

Codage distribué des images fixes

R. Guerhazi *, K. Idrissi *, C. Ben Amar **, A. Baskurt *

*LIRIS, Université Claude Bernard, Lyon 1
Bât. Nautibus, 8 boulevard Niels Bohr, 69622 Villeurbanne cedex – France*

**{rguermaz, kidrissi, abaskurt}@liris.cnrs.fr*

*REGIM, Ecole Nationale d'ingénieurs de Sfax
Route de Soukra, B.P. W, 3038 Sfax – Tunisie*

*** Chokri.BenAmar@enis.rnu.tn*

Résumé : Le concept du codage de sources distribuées ou DSC pour «Distributed Source Coding» fut introduit en 1973 par Slepian et Wolf. Ce domaine de la théorie de l'information traite la compression séparée des sources corrélées et a pour but la minimisation du coût de leur description conjointe.

Dans le cadre de la transmission à partir de terminaux portables, il est important que le schéma fonctionnel de la chaîne soit d'une part constitué d'un codeur simple, et d'autre part robuste aux erreurs de transmission afin d'être adapté aux caractéristiques d'un réseau sans fil à faible débit.

Cet article propose une méthode innovante de compression sans pertes d'images fixes grâce à l'utilisation du théorème de Slepian et Wolf, et adaptée à l'environnement pervasif, déportant ainsi la complexité vers le décodeur.

Mots clés: Codage distribué, compression sans pertes, images fixes.

1. Introduction

Dans les systèmes classiques de compression d'images et de vidéos tels que JPEG, SPIHT, MPEG ou H.26x, le codeur doit réaliser de nombreux calculs (estimation de vecteurs mouvements, décomposition en ondelettes, transformation en DCT...) afin d'exploiter les similarités entre les pixels pour les images fixes, et entre les trames successives d'une séquence pour la vidéo. Il est donc typiquement 5 à 10 fois plus complexe que le décodeur. Ces standards sont naturellement bien conçus pour un système composé de terminaux peu complexes, qui réceptionnent des données depuis un serveur. En revanche ils risquent de ne plus être adaptés aux nouveaux scénarios de transmissions, en particulier le schéma inverse dans lequel le codeur est embarqué. Ce dernier n'est plus sur un serveur puissant, mais sur un appareil portatif incapable de réaliser les nombreuses opérations requises pour le codage, a fortiori si ces opérations doivent répondre à des critères de temps réel. Il est alors nécessaire d'envisager une autre architecture de codecs permettant, à qualité équivalente, de simplifier la partie codage quitte à complexifier la partie décodage, et de rendre le flux moins sensible aux erreurs de transmission, notamment, en exploitant les statistiques de la source uniquement au décodeur. Ce résultat est la conséquence des théorèmes sur le codage de sources distribuées établis en 1973 par

Slepian et Wolf pour le codage sans perte, puis en 1976 par Wyner et Ziv pour le codage avec pertes.

Pradhan et Ramchandran ont proposé une méthode concrète permettant d'approcher les limites de Wyner et Ziv [Pradhan & al, 99], ouvrant ainsi la voie à la recherche dans ce domaine. Depuis, deux méthodes dédiées à la compression vidéo ont été proposées. Dans [Aaron & al, 02], l'auteur présente un codeur opérant dans le domaine spatial qui utilise un codec de Wyner-Ziv et génère l'information parallèle à partir des trames clés codés en intra par la partie basse du codeur. La seconde méthode [Puri & al, 02] utilise une DCT bloc où seuls les coefficients basses fréquences sont compressés par l'utilisation des syndromes via des treillis. Les travaux actuels sont essentiellement tournés vers l'utilisation du théorème de Wyner-Ziv pour la compression vidéo mais plusieurs autres domaines commencent également à être concernés par l'utilisation du DSC, comme le réseau de capteurs [Kusuma & al, 01], [Pradhan & al, 00] et [Pradhan & al, 02], le stockage des images [Zhu & al, 03], l'amélioration numérique de la transmission analogique [Pradhan & al, 01], le watermarking multimédia ou data hiding [Chou & al, 99], la construction 3D [Girod & al, 03], et la compression d'images fixes [Ozonat, 00], [Liveris & al, 02] et [Varodayan, 04]. C'est cette dernière problématique qui nous concerne, et que sera développée dans cet article.

Dans la section suivante, nous présentons la théorie du codage distribué. La section 3 sera consacrée au codage distribué des images fixes. La description du schéma global proposé pour la compression d'images fixes sera détaillée dans la section 4. La section 5 sera réservée aux résultats de l'expérimentation et enfin des conclusions feront l'objet du dernier paragraphe.

2. Théorie du codage distribué

Le codage de sources distribuées fait référence au codage d'au moins deux sources aléatoires corrélées **X** et **Y**, codées par un codeur propre à chaque source et ne communiquant pas avec les autres codeurs, mais décodées par un décodeur conjoint (Figure 1).

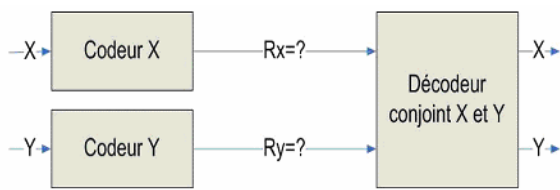


Figure 1. X et Y sont codées séparément mais décodées conjointement.

Le théorème de Slepian Wolf [Slepian & al, 73] définit les relations théoriques suivantes :

$$R_x + R_y \geq H(X,Y)$$

$$R_x \geq H(X/Y)$$

$$R_y \geq H(Y/X)$$

où R_x , R_y représentent respectivement le débit binaire de la Source X et Y.

Il est donc théoriquement possible que la somme des débits R_x+R_y atteigne l'entropie conjointe $H(X,Y)$ (Figure 3), grâce au modèle de « l'information parallèle présente uniquement au décodeur » (Figure 2) qui est un cas particulier du codage de sources distribuées présenté dans la figure 1.

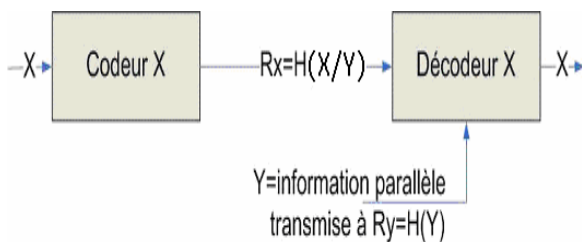


Figure 2. Compression d'une source aléatoire X en utilisant les statistiques communes à X et à Y, dans le cas où Y est uniquement disponible au décodeur.

L'entropie conjointe peut théoriquement être atteinte dans certains cas :

- * Tout d'abord ceux correspondant aux points A et B. Ainsi, en A, si l'information parallèle Y connue au décodeur (Figure. 2) est codée au débit $R_y = H(Y)$ alors il est possible de coder X au débit de $R_x = H(X/Y)$. De même en B, en inversant X et Y.
- * Un autre cas plus intéressant, est celui d'une

configuration intermédiaire représentée sur la figure 3 par le segment oblique reliant les points A et B. Tous les points de ce segment peuvent être atteints si l'on effectue un codage symétrique, où les codeurs jouent tour à tour le rôle de maître et d'esclave. On aura donc une souplesse dans l'utilisation d'un débit binaire de X compris entre $H(X/Y)$ et $H(X)$ et celui de Y compris entre $H(Y/X)$ et $H(Y)$ de telle sorte que la somme de ces débits soit égale à l'entropie conjointe $H(X, Y)$ [Pradhan & al, 00].

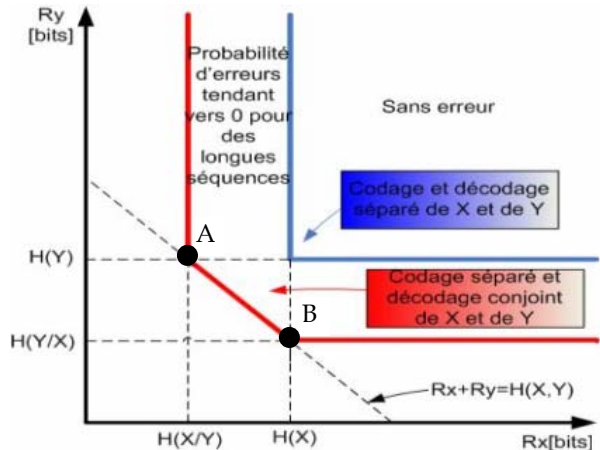


Figure 3. Débits théoriquement atteignables d'après le théorème de SlepianWolf pour la compression distribuée de deux séquences corrélées X et Y.

3. Codage distribué des images fixes

Peu de travaux existent dans la littérature sur le codage distribué des images fixes. On peut citer [Liveris & al, 02], [Varodayan, 04] qui ont utilisé un turbocodeur, de même que [Ozonat,00] dont l'approche se base sur la notion des cosets. Toutefois, certaines hypothèses émises nous semblent inadaptées au cadre d'étude. Ainsi, l'approche proposée a été validée dans le cadre du codage de sources distribuées par l'application du théorème de Slepian et Wolf du codage sans perte, mais il y a toujours un pourcentage non nul d'erreurs de décodage. En effet si $pixel_{InformationParallèle} = pixel_{ImageOriginale} \pm D/2$ (D est la distance maximale entre l'image originale et l'information parallèle) alors il y a une ambiguïté au décodage entre deux valeurs possibles et le pixel risque d'être mal décodé. Il est plus judicieux de choisir D au codeur tel que (Image Originale - Information Parallèle) appartient à l'intervalle $]-D/2;D/2[$ et non pas à $[-D/2;D/2]$. De plus, l'information parallèle est aussi volumineuse que l'image originale en terme de bits à transmettre. Cela rend l'approche totalement inintéressante si on considère que l'information parallèle est générée puis transmise par le codeur. La méthode se place toujours dans le scénario où l'information parallèle est présente directement au décodeur, et ne fait que comparer le volume d'information à transmettre, $H_{DSC}(X)$ (entropie de la source X en utilisant le codage des sources distribuées), par rapport au volume qu'aurait à transmettre un codeur conjoint, $H(X/Y)$. Le dernier problème est fortement lié au précédent. En effet, dans

une application de compression distribuée où l'information parallèle serait construite au codeur, le cumul des bits envoyés par les codeurs X et Y doit être inférieur à celui nécessaire pour l'envoi de l'image originale. Mais la diminution de l'entropie de l'information parallèle va entraîner une perte de la maîtrise et de la connaissance de la distance D (cf 4.3) entre l'image originale et l'information parallèle, aussi bien au codeur et au décodeur.

L'approche proposée dans ce papier va donc remédier à ces limites par la génération d'une information parallèle de taille correcte qui minimise le débit, et donc qui maximise le taux de compression.

4. Approche proposée

Le codec proposé (Figure 4) s'inspire de travaux utilisant le codage distribué pour la vidéo [Puri & al, 02], [Aaron & al,02] et pour les images fixes [Ozonat,00]. On retrouve ainsi une partie haute dédiée au codage et décodage de l'image et une partie basse dédiée à la construction de l'information parallèle (Side Information, notée SI).

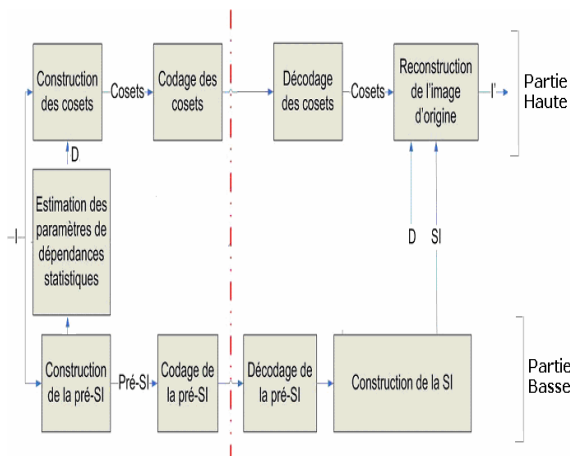


Figure 4. Schéma du codec proposé.

L'approche proposée se distingue des méthodes précédentes par deux aspects :

- * L'adaptation des dépendances statistiques entre l'image originale (I) et l'information parallèle en fonction de l'image elle-même, contrairement aux méthodes existantes, et pour lesquelles la dépendance statistique est fixée dans le codeur et le décodeur dès le départ.

- * La génération de l'information parallèle. Celle-ci se fait dans le codeur de manière à réduire la taille des données à transmettre, et augmenter la corrélation entre l'image originale et l'information parallèle. A noter que pour les codecs vidéo, l'information parallèle est générée par la partie basse du décodeur à partir des trames adjacentes à la trame de Wyner-Ziv, transmises par le canal de la partie basse du codeur. Ainsi, le codeur partie basse envoie les vraies valeurs des pixels (Figure 5).

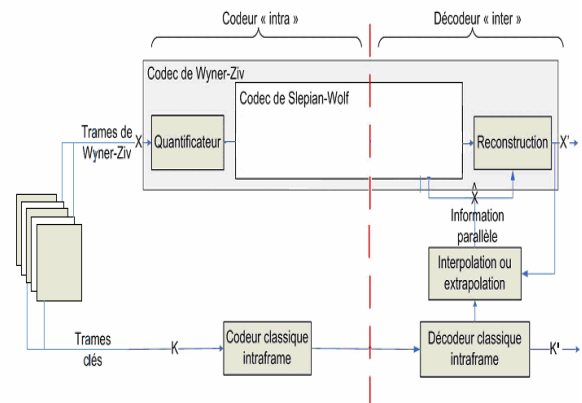


Figure 5. Génération de l'information parallèle à partir des trames adjacentes à la trame de Wyner-Ziv dans le codec proposé par Stanford.

Dans ce papier, deux méthodes seront utilisées pour la génération de l'information parallèle, soit par sous échantillonnage, soit par décomposition en sous bandes. Afin d'avoir une vue d'ensemble du fonctionnement de l'approche proposée, les différents blocs du codec sont détaillés ci-après (Figure. 4).

4.1 Partie basse du codec

La partie basse du décodeur génère l'information parallèle par interpolation, à partir de données extraites de l'image originale, puis codées et transmises par la partie basse du codeur. Ces données seront désignées par la suite par "pré-information parallèle". Deux méthodes de construction de la pré-information parallèle ont été utilisées :

- * La première envoie de « vrais pixels » : l'image originale va subir un sous échantillonnage de facteur n permettant de réduire le nombre de lignes et de colonnes. L'information parallèle est construite au décodeur par interpolation bidimensionnelle bilinéaire.

- * La seconde méthode fait appel à la décomposition en ondelettes : l'image originale est ainsi décomposée en sous bande au niveau $-n$, en utilisant le filtre 7/9 [Antonini & al, 92]. Nous conservons le $(1/4)^{n^{\text{ème}}}$ de l'image obtenue comme pré-information parallèle. L'information parallèle est reconstruite dans le décodeur par la recombinaison de l'image en utilisant le même filtre 7/9 et en annulant le reste des coefficients.

Un codage de Huffman est utilisé pour la compression de la pré-information parallèle et la minimisation de son coût de transmission.

Les premiers résultats expérimentaux montrent que les deux méthodes donnent des performances similaires. Or l'approche par sous échantillonnage, de part sa simplicité, semble plus adaptée au schéma de transmission visé.

4.2 Partie haute du codec

Il s'agit à ce niveau de transmettre l'information susceptible de lever l'ambiguïté de l'information parallèle, pour retrouver l'image originale. Notre approche fait appel aux cosets, initialement introduits

par [Ozonat, 00]. L'idée ici d'utiliser les techniques de codage canal, et de n'envoyer que les syndromes pour recouvrir l'information originale.

4.2.1 Fabrication des cosets

N pixels vont être codés ensemble pour former un « coset » de longueur N. A l'intérieur de celui-ci, chaque pixel est codé par un syndrome qui sera décrit par un nombre réduit de bits. Tous les cosets ont la même taille N. Sans autre information que les syndromes, le décodeur est incapable de retrouver la valeur initiale du pixel, un même syndrome pouvant correspondre à plusieurs valeurs de pixel possibles. C'est donc l'information parallèle qui va permettre au décodeur de lever l'ambiguïté en estimant le niveau de gris (NG) le plus vraisemblable. Pour faciliter l'estimation au décodeur, deux conditions doivent être réunies :

- * Un même syndrome doit coder des NG très éloignés les uns des autres,
- * La valeur d'un pixel dans l'information parallèle doit être la plus proche possible de la valeur du pixel correspondant dans l'image originale.

La première condition est remplie par le codage des pixels en syndromes. En effet, connaissant la distance maximale D entre image originale et SI au codeur, tous les pixels sont codés en modulo D, sur $\log_2(D)$ bits. La seconde est directement liée à la méthode choisie pour générer l'information parallèle SI.

4.2.2 Codage des cosets

Après calcul, les cosets sont codés par le codeur de Huffman puis transmis vers le décodeur. C'est au cours de cette étape que l'on va gagner en débit sur les N - 1 derniers pixels de chaque coset, car les valeurs stockées vont être majoritairement très faibles.

4.2.3 Décodage des cosets et reconstruction de l'image originale :

Le décodeur reconstruit les cosets originaux à partir des cosets codés. Ensuite le bloc «reconstruction de l'image» se sert de toute l'information disponible au décodeur, à savoir les cosets, l'information parallèle et la connaissance de D (la dépendance statistique entre les deux sources) pour reconstruire l'image originale.

L'algorithme de reconstruction de l'image d'origine à partir des cosets s'organise en deux temps :

- * Le décodage de la valeur correspondant au premier pixel d'un coset est effectué en prenant parmi l'ensemble des valeurs possibles des NG donnant ce modulo, le NG le plus proche de la valeur du pixel correspondant dans SI. L'erreur maximale entre I et SI étant connue ($\pm D/2$), il n'y a aucune ambiguïté possible.

- * Le décodage des N-1 syndromes suivants de chaque coset est réalisé de la même manière en prenant parmi l'ensemble des valeurs possibles de NG, qui vérifient que la différence avec le pixel en avant donne ce modulo, le NG le plus proche de la valeur du pixel correspondant dans SI.

4.3 Estimation du paramètre de dépendances statistiques :

L'image originale I et l'information parallèle SI sont utilisées pour dresser l'histogramme de différences entre elles. Ensuite, la distance maximale entre I et SI, notée D, est calculée pour tous les pixels. Cette distance va être utilisée pour la fabrication des cosets comme cela est détaillé dans 4.2.1.

5. Expérimentation

Les tests ont été réalisés sur les six images (Barbara, Cameraman, Lena, Baboon, Peppers, Aerial), utilisées comme références dans le domaine de l'imagerie, car elles sont représentatives d'une certaine diversité spatiale et fréquentielle. Deux paramètres ont été utilisés dans l'analyse de l'évolution du taux de compression, à savoir, le nombre de décomposition en sous bandes ou le facteur de sous échantillonnage de l'image dans la partie basse du codeur pour la génération de l'information parallèle et la taille des cosets.

La figure 6 illustre l'évolution du taux de compression obtenu par l'utilisation des deux méthodes de construction de l'information parallèle, en fonction du facteur de sous échantillonnage ou du niveau de décomposition en sous bandes. Le taux de compression augmente puis il se stabilise à partir du niveau 5 étant donné qu'à partir de ce niveau, la taille de la pré-information devient tellement négligeable par rapport à la taille de l'image originale que le taux de compression reste le même. Le niveau 5 semble donc suffisant pour la décomposition ou le sous échantillonnage de l'image afin de minimiser le coût de calcul dans la génération de l'information parallèle et de maximiser la simplicité de notre codeur.

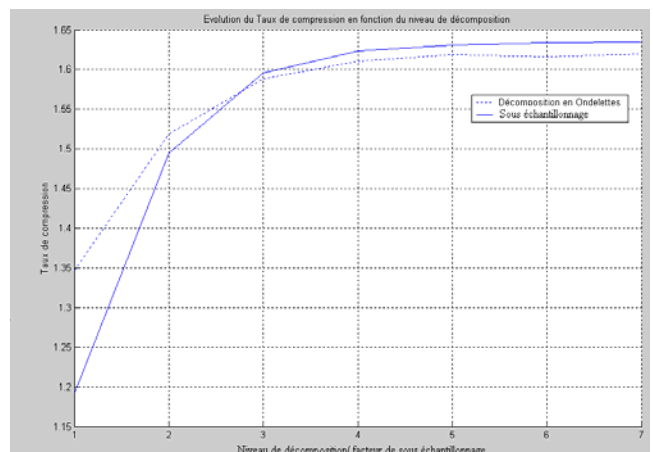


Figure 6. Evolution du taux de compression en fonction du nombre de décomposition (Lena).

La figure 7 présente l'évolution du taux de compression en fonction de la taille du coset par l'utilisation des deux méthodes de génération de l'information parallèle pour un niveau de décomposition et un facteur de sous échantillonnage égale à 7 pour l'image de barbara. Pour les deux méthodes, le taux de compression augmente en

fonction de la taille du coset et ceci reflète l'effet de l'exploitation de la redondance intra dans l'image lors de l'augmentation de la taille des cosets, puisqu'on se base sur la différence entre les pixels dans leur fabrication.

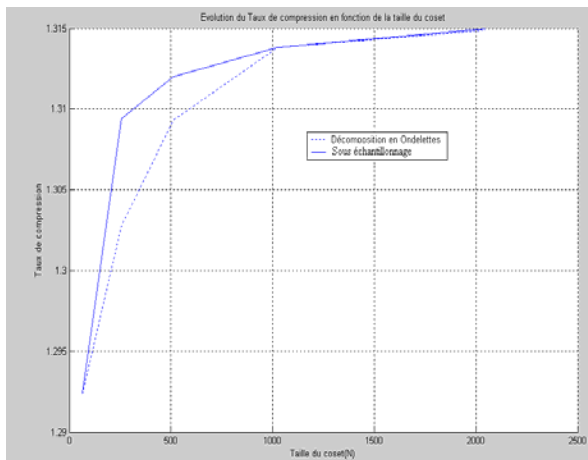


Figure 7. Evolution du taux de compression en fonction de la taille des cosets (barbara).

Pour évaluer la performance de notre méthode de compression sans perte, en utilisant les deux types de la génération de l'information parallèle, nous avons effectué une étude comparative de notre approche (Codage Distribué), avec d'autres techniques de compression sans perte existantes : JPEG sans perte (LJPG) [Wallace & al, 91], le codage EZW [Shapiro, 93], le codage SPHIT (Set Partitioning into Hierarchical Trees coder) [Said & al, 96] et le codage BTPC (Binary Tree Predictive Coding) [Robinson & al, 97] (Table1).

Image	LJPG	Codage Distribué (décomposition en Ondelettes)	Codage Distribué (Sous échantillonnage)	EZW	BTPC	SPHIT
Aerial 256x256	4.40	5.68	5.68	4.22	4.56	4.11
Barbara 512x512	5.53	6.07	6.07	4.75	5.40	4.58
Camera n256x256	5.03	5.18	5.13	4.73	4.98	4.51
Lena, 512x512	4.70	4.89	4.88	4.31	4.51	4.20
Baboon, 512x512	6.58	6.17	6.16	6.09	6.51	5.96
Peppers 512x512	5.19	4.91	4.90	4.68	4.84	4.62

Table1. Débit binaire obtenu par les différentes méthodes de compression.

Les deux méthodes de génération de l'information parallèle donnent des résultats similaires, mais l'approche par sous échantillonnage de l'image sera privilégiée car elle offre une plus grande simplicité se trouve plus adaptée. Les valeurs des taux de compression trouvées pour le codage distribué sont du même ordre de grandeur que celles issues des méthodes classiques, sans toutefois pouvoir être meilleures.

6. Conclusions

Dans cet article, une nouvelle approche pour le codage distribué des images fixes a été présentée.

L'approche proposée innove par rapport aux travaux référencés sur deux points. D'une part la génération de l'information parallèle où deux méthodes de construction sont évaluées, à savoir la décomposition en ondelettes et le sous échantillonnage de l'image. Et d'autre part, par l'adaptation des dépendances statistiques entre l'image originale et l'information parallèle en fonction de l'image.

Les performances de l'approche proposée du point de vue du taux de compression sont encore limitées par rapport au standards du marché, en revanche, la méthode offre la simplicité au codeur, qui pourra être indispensable dans des applications de transmissions via un réseau sans fil à faible débit.

Ce travail constitue une première ébauche et plusieurs pistes d'améliorations sont possibles : utilisation de cosets de taille variable dépendant du contenu de l'image et reflétant les zones homogènes de celle-ci, utilisation d'une distance spécifique pour chaque coset, amélioration de la corrélation entre l'information parallèle et l'image originale. Nos travaux actuels explorent ces différentes voies.

7. Références

- [Aaron & al, 02] A. Aaron, R. Zhang et B. Girod, « Wynerziv coding of motion video », in Proc. Asilomar Conference on Signals and Systems, Pacific Grove, CA, nov. 2002.
- [Antonini & al, 92] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, « Image coding using wavelet transform », IEEE Transactions on Image Processing", Vol. pp. 205-220, 1992.
- [Chou & al, 99] J. Chou, S. S. Pradhan et K. Ramchandran, « Systematic lossy forward error protection for video waveforms », in proc. 33rd Asilomar Conference on Signals, Systems and Computers, p. 1503–1507, nov. 1999.
- [Girod & al, 03] B. Girod, C-L. Chang, P. Ramanathan, and X. Zhu « light field compression using disparity-compensated lifting » in Proc. IEEE International Conference on Acoustics, Speech and Signal Processing, vol.IV, Hong Kong, China, Apr.2003, pp. 761-764
- [Kusuma & al, 01] J. Kusuma, L. Doherty et K. Ramchandran, « Distributed compression for sensor networks », in Proc. IEEE International Conference on Image Processing (ICIP), Thessaloniki, Greece, vol. 1, p. 82–85, oct. 2001.
- [Liveris & al, 02] A. Liveris, Z. Xiong, et C. Georgiades « A Distributed Source Coding Technique for Correlated Images Using Turbo-Codes », IEEE Communications Letters, Vol. 6, N° 9, Septembre 2002.
- [Ozonat, 00] K. Ozonat, "Lossless distributed source coding for highly correlated still images,"

Electrical Eng. Dept., Stanford Univ., Stanford, CA, Tech. Rep., 2000.

- [Pradhan & al, 99] S. S. Pradhan et K. Ramchandran, « *Distributed source coding using syndromes (DISCUS) : Design and construction* », in proc. IEEE Data Compression Conference, Snowbird, UT, p. 158–167, mars 1999.
- [Pradhan & al, 00] S. S. Pradhan et K. Ramchandran, « *Distributed source coding : Symmetric rates and applications to sensor networks* », in proc. IEEE Data Compression Conference, Snowbird, UT, p. 363–372, mars 2000.
- [Pradhan & al, 01] S. S. Pradhan et K. Ramchandran, « *Enhancing analog image transmission systems using digital side information : a new wavelet based image coding paradigm* », in Proc. IEEE Data Compression Conference, Snowbird, UT, p. 63–72, mars 2001.
- [Pradhan & al, 02] S. S. Pradhan, J. Kusuma et K. Ramchandran, « *Distributed compression in a dense microsensor network* », IEEE Signal Processing Magazine, vol. 19, p. 51–60, mars 2002.
- [Puri & al, 02] R. Puri et K. Ramchandran, « *Prism : A new robust video coding architecture based on distributed compression principles* », in Proc. Allerton Conference on Communication, Control, and Computing, Allerton, IL, oct. 2002.
- [Robinson & al, 97] J. A. Robinson, « *Efficient general-purpose image compression with binary tree predictive coding* », IEEE Trans. image Processing, 6, 1997, 601-608.
- [Said & al, 96] A. Said et W. Pearlman, « *A new fast and efficient image codec based on set partitioning in hierarchical trees* », IEEE Trans. on Circuits and Systems Video Technology, 6, 1996, 243-250.
- [Shapiro, 93] J. M. Shapiro, « *Embedded image coding using zerotrees of wavelet coefficients* », IEEE Trans. Signal Processing, 41, 1993, 3445-3462.
- [Slepian & al, 73] D. Slepian et J. Wolf, « *Noiseless coding of correlated information sources* », IEEE Transactions on Information Theory, vol. IT19, p. 471–480, juil. 1973.
- [Varodayan, 04] D. P. VARODAYAN, « *Wynerziv coding of still images with rate estimation at the encoder* », EE398B Project Final Report, mai 2004.
- [Wallace & al, 91] G. K. Wallace, « *The JPEG still picture compression standard* », Comm. ACM, 34, 1991, 30-44.
- [Zhu & al, 03] X. Zhu, A. Aaron et B. Girod, « *Distributed compression for large camera arrays* », in Proc. IEEE Workshop on Statistical Signal Processing, St. Louis, MO, p. 21–24, sept. 2003.