

Analyse des performances d'un système de compression distribuée de séquences vidéo transmises sur un lien non fiable

Joumana Farah¹, Charles Yaacoub¹, François Marx², Béatrice Pesquet-Popescu³

¹Faculté des sciences et de génie informatique, Université St-Esprit de Kaslik, B.P. 446 Jounieh, Liban.

{joumanafarah, charlesyaacoub} @ usek.edu.lb

²France Telecom ROSI, Paris, France

francois.marx @ orange-ftgroup.com

³Département Traitement du Signal et des Images, Ecole Nationale Supérieure des Télécommunications, 75634 Paris Cedex 13, France.

beatrice.pesquet @ enst.fr

Résumé – Dans cet article, nous considérons un système de transmission de séquences vidéo comprimées selon le concept Wyner-Ziv et transmises en présence de bruit. Nous présentons une nouvelle méthode d'estimation des limites théoriques de compression de ce système en prenant en compte la quantité de mouvement dans chaque trame de la séquence ainsi que les conditions de transmission sur le canal. Les études réalisées jusqu'à présent sur de tels systèmes de codage ne prennent pas en compte l'influence d'une transmission sur un lien non fiable. De plus, elles ne présentent pas d'estimations des limites théoriques de compression. Dans ce travail, nous montrons qu'en se basant sur les limites de compression estimées, on peut se débarrasser du canal de retour dans le cas d'applications de diffusion. De plus, la perte en performances reste minimale par rapport aux systèmes classiques avec canal de retour.

Abstract – In this paper, we consider a video transmission system based on Wyner-Ziv coding. Unlike previous works that considered error-free transmission of the encoded data, we study the influence of noise on the system performances. We also present a novel approach for estimating the theoretical compression bound of each video frame based on the amount of motion inside the frame and on the channel conditions. Furthermore, we show how the entropy limits can be used for estimating the optimal amount of information to be transmitted per frame, in case of a broadcasting application where feedback information from the receiver is not available. Simulation results show a small degradation in the system rate-distortion performance, compared to the case of a source-channel coding system with a feedback channel.

1. Introduction

Le codage distribué [1-4] a connu dernièrement d'importants progrès dans le domaine de transmission d'images. Le but principal est de contrer la complexité élevée, dans l'encodeur, des techniques de compression classiques telles que les normes MPEG basées sur le codage prédictif inter-frames. Le concept a émergé de la théorie de Slepian et Wolf [5] : étant donné deux sources corrélées X et Y , avec Y , l'information de bord, comprimée à sa limite entropique $H(Y)$, X peut être transmise à un débit très proche de l'entropie conditionnelle $H(X/Y)$. L'application de ce concept au codage source avec pertes est connu par le codage de Wyner-Ziv [6] qui a été appliqué dans les réseaux de capteurs [3] et dans les codeurs vidéo à basse complexité [7]. Dans ce dernier cas, une compression intra-trame de la

source est réalisée par un poinçonnage adéquat des bits de parité à la sortie d'un turbo-codeur. Les informations systématiques ne sont pas transmises mais remplacées par l'information de bord dans le décodeur inter-frames. L'un des inconvénients majeurs de cette structure est qu'elle se base sur un canal de retour entre l'émetteur et le récepteur, ce qui rend son application dans un environnement temps-réel ou une application de diffusion (Broadcasting) quasi-impossible. De plus, le comportement d'un tel système dans des conditions de transmission réalistes n'a pas été considéré. Dans ce travail, nous considérons tout d'abord un système de codage avec canal de retour et étudions ses performances lors d'une transmission sur un canal à bruit blanc additif Gaussien (AWGN). Nous déterminons ensuite les limites théoriques de compression d'un tel système de codage.

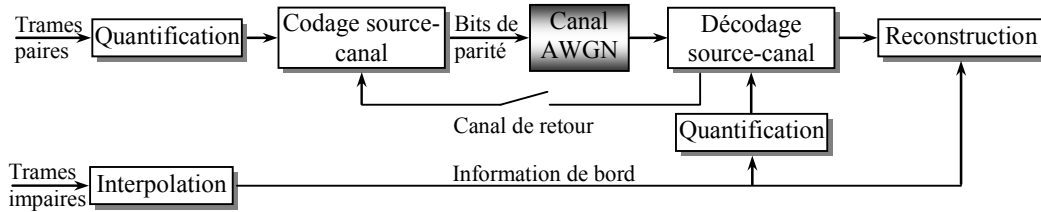


FIG. 1 : Schéma-bloc du système de codage conjoint source-canal.

Ces limites seront utilisées dans le système de diffusion, de façon à faire varier le niveau de compression d'une trame à l'autre en prenant en compte la quantité de mouvement dans chaque trame ainsi que les conditions de transmission.

2. Description du système

Le système de codage conjoint distribué considéré dans cette étude peut être représenté par le schéma-bloc de la fig.1. Le codeur intra-trame comprime uniquement les trames paires de la séquence vidéo. Les trames impaires sont supposées être parfaitement reçues au décodeur où elles constituent les informations de bord. Ces dernières seront utilisées en tant que données systématiques par le décodeur source-canal ainsi que dans la reconstruction des trames paires. L'information de bord pour une trame paire particulière est générée par une interpolation moyenne des trames impaires adjacentes [7]. Par ailleurs, on peut vérifier que le signal résiduel d , différence de la trame paire avec la trame interpolée, suit une distribution Laplacienne :

$$P(d) = \frac{\alpha}{2} e^{-\alpha|d|}. \quad (1)$$

Le paramètre α peut être estimé dans le récepteur en se basant sur les trames impaires ou dans le codeur (et transmis comme information de bord au récepteur).

Une quantification scalaire uniforme est appliquée sur les trames paires pour obtenir une représentation sur M bits des pixels de chaque image.

Dans ce travail, nous considérons $M = 1, 2$ ou 4 pour obtenir différents débits de codage source. Le codeur source-canal consiste en une concaténation parallèle de deux codeurs convolutifs quadri-binaires à 16 états séparés d'un entrelaceur interne et résultant en un rendement de codage global minimal de $2/3$. Les polynômes générateurs sont $(23, 35, 31, 37, 27)_8$ de [8]. La taille de l'entrelaceur est égale à celle de la trame quantifiée. A la sortie du codeur, seules les informations de parité sont poinçonnées et transmises au décodeur. De cette façon, le taux de compression maximal du système est 0.5 , ce qui correspond à la transmission de la totalité des bits de parité. Du côté du récepteur, le turbo-décodage est réalisé à l'aide de décodeurs itératifs à entrées souples et sorties souples (SISO) basés sur l'algorithme Max-Log-MAP (Maximum A Posteriori) [9]. Cependant, nous avons modifié le calcul des métriques de façon à prendre en compte les 16 transitions possibles entre deux états

quelconques du treillis. De plus, les probabilités conditionnelles dans le processus de turbo-décodage doivent prendre en compte la statistique du signal résiduel entre les trames paires et l'information de bord ainsi que l'état du canal de transmission. Dans notre cas, l'information de parité est transmise sur un canal AWGN.

Dans le cas d'un système de transmission avec un canal de retour, le processus de décodage débute avec une quantité minimale de bits de parité. Le décodeur demande des bits de parité supplémentaires de la part de l'émetteur jusqu'à ce qu'un taux d'erreurs par symbole de 10^{-3} soit atteint. Cependant, dans le cas d'une application de diffusion, le système doit se baser sur une quantité prédéterminée de bits de parité : cette quantité peut être estimée à partir des limites théoriques de compression (H_f) des trames paires calculées par l'émetteur. Pour un débit de transmission fixe (par exemple, le débit moyen D_t obtenu par le système avec canal de retour), l'émetteur détermine le schéma de répartition quasi-optimal de la quantité totale d'informations de parité entre les différentes trames. Le débit de transmission pour une trame particulière f sera :

$$D_f = D_t \frac{H_f}{H_t}, \quad (2)$$

où H_t est la somme des limites de compression de toutes les trames de la séquence.

Après décodage source-canal, le bloc de reconstruction permet de récupérer une version à 8 bits des pixels de la trame, en se basant sur l'information de bord [7] : chaque symbole décodé est comparé à l'info de bord correspondante. Si cette dernière se trouve dans l'intervalle de quantification du symbole décodé, le pixel reconstruit prend la valeur de l'info de bord. Sinon, il prend la valeur de la limite d'intervalle la plus proche de l'info de bord.

3. Estimation de la limite de compression

La limite théorique de compression est calculée, en l'absence de bruit, à partir de l'entropie conditionnelle $H(X|Y)$, où X est la source discrète transmise et Y l'information de bord. X et Y sont des pixels quantifiés à M bits, par suite :

$$H(X|Y) = - \sum_{i=0}^{2^M-1} \sum_{j=0}^{2^M-1} P(Y=j) \cdot g[P(X=i|Y=j)], \quad (3)$$

où $g(x) = x \cdot \log_2(x)$.

Par ailleurs, on a :

$$P(X - Y = d) = c \frac{\alpha}{2} e^{-\alpha |d_{X-Y}|} \quad (4)$$

où $d_{X-Y} = d \cdot 2^{8-M}$ et c est un facteur qui permet de prendre en compte la nature discrète et limitée du signal résiduel entre X et Y :

$$c = (2/\alpha) \left/ \sum_{i=-(2^M-1)}^{2^M-1} e^{-\alpha |d_i|} \right., \quad (5)$$

avec $d_i = i \cdot 2^{8-M}$.

D'un autre côté :

$$P(X - Y = d) = \sum_j P(X = i = d + j | Y = j) P(Y = j). \quad (6)$$

Le nombre de termes non nuls dans cette sommation est égal au nombre $L_{d_{i-j}}$ de couples (i, j) qui donnent la différence $i - j = d$. Dans le cas d'une source équiprobable, on peut supposer que ces couples ont tous la même probabilité d'occurrence. Par exemple, si on considère le cas d'une quantification à 2 bits, les valeurs possibles de d sont $0, \pm 64, \pm 128$ et ± 192 , avec un ordre décroissant de la probabilité d'occurrence. Pour une valeur particulière de la différence, par exemple $d = 64$, les couples $(64, 0)$, $(128, 64)$ ou $(192, 128)$ ont la même probabilité d'avoir lieu. Par suite :

$$P(X = i = d + j | Y = j) = \frac{2^M}{L_{d_{i-j}}} P(X - Y = d). \quad (7)$$

Généralement, la différence résiduelle prend les valeurs suivantes :

$d_i = i \cdot 2^{8-M}$, $i = -(2^M - 1), \dots, 2^M - 1$, et l'on peut exprimer L_{d_i} en fonction de i par :

$$L_{d_i} = 2^M - |i|. \quad (8)$$

Finalement :

$$H(X | Y) = - \sum_{i=0}^{2^M-1} \sum_{j=0}^{2^M-1} \frac{1}{2^M} \log \left[c \frac{\alpha}{2} \frac{2^M e^{-\alpha |d_{i-j}|}}{L_{d_{i-j}}} \right]. \quad (9)$$

Dans le cas d'une transmission des bits d'information sur un canal binaire symétrique (BSC) de probabilité de transition p , il est connu que la limite de compression devient plus importante que $H(X|Y)$ et peut s'écrire [1] :

$$H_N(X | Y) = \frac{H(X | Y)}{C(p)}, \quad (10)$$

où $C(p)$ est la capacité du canal BSC. Afin de dériver la limite théorique de notre système de compression pour un canal AWGN, nous avons effectué un "mapping" de ce dernier sur un canal BSC en utilisant l'équivalence des fonctions de stabilité des deux types de canaux [10]. Cette équivalence a été adoptée dans [11] pour la conception de schémas particuliers de codage conjoint source-canal. La relation entre p et le rapport de l'énergie du symbole sur la densité spectrale de puissance du bruit (E_s/N_0) du canal AWGN peut s'écrire :

$$p = \frac{1}{2} \left(1 - \sqrt{1 - \exp\left(-2 \frac{E_s}{N_0}\right)} \right). \quad (11)$$

4. Résultats

Dans les figures 2 et 3, nous montrons le taux de compression (R: 1dB) obtenu en simulant le système de codage conjoint de la fig.1, ainsi que la limite théorique inférieure (H: 1dB), pour chaque trame, obtenus avec $E_s/N_0 = 1$ dB et $M = 4$ et 1 (respectivement). Nous avons considéré 50 trames de la séquence vidéo Carphone échantillonnée à 15 trames paires par seconde. Les courbes 'R' et 'H' sont obtenues pour une transmission sans bruit. Pour $M = 4$, on remarque, en l'absence de bruit, une différence de performance de la compression réalisée entre 0.06 et 0.14 par rapport aux courbes théoriques. Lorsque $E_s/N_0 = 1$ dB, la gamme de différence augmente à [0.11 ; 0.18]. Notons que les valeurs importantes des limites de compression pratiques ou théoriques correspondent aux faibles valeurs du paramètre α estimé à $M = 4$. En fait, une faible valeur de α indique un niveau de mouvement élevé dans la trame considérée, ce qui rend l'information de bord correspondante moins fiable pour le processus de turbo-décodage. Par suite, une grande quantité d'informations de parité sera requise de la part du codeur.

Des effets similaires ont été observés pour $M = 2$ et $M = 1$. Cependant, à $E_s/N_0 = 1$ dB, la gamme de différence augmente à [0.14 ; 0.25] pour $M = 2$ et [0.12 ; 0.33] pour $M = 1$ (fig. 3). Les valeurs élevées des niveaux de compression pratiques, pour les deux cas d'une transmission sans bruit ou en présence de bruit, sont dues au fait que, lorsque le nombre de niveaux de quantification est faible, la corrélation entre les pixels quantifiés des trames paires et ceux des trames interpolées diminue de façon dramatique. Ceci conduit à des fluctuations importantes du niveau de compression d'une trame à l'autre et explique les gammes de différence plus larges par rapport à celles obtenues pour $M = 4$.

Finalement, dans la fig. 4, nous montrons les courbes de débit-distorsion, i.e. le PSNR (Peak Signal to Noise Ratio) en fonction du débit de transmission moyen, obtenues pour différentes valeurs du rapport E_s/N_0 . Les limites théoriques (courbes 'TL') sont estimées en considérant uniquement l'influence de la quantification; c. à d. qu'on suppose un décodage parfait des trames à M bits à l'entrée du bloc de reconstruction. A $M = 4$, on remarque une différence entre les débits théoriques et pratiques entre 145 kbps (pour une transmission sans bruit) et 218 kbps (pour $E_s/N_0 = 1$ dB). Pour $M = 2$, cette différence est d'environ 150 kbps à $E_s/N_0 = 1$ dB. Elle décroît à 91 kbps pour $M = 1$. La perte en débit binaire dans le système de compression pratique, due à l'influence du bruit et mesurée à $M = 4$, est de 13 kbps à $E_s/N_0 = 5$ dB, 46 kbps à $E_s/N_0 = 3$ dB, et atteint 124 kbps à $E_s/N_0 = 1$ dB.

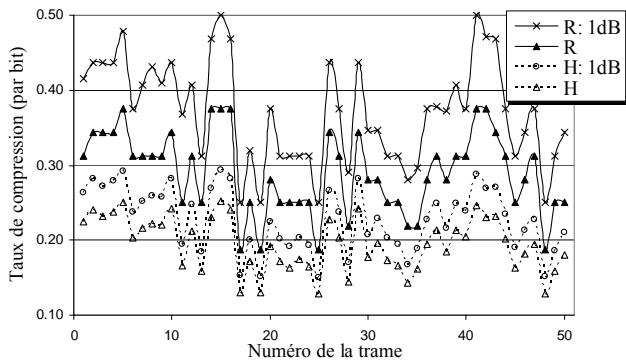


FIG. 2 : Limites de compression théoriques et pratiques pour $M = 4$.

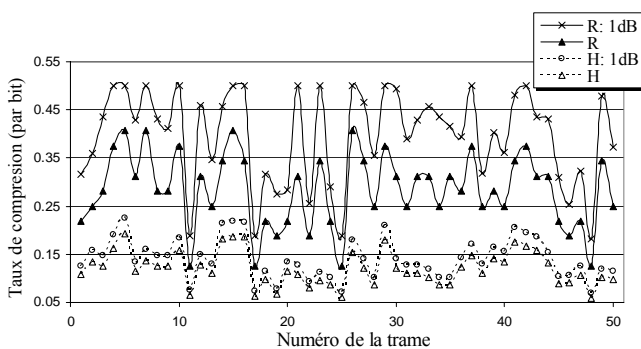


FIG. 3 : Limites de compression théoriques et pratiques pour $M = 1$.

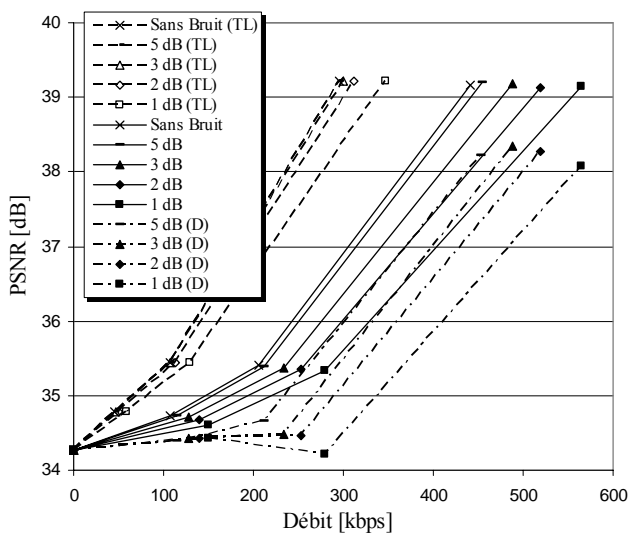


FIG. 4 : Courbes de débit-distorsion pour la séquence Carphone transmise en présence de bruit.

Dans le cas d'un système de diffusion (courbes 'D'), pour un même débit binaire, une perte en PSNR entre 0.85 et 1.1 dB est observée pour $M = 2$ et 4, par rapport au système avec canal de retour. En observant les performances du système de diffusion à bas débits, on remarque que lorsque le niveau de bruit est élevé, il est préférable que le décodeur se base beaucoup plus sur

l'information de bord que sur les informations de parité puisque ces dernières ne permettent par une amélioration notable du PSNR.

5. Conclusion

En observant les niveaux de compression trame par trame, nous avons montré que le système de compression pratique présente un comportement comparable à celui des résultats théoriques : la limite de compression réalisable pour une trame dépend de son contenu et des conditions de transmission. Par suite, nous pouvons utiliser les résultats analytiques pour prédire le niveau de compression pour chaque trame dans le codeur et l'utiliser dans un système de diffusion moyennant une faible perte en PSNR, par rapport au système classique avec canal de retour.

Remerciements

Ce travail a été en grande partie financé par le CNRS Libanais.

Références

- [1] A. Aaron et B. Girod, "Compression with Side Information Using Turbo Codes", *Data Compression Conference*, pp. 252- 261, Avril 2002.
- [2] M. Jeanne, R. Zhang, B. Girod et P. Siohan, "Codage de sources distribuées : comparaison de 2 systèmes approchant la borne de Wyner-Ziv", *CORESA*, France, Janvier 2003.
- [3] S. S. Pradhan et K. Ramchandran, "Distributed Source Coding: Symmetric rates and Applications to Sensor Networks", *Data Compression conference*, pp. 363-372, Mars 2000.
- [4] J. Farah, C. Yaacoub, N. Rachkidy et F. Marx, "Binary and non-Binary Turbo Codes for the Compression of Correlated Sources Transmitted through Error-Prone Channels", *4th International Symposium on Turbo Codes & Related Topics with the 6th International ITG-Conference on Source and Channel Coding*, Allemagne, Avril 2006.
- [5] D. Slepian et J.K. Wolf, "Noiseless Coding of Correlated Information Sources", *IEEE Transactions on Information Theory*, Vol. IT-19, pp. 471-480, Juillet 1973.
- [6] D. Wyner et J. Ziv, "The Rate-Distortion Function for Source Coding with Side Information at the Decoder", *IEEE Transactions on Information Theory*, Vol. IT-22, pp. 1-10, Janvier 1976.
- [7] A. Aaron, R. Zhang et B.Girod, "Wyner-Ziv Coding of Motion Video", *Thirty-Sixth Asilomar Conference on Signals, Systems and Computers*, pp. 240-244, Nov. 2002.
- [8] D. Divsalar et F. Pollara, "Multiple Turbo Codes", *Military Communication Conference*, pp. 279-285, 1995.
- [9] P. Robertson, P. Hoeher et E. Villebrun, "Optimal and Suboptimal Maximum A Posteriori Algorithms Suitable for Turbo Decoding", *European Transactions on Telecommunications*, pp. 119-125, Vol. 8, Mars-Avril 1997.
- [10] S.-Y. Chung, "On the Construction of Some Capacity-Approaching Coding Schemes", Ph.D. thesis, Massachusetts Institute of Technology, 2000, pp. 137-155.
- [11] A.D. Liveris, Z. Xiong et C. N. Georgiades, "Joint Source-Channel Coding of Binary Sources with Side Information at the Decoder Using IRA codes", *IEEE Workshop on Multimedia Signal Processing*, pp. 53-56, Dec. 2002.