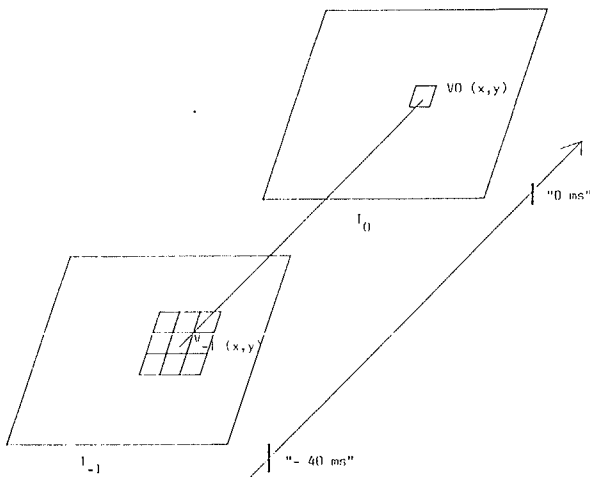


Figure 4

Le nombre de cas possibles apporté par cette modification est décrit dans le tableau IV. Cette nouvelle méthode d'estimation de mouvement permet de respecter le débit imposé. Nous avons alors vérifié les performances du système HDMAC muni de ce nouvel estimateur. D'un point de vue subjectif, aucune dégradation sur l'image n'a été observée. D'autre part, le nombre moyen de blocs traités dans le mode 40CM est resté inchangé. De plus, d'un point de vue distorsion, en n'utilisant que la voie 40CM, nous avons pu mesurer par rapport à l'estimateur précédent qui opérait de la même façon sur la première et la deuxième image de l'intervalle de 80 ms un accroissement des performances, variant suivant les séquences de -2 à -0.2 dB en terme d'énergie de l'erreur de compensation. Plus la corrélation temporelle dans la séquence était grande, plus important était ce gain. Ce résultat souligne qu'une solution au problème des minima locaux rencontrés dans des algorithmes de BMA à recherche non exhaustive a été trouvée pour des scènes naturelles. Il est important de noter que ce principe de réduction de débit peut être appliqué avec n'importe lequel des estimateurs par blocs (bloc récursif, méthode du plan de phase, ...). Par exemple, en utilisant la méthode du plan de phase [4] pour estimer les vecteurs sur la première image de l'intervalle de 80 ms, nous avons obtenu des résultats équivalents.

T1	T2	T3	T4	Numéro de la voie	Nombre de cas possibles par voie
20	20	20	20	1	1
20	20	40CM	40CM	2	8
40CM	40CM	20	20	3	169
40CM	40CM	40CM	40CM	4	169*9
80	80	80	80	5	1

TABLÉAU 4

Par rapport à la méthode exposée dans le paragraphe 2.2 ($\Delta N = 10$), les résultats subjectifs sont équivalents. Cependant, un meilleur rapport S/B sur l'image, l'absence de dictionnaire à transmettre, et la complexité matérielle évaluée nous ont amené à choisir cette simple modification de l'estimateur de mouvement comme technique de

réduction de débit. Cette technique est celle qui a été adoptée par Philips puis par Euréka 95 comme élément du système de référence devant servir de base à l'élaboration d'une proposition de norme HDMAC pour le CCIR.

CONCLUSION

Dans le cadre de la transmission d'images de TVHD dans un canal MAC, la compression qu'il est nécessaire d'appliquer au signal d'assistance numérique (AN) pour sa transmission est peu importante. Néanmoins les performances du système de codage HDMAC sont étroitement liées à la technique de réduction de débit adoptée. Après avoir présenté la nature du signal AN, essentiellement constitué de vecteurs mouvement, nous avons montré comment il était possible de façon adaptative et quasiment sans distorsion d'atteindre le débit requis. Cette méthode qui repose sur la limitation du nombre de vecteurs (vecteurs ou vecteurs variations) nécessite la transmission d'un ou de plusieurs dictionnaires. De ce fait, cette technique présente une relative sensibilité au bruit. Nous avons alors décrit un nouvel algorithme d'estimation de mouvement qui repose sur l'importante corrélation temporelle observée dans la mesure du mouvement. Les performances de cet algorithme qui en lui-même permet une réduction suffisante du débit ont été comparées à celles obtenues avec la méthode précédente. Une meilleure immunité au bruit, une simplicité de réalisation plus grande, font partie des éléments qui nous ont amené à choisir ce nouvel estimateur de mouvement comme système de réduction de débit. Les codes choisis pour la transmission du signal AN sont des codes à longueur fixe qui assurent une bonne robustesse au bruit en terme de propagation d'erreur et qui ne posent pas de problèmes de synchronisation particuliers. C'est pour toutes ces raisons que ce système d'estimation de mouvement à débit de données réduit a été retenu comme un élément du système de transmission HDMAC devant servir comme base à la proposition de norme de HDTV auprès du CCIR.

REFERENCES

- [1] F. Wreeswijk et al "HD-MAC Coding of high definition television signals", IBC 88
- [2] J.P. Arragon et al "Motion Compensated Interpolation techniques for HD-MAC", IBC 88
- [3] J. Musman et al "Advances in picture coding", Proceeding of the IEEE, Vol 73 n°4, April 1985
- [4] G.A. Thomas "Television Motion Measurement for DATV and other applications", BBC Research Department Report n° 1987/11

ANNEXE : Description des séquences tests

DOLL : Le fond de la scène est une vue du port de (source Baltimore avec des gratte-ciel. Au premier Philips) plan, une poupée avec un châle à carreaux est posée sur une chaise animée d'un lent mouvement de rotation. Un zoom avant est effectué à une vitesse très réduite.

GIRL : Deux jeunes fillettes assises dans une (source salle de jeux manipulent des jouets. Le CGETT) fond de la scène est constitué par un rideau contrasté bougeant légèrement.

INTERVIEW: Deux femmes assises sur un divan dialo- (source guent. Au bout d'un moment l'une d'elles BBC) se lève et son mouvement est suivi par la caméra.

CAR PRL2 : Une voiture blanche roule devant un (source immeuble dont les fenêtres sont garnies de Philips) stores vénitiens. La caméra suit le mouvement de la voiture qui est très rapide.

BBC DISK : Sur un fond noir uniforme, un calendrier (source de la BBC est fixé sur un disque animé BBC) d'une importante vitesse angulaire.