



Intérêt d'une prise de vue cyclopéenne dans un système de transmission d'images de télévision stéréoscopiques

P. Corlay, M. Gazalet, J.C. Kastellk, M. Gharbi
and F.X. Coudoux

IEMN, Département OAE,
Le Mont Houy, BP 311, 59304 Valenciennes Cedex, FRANCE

RESUME

Un système de transmission d'images de télévision stéréoscopiques basé sur une prise de vue à trois caméras: gauche, cyclopéenne et droite est présenté. La caméra cyclopéenne permet d'obtenir directement l'image compatible avec le système standard 2D et une symétrie de qualité entre les images gauche et droite reconstruites pour la visualisation en relief. L'image compatible est transmise avec des données d'assistance, ces données sont constituées de cartes de disparité et d'erreurs d'estimation et permettent de reconstruire les vues gauche et droite à la réception. Le schéma général et le codage des données sont détaillés. Une comparaison avec un système classique comprenant seulement deux caméras est effectuée.

ABSTRACT

A transmission system for stereoscopic television is proposed in this paper. It is based on three cameras taking pictures: left, cyclopean, right. The cyclopean camera allows to directly obtain the 2D compatible TV picture as a symmetry of quality between the left and right restored pictures for the 3D visualization. The 2D compatible picture is transmitted with assistance data. These data consist on maps of disparity and estimation errors; they allow to reconstruct left and right views before final display. The proposed general scheme as the data coding are presented. This scheme is compared to the classical coding one using only two cameras.

INTRODUCTION

Le développement de la TV stéréoscopique, dans une optique de transmission "grand public" ne pourra se faire que par remplacement progressif des systèmes de télévision 2-D. La compatibilité entre les deux systèmes sera alors nécessaire. On transmettra une image 2D dite "compatible", associée à des données d'assistance (D.A.) permettant de reconstruire les vues stéréoscopiques à la réception. La prise de vue stéréoscopique est constituée de deux caméras fournissant deux images droite (D) et gauche (G) de la même scène. Ces deux images étant fortement corrélées, on cherchera naturellement à éliminer cette redondance.

A cet effet, deux types de schémas de transmission sont couramment proposés:

- le schéma le plus classique consiste à transmettre l'image G (ou D) comme image "compatible" et à reconstruire l'autre vue D (respectivement G) à partir des D.A. [1][2]. Ces schémas ont une image compatible de "bonne" qualité mais à la reconstruction les qualités des images D et G diffèrent.

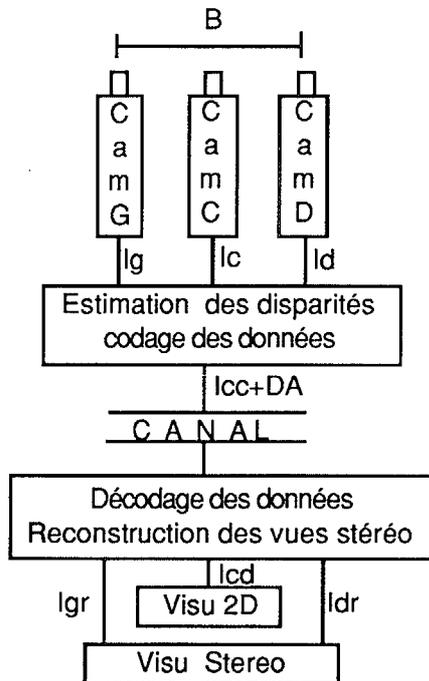
- Certains schémas proposent de symétriser la qualité des deux images stéréoscopiques en les reconstruisant toutes deux à partir de l'image compatible 2D. Cette dernière est obtenue dans l'émetteur à partir des vues droite et gauche, soit par calcul en estimant l'image cyclopéenne [3], soit en créant une image qui contient en partie des points sources des images droite et gauche [4]. De ce fait, l'image compatible n'est pas une image "réelle" et la qualité de la restitution 2D en est affectée.

Afin de pallier les défauts des deux schémas précédents, nous proposons ici un système de codage basé sur une prise de vue comprenant trois caméras: une caméra "cyclopéenne" est associée aux deux caméras G et D (voir figure 1). Dans cet article, nous comparerons ce schéma à celui comportant seulement les caméras G et D et basé sur un codage similaire. Nous détaillerons le codage des données d'assistance et donnerons les résultats de la comparaison entre les systèmes à deux et trois caméras.



SYNOPTIQUE DU SYSTEME A 3 CAMERAS

La troisième caméra est placée entre les deux caméras G et D habituelles (figure 1).



Avec B = base de la prise de vue; I_d = Image droite; I_g = Image gauche; I_c = Image cyclopéenne = Image compatible ; I_{cc} = Image compatible transmise et compressée; DA = Données d'Assistance; I_{cd} = Image compatible transmise et décompressée; I_{dr} = Image droite reconstruite; I_{gr} = Image gauche reconstruite.

FIGURE 1 : SYSTEME A TROIS CAMERAS

La caméra C fournit directement l'image compatible qui sera donc de "bonne" qualité. En codant les images G et D à partir de l'image C, on répartit les dégradations possibles entre les vues G et D.

La symétrie par rapport à la caméra C permet d'obtenir deux champs de disparité gauche-centre (GC) et droite-centre (DC) largement symétriques et leur estimation sera d'autant plus précise que la proportion des faces cachées sera diminuée de moitié par rapport aux couples (GD) (ceci étant dû à la diminution de la base entre les caméras B→B/2).

CODAGE DES DONNEES D'ASSISTANCE

Les données d'assistance sont normalement constituées de deux cartes de disparité (GC et DC) ainsi que des deux images DFD (Displaced Frame Difference) traduisant les erreurs de prédictions des images D et G à partir de l'image C et des disparités.

a) Estimation et codage des champs de disparité

Il existe aujourd'hui différents algorithmes pour estimer les cartes de disparité[5]. Le but étant de comparer deux schémas de codage, nous utiliserons dans un premier temps, une simple mise en correspondance (Block Matching) des macroblocs 16*16 pixels. De plus nous n'avons pas tenu compte ici de la redondance temporelle. Nous avons choisi dans l'immédiat et dans un but de simplicité de placer les caméras de sorte que leurs axes optiques soient parallèles. Les cartes de disparité seront donc représentées par des vecteurs à une composante d.

Celle-ci est obtenue de manière à minimiser pour chaque macrobloc de l'image G ou D, la quantité suivante :

$$\sum_{x, y} \text{du bloc} |DFD_{G \text{ ou } D}(x, y, d)|$$

avec

$$DFD_{G \text{ ou } D}(x, y, d) = I_{G \text{ ou } D}(x, y) - I_C(x - d, y)$$

où $I_{G \text{ ou } D}(x, y)$, $I_C(x - d, y)$ sont les intensités lumineuses au point (x,y) de l'image G ou D, au point (x-d,y) de l'image C.

Les cartes de disparité étant par nature largement continues, il est intéressant de les coder différenciellement. Pour nos images sources (de taille 768*512 pixels) une entropie de l'ordre de 3 bits par macrobloc pour les cartes de disparité différentielles est obtenue pour chacun des couples (GC) et (DC).

La symétrie des cartes de disparité GC et DC pourrait être exploitée, par exemple en estimant la carte (GC) à partir de l'autre carte (DC). Il suffirait alors de transmettre la carte (DC) et l'erreur d'estimation de la carte (GC). Toutefois, nos résultats expérimentaux ont montré que l'on n'obtient pas de débits significativement inférieurs à ceux obtenus lors du codage différentiel séparé des deux cartes (DC) et (GC).



b) codage des images DFD

A partir de l'image compatible, et de l'estimation des disparités, on obtient deux images DFD qui sont codées et transmises et qui permettent après décodage de reconstruire les vues G et D.

Nous avons choisi le schéma de compression le plus simple associant une transformation de l'image DFD par Transformée en Cosinus Discrète 8*8 pixels suivie d'une quantification avec zone morte afin de ne conserver que les composantes fréquentielles dont l'amplitude dépasse un certain seuil. La quantification peut être identique pour toutes les composantes DCT ou tenir compte de l'importance visuelle des erreurs de quantification par l'intermédiaire d'une pondération appropriée [6].

RESULTATS ET COMPARAISON

Les tests ont été effectués sur une scène de bureau en environnement informatique. Nous nous proposons de comparer les débits après codage des D.A. pour un même rapport signal à bruit entre les systèmes à deux et trois caméras.

Les résultats significatifs représenteront l'évolution du rapport signal à bruit (S/B) pondéré pour tenir compte de la perception des dégradations pour chaque composante DCT, en fonction du débit total (Images DFD + cartes de disparité). Nous avons testé différentes tailles de zones mortes (z_m de 0 à 3). La table de quantification retenue est le produit d'un paramètre scalaire par une table de base optimisée pour un meilleur rapport S/B pondéré.

Pour le schéma à deux caméras, nous avons considéré les deux cas possibles: l'image reconstruite est l'image D ou G. Pour celui à trois caméras, nous présentons les courbes correspondantes aux deux images G et D reconstruites.

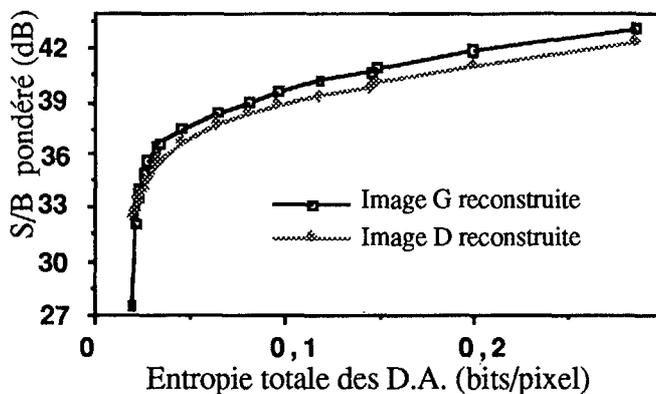


FIGURE 2 : RAPPORT S/B EN FONCTION DU DEBIT TOTAL DES D.A. POUR LE SYSTEME A TROIS CAMERAS

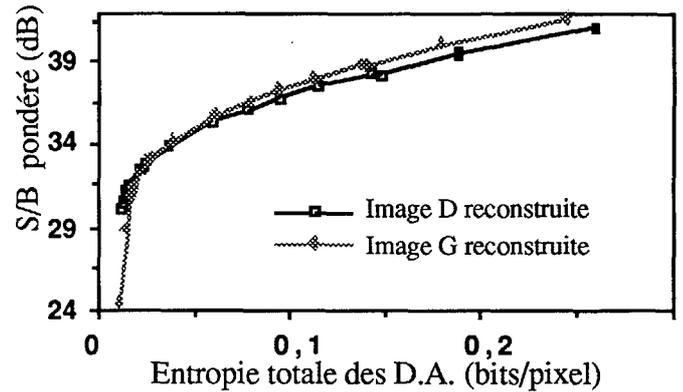


FIGURE 3 : RAPPORT S/B EN FONCTION DU DEBIT TOTAL DES D.A. POUR LE SYSTEME A DEUX CAMERAS

Nous avons constaté que les courbes, quel que soit le débit, sont très peu affectées par la largeur de la zone morte (z_m variant de 1 à 3), l'optimum correspondant à une zone morte de 1 et la quantification sans zone morte ($z_m=0$) donne des résultats sensiblement inférieurs. Les figures 2 et 3 représentent les résultats obtenus pour cette valeur idéale $z_m=1$. La comparaison montre que, même si ce n'était pas le but recherché, l'amélioration souhaitée se situant au niveau de la meilleure compatibilité 2D-3D, le débit à qualité égale est en faveur du système à trois caméras, au moins pour nos images tests. Ceci s'explique par le fait qu'en utilisant la caméra C, nous améliorons nettement l'estimation des images codées par disparité. Cette comparaison est effectuée entre le cas le moins favorable pour la configuration à trois caméras (avec nos images, l'image D) et le cas le plus favorable pour la configuration à deux caméras (c.a.d quand l'image reconstruite 2D est l'image G). Par exemple pour un débit total des D.A. de l'ordre de 0.1 bit par pixel, nous obtenons un rapport S/B légèrement supérieur à 38 dB pour notre système contre 37 dB pour le schéma à deux caméras.

Une partie, de l'ordre du tiers, de l'image erreur gauche (différence entre l'image source G et l'image reconstruite G) est représentée pour le système à trois caméras (figure 4) et pour celui à deux caméras (figure 5). Le débit des D.A. dans ces deux cas est de l'ordre de 0.1 bit par pixel.

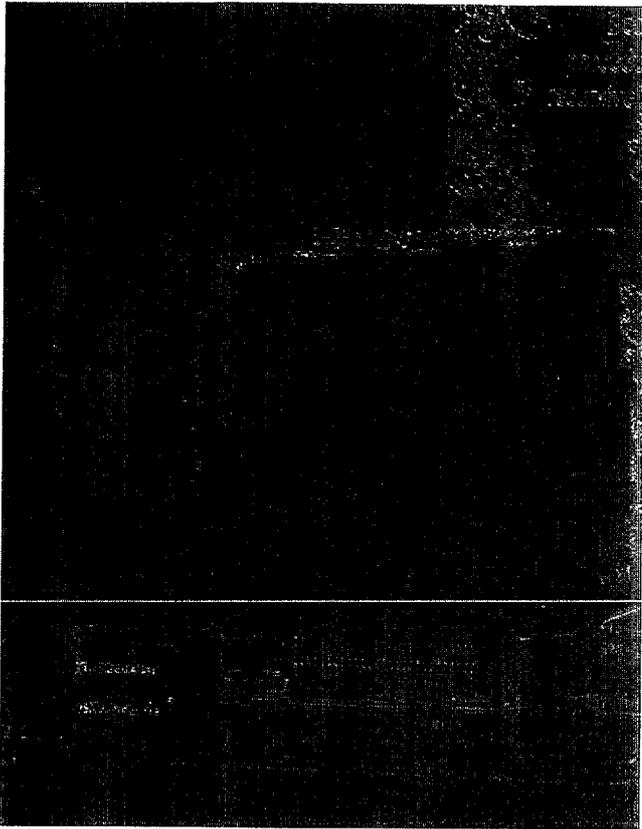


FIGURE 4 : PARTIE DE L'IMAGE ERREUR GAUCHE
POUR LE SYSTEME A TROIS CAMERAS

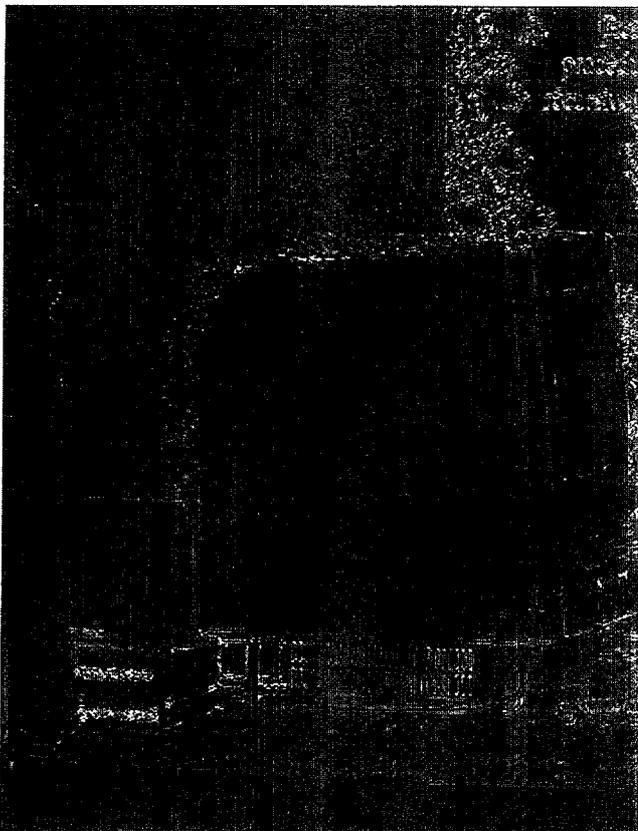


FIGURE 5 : PARTIE DE L'IMAGE ERREUR GAUCHE
POUR LE SYSTEME A DEUX CAMERAS

CONCLUSION

Une méthode originale de transmission d'images de télévision stéréoscopiques a été proposée. Elle est basée sur une prise de vue à trois caméras: gauche, cyclopéenne et droite.

Cette méthode permet d'obtenir :

- une compatibilité totale avec le système de télévision standard.
- une symétrie de qualité entre les images stéréoscopiques après reconstruction lors de la visualisation en relief.

Une comparaison de ce système à un schéma comprenant seulement deux caméras a été effectuée. Les premiers résultats de cette comparaison sont satisfaisants: les images reconstruites sont de meilleure qualité pour un même débit des données additionnelles.

BIBLIOGRAPHIE

- [1] M.G. Perkins, "Data compression of stereopairs", IEEE Transactions on Communications, Vol. 40, No. 4, p. 684-696, Avril 1992.
- [2] I.Dinstein, M.G. Kim, A.Henik, J. Tzelgov, "Compression of stereo images using subsampling and transform coding", Optical Engineering, Vol. 30, No. 9, September 1991.
- [3] D. De Vleeschauwer, "Coding of the disparity field and the correction signal in predictive 3DTV coding", First International Symposium on 3D Images, Paris, September 26-28, 1991.
- [4] F.Chassaing, B.Choquet and D.Pelé, "A stereoscopic television system (3D-TV) and compatible transmission on a MAC channel (3D-MAC)", Signal Processing : *Image Communications*, No. 4, p. 33-43, 1991.
- [5] B.Choquet, "Télévision 3-D: codage et traitement des images", Journée d'études VISION 3D, CNET Paris, 06 Mai 1993.
- [6] L.Vandendorpe, "Human visual quantization: a unified approach for subband and pyramidal image coders", Ann. Telecommun., Vol. 47, No. 7-8, p. 282-292, 1992.