

Interleaved S+P : Décomposition pyramidale entrelacée à contexte de prédiction enrichi

Marie BABEL, Olivier DÉFORGES, Joseph RONSIN

IETR CNRS UMR 6164 Groupe Image / INSA de Rennes
20, av. des Buttes de Coësmes, CS 14315, 35043 Rennes, France
{mbabel, odeforge, ronsin}@insa-rennes.fr

Résumé – La progressivité et les fonctionnalités du type codage par Région d’Intérêt constituent des composantes essentielles d’un codeur d’images efficace. L’état de l’art, dans le contexte de la compression sans perte, est représenté par les techniques S+P et CALIC. Dans un contexte pyramidal, l’algorithme ici proposé, l’Interleaved S+P, non seulement affiche des performances supérieures à ces méthodes, mais également procure les fonctionnalités avancées souhaitées. Une décomposition pyramidale originale, associée à un schéma MICD et fondée sur la méthode LAR (Locally Adaptive Resolution) est décrite dans cet article. Cette solution de codage met en œuvre la transformée en S de telle sorte que, pour chaque phase d’estimation, un contexte de prédiction enrichi soit disponible. Pour ce faire, la reconstruction sans perte d’une image est effectuée en deux étapes successives, où la première fournit une image LAR basse résolution de très bonne qualité.

Abstract – Scalability and functionalities such as Region of Interest coding constitute essential components of efficient image coder. In lossless compression context, state-of-the-art techniques are S+P and CALIC ones. We propose in a pyramidal context a new algorithm, the Interleaved S+P, that achieves better results than the previous methods, and provides also the desired functionalities. An original pyramidal decomposition used in association with a DPCM scheme and based on the LAR method (Locally Adaptive Resolution) is described. This solution adapts the S-Transform so that a refined prediction context is available for each estimation process. The lossless reconstruction of the image is thus realized in two successive steps, where the first one provides a LAR low-resolution image of very good quality.

1 Introduction

Du fait de la généralisation de l’utilisation des réseaux à bas-débits du type Internet, les méthodes de compression d’images tendent à satisfaire au besoin de scalabilité de la source. En effet, des applications spécifiques, telle la télémédecine, nécessitent une transmission progressive des images en termes de résolution et de qualité, et ce au moyen d’un unique codeur.

De ce fait, bien que les techniques opérant en pleine résolution de l’image soient les plus répandues dans le domaine de la compression sans perte (CALIC [1]), les codeurs à structure pyramidale continuent à faire l’objet de recherches. Si l’utilisation conjointe de schémas MICD et de transformées réversibles est rarement mise en œuvre au sein d’un même procédé de codage, l’algorithme S+P [2] propose une décomposition multirésolution de l’image, associant la transformée en S et des étapes de prédiction simples.

La méthode LAR (Locally Adaptive Resolution), fondée sur une représentation de l’image à taille de blocs variable, définit une technique efficace de compression avec pertes [3]. Nous avons introduit dans [4, 5] une approche pyramidale à redondance minimale (le LAR-APP), où les étapes de prédiction successives tirent à la fois parti d’un voisinage à 360° et d’une modélisation de contexte implicite. Cet article propose une solution originale de codage d’images : profitant de la description pyramidale du LAR-APP, une implantation particulière de la transformée en S conduit à la définition de l’Interleaved S+P.

Après avoir évoqué les particularités du LAR simple (§2), nous donnons la méthode de construction entrelacée de la pyramide Interleaved S+P (§3). La décomposition pyramidale est par la suite décrite en association avec les estimateurs mis

en œuvre (§4). Enfin, les résultats de l’Interleaved S+P sont présentés en section 5.

2 La méthode LAR pour les images en niveaux de gris

La méthode LAR (Locally Adaptive Resolution), définie en premier lieu pour la compression à bas débits d’images en niveaux de gris, est constituée d’un codec à deux couches : le codeur spatial, fournissant une image basse résolution, et le codeur spectral traitant de la texture. L’originalité de l’algorithme repose sur le principe suivant : la résolution locale, *i.e.* la taille des pixels, peut varier en fonction de l’activité locale. Ainsi, les pixels de petite taille se situent naturellement sur les contours de l’image, alors que les zones uniformes sont représentées par un bloc de grande taille (figure 1).

Cette grille de résolution spécifique, s’appuyant sur une topologie Quadtree déterminée par un gradient morphologique, a été en particulier exploitée dans l’élaboration d’une représentation en régions auto-extractible [6]. L’Interleaved S+P présenté dans cet article ajoute la propriété de scalabilité de codage de la source.

3 Construction de la pyramide Interleaved S+P

Dans cette section, les principes généraux de l’algorithme sont tout d’abord introduits, avant d’exposer la construction de



FIG. 1 – (a) Image originale (b) Grille (0.032bpp). (c) Image basse résolution (0.4bpp) - Première descente de pyramide. (d) Image post-traitée.

la pyramide. Par la suite, nous notons z_0 et z_1 les coefficients transformés d'un couple (u_0, u_1) tels que

$$\begin{aligned} z_0 &= \lfloor (u_0 + u_1)/2 \rfloor, \\ z_1 &= u_1 - u_0, \end{aligned} \quad (1)$$

avec $\lfloor \cdot \rfloor$ opérateur d'arrondi à l'entier immédiatement inférieur.

3.1 Principes généraux

La décomposition pyramidale du LAR est issue de l'adaptation du prédicteur de Wu décrit dans [7]. Pour une image pleine résolution, les erreurs de prédiction sont encodées par trois échantillonnages successifs de l'image originale. Ainsi, pour un pixel donné, l'estimateur peut s'appuyer sur un contexte de configuration spatiale dite à 360°, de telle sorte que l'entropie résultante soit considérablement réduite.

La première phase de l'Interleaved S+P consiste en l'application de la transformée en S 1D sur les deux vecteurs constitué chacun des deux pixels diagonalement adjacents d'un bloc 2×2 donné (figure 2). Par la suite, nous notons z_i^k , avec $i \in \{0, 1\}$ et $k \in \{1, 2, 3\}$, les coefficients z_i codés par la $k^{\text{ème}}$ passe de l'algorithme.

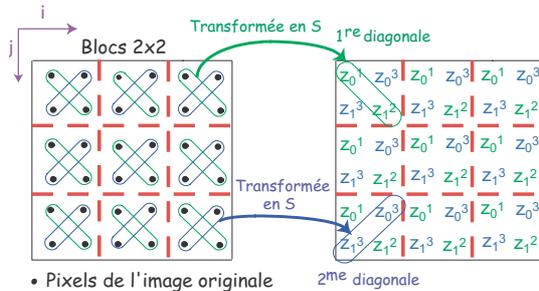


FIG. 2 – Application de la transformée en S

La première passe encode, par un MICD classique, l'image sous-échantillonnée formée par l'ensemble des valeurs de moyenne de deux pixels adjacents (coefficients z_0^1 , cf figure 2).

Par la suite, la seconde passe estime la valeur des coefficients z_1^2 afin de pouvoir reconstruire les deux pixels de la première diagonale. Pour cette seconde passe, le voisinage à 360° est constitué des valeurs déjà prédites (voisinage causal), ainsi que de l'ensemble des moyennes de diagonale issues de la première passe. Enfin, la troisième passe traite du reste de l'image originale, à savoir les coefficients transformés z_0^3 et z_1^3 : de nouveau, le contexte riche disponible permet d'estimer efficacement leur valeur.

3.2 Construction de la pyramide - Entrelacement

Soit I l'image originale de taille $N_x \times N_y$. La représentation multirésolution d'une image est donnée par l'ensemble $\{Y_l\}_{l=0}^{l_{max}}$, où l_{max} désigne le sommet de la pyramide et $l = 0$ la pleine résolution. Par extension de la méthode de Wu, quatre blocs $\frac{N}{2} \times \frac{N}{2}$ sont représentés au niveau supérieur par un bloc $N \times N$ de valeur égale à la moyenne des deux blocs sous-jacents de la première diagonale (première pyramide en S, figure 3) :

$$\begin{cases} l = 0, & Y_0(i, j) = I(i, j); \\ l > 0, & Y_l(i, j) = \left\lfloor \frac{Y_{l-1}(2i, 2j) + Y_{l-1}(2i+1, 2j+1)}{2} \right\rfloor, \end{cases} \quad (2)$$

avec $0 \leq i \leq N_x^l$, $0 \leq j \leq N_y^l$, où $N_x^l = N_x/l$ et $N_y^l = N_y/l$.

L'originalité de l'algorithme Interleaved S+P tient dans le fait qu'il est possible de **décrire la construction de la pyramide entière au moyen de deux pyramides entrelacées** (figure 3). La transformation de la deuxième diagonale d'un bloc 2×2 donné peut en effet aussi être vue comme la réalisation d'une seconde pyramide en S : l'estimation de la valeur d'un pixel dépend alors de celles présentes au niveau inférieur de la première pyramide en S.

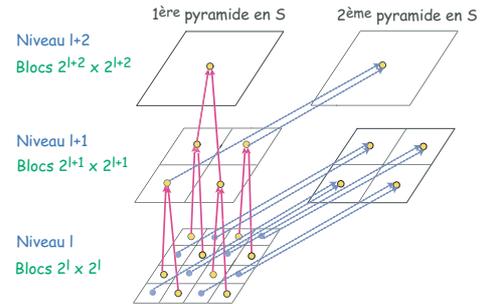


FIG. 3 – Construction de la pyramide par entrelacement de deux pyramides en S

4 Décomposition de la pyramide

Le processus de décomposition pyramidale de l'Interleaved S+P résulte de l'extension de la méthode de prédiction précédemment décrite à la pyramide entière. Dans cette section sont décrites l'ensemble des phases de prédiction.

4.1 Principes de la décomposition pyramidale

La reconstruction d'un bloc 2×2 situé à un niveau donné l est réalisée en deux passes. Tout d'abord, la transformée en S

s'applique sur les pixels $Y_l(2i, 2j)$ et $Y_l(2i + 1, 2j + 1)$:

$$z_0^{l,1}(2i, 2j) = \left\lfloor \frac{Y_l(2i, 2j) + Y_l(2i + 1, 2j + 1)}{2} \right\rfloor, \quad (3)$$

$$z_1^{l,2}(2i + 1, 2j + 1) = Y_l(2i, 2j) - Y_l(2i + 1, 2j + 1).$$

Nous remarquons immédiatement que le coefficient $z_0^{l,1}(2i, 2j)$ est égal à la valeur du pixel reconstruit au niveau $l + 1$ de la pyramide :

$$z_0^{l,1}(2i, 2j) = Y_{l+1}(i, j). \quad (4)$$

De façon similaire, les pixels de la seconde diagonale sont transformés selon :

$$z_0^{l,3}(2i + 1, 2j) = \left\lfloor \frac{Y_l(2i, 2j + 1) + Y_l(2i + 1, 2j)}{2} \right\rfloor, \quad (5)$$

$$z_1^{l,3}(2i, 2j + 1) = Y_l(2i + 1, 2j) - Y_l(2i, 2j + 1).$$

Comme mentionné au paragraphe 2, la pyramide est alors décomposée selon deux descentes successives. Si la première descente vise à reconstruire l'image des blocs, la seconde encode l'image de texture : la grille LAR y est alors exploitée (voir figure 4). Pour ce faire, une image notée *Siz* contenant l'information de taille de blocs est élaborée à partir de notre partitionnement Quadtree. La figure 1.c donne l'image des blocs du LAR, résultat de la première descente partielle de notre pyramide.

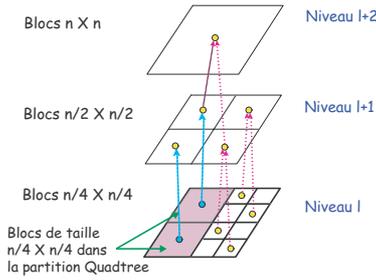


FIG. 4 – Décomposition : extraction de l'image LAR des blocs

4.2 Sommet de la pyramide - MICD classique

Au niveau le plus haut de la pyramide, les valeurs des pixels $Y_{l_{max}}(i, j) = z_0^{l_{max}-1,1}(2i, 2j)$ subissent une unique phase d'estimation. Le prédicteur classique MED, utilisé dans la méthode de compression method LOCO-I [8], est alors mis en œuvre.

4.3 Définition des prédicteurs - Image des blocs

La reconstruction de l'image LAR basse résolution (image des blocs) induit, pour un niveau de la pyramide donné, le traitement des seuls pixels de taille $Siz(x \times l, y \times l) \leq 2^l$. Inversement, si $Siz(x \times l, y \times l) > 2^l$, la valeur du bloc est simplement recopiée de telle sorte que $Y_l(2i, 2j) = Y_l(2i + 1, 2j) = Y_l(2i, 2j + 1) = Y_l(2i + 1, 2j + 1) = Y_{l+1}(i, j)$.

Nous notons par la suite \check{z} la valeur estimée du coefficient transformé z .

Première pyramide en S. La valeur du coefficient $z_0^{l,1}$ étant déjà connue (niveau supérieur de la première pyramide en S), il nous suffit de procéder à l'estimation du coefficient transformé $z_1^{l,2}$, valeur gradient de la première diagonale d'un bloc 2×2

donné. Dans l'objectif d'obtenir l'image des blocs (zones de forte activité), nous avons remarqué qu'une estimation linéaire des coefficients demeure la plus efficace :

$$\begin{aligned} z_1^{l,2}(2i + 1, 2j + 1) = 2.1 \left[0.9Y_{l+1}(i, j) + \frac{1}{6} \left(Y_l(2i + 1, 2j - 1) \right. \right. \\ \left. \left. + Y_l(2i - 1, 2j - 1) + Y_l(2i - 1, 2j + 1) \right) \right. \\ \left. - 0.05 \left(Y_l(2i, 2j - 2) + Y_l(2i - 2, 2j) \right) \right. \\ \left. - 0.15 \left(Y_{l+1}(i, j + 1) + Y_{l+1}(i + 1, j) \right) - Y_{l+1}(i, j) \right]. \end{aligned} \quad (6)$$

Deuxième pyramide en S. La troisième passe de l'algorithme de Wu adapté à notre problématique encode les coefficients transformés issus de la deuxième diagonale d'un bloc 2×2 , appartenant de fait à la seconde pyramide en S. L'estimation de $z_0^{l,3}$ fait intervenir une prédiction inter- et intra-niveau :

$$\begin{aligned} z_0^{l,3}(2i + 1, 2j) = \alpha_0 \frac{1}{4} \left(Y_l(2i - 1, 2j + 1) + Y_l(2i, 2j + 2) \right. \\ \left. + Y_l(2i + 2, 2j) + Y_l(2i + 1, 2j - 1) \right) \\ \left. + \beta_0 z_0^{l,1}(2i, 2j), \end{aligned} \quad (7)$$

où $(\alpha_0, \beta_0) = (0.25, 0.75)$, et $z_0^{l,1}$ représente la valeur reconstruite du coefficient $z_0^{l,1}$. La valeur estimée $z_1^{l,3}(2i, 2j)$ se calcule selon :

$$\begin{aligned} z_1^{l,3}(2i, 2j + 1) = \alpha_1 \left(Y_l(2i - 1, 2j + 1) + Y_l(2i, 2j + 2) \right. \\ \left. - Y_l(2i + 1, 2j - 1) - Y_l(2i + 2, 2j) \right) \\ - \beta_1 \left(Y_l(2i - 1, 2j) + Y_l(2i - 1, 2j + 2) \right. \\ \left. - Y_l(2i, 2j - 1) - \check{Y}_l(2i, 2j + 1) \right). \end{aligned} \quad (8)$$

avec $(\alpha_1, \beta_1) = (3/8, 1/8)$. $\check{Y}_l(2i, 2j + 1)$ correspond au prédicteur de Wu [7] de la troisième passe appliquée au pixel $Y_l(2i, 2j + 1)$.

4.4 Informations de texture et prédicteurs

La réalisation de la deuxième phase de décomposition de la pyramide permet la reconstruction de la texture. Rappelons qu'à chaque niveau de la pyramide, seuls les pixels de taille $Siz(x \times l, y \times l) > 2^l$ subissent une étape de prédiction, les autres ayant été traités lors du premier processus de descente. Il est de nouveau possible de distinguer le cas des valeurs issues de la première pyramide en S (première diagonale), de celles de la deuxième pyramide en S (seconde diagonale).

Première pyramide en S. L'information de texture se caractérise par une faible activité locale. Les valeurs de gradient $z_1^{l,2}$ sont difficilement appréciables à travers un processus linéaire. C'est pourquoi nous nous appuyons sur des opérateurs non linéaires de type médian.

Soit $m_e(u_1, u_2, \dots, u_n)$ la valeur médiane d'un ensemble (u_1, u_2, \dots, u_n) de n valeurs. La valeur estimée du coefficient

$$z_1^{l,2}(2i, 2j) \text{ situé dans une zone de texture est donnée par}$$

$$z_1^{l,2}(2i + 1, 2j + 1) = \frac{1}{4} \left(m_e(Y_i(2i - 2, 2j), Y_i(2i, 2j - 2), \right. \\ \left. Y_i(2i - 1, 2j - 1)) \right. \\ \left. + m_e(Y_{l+1}(i + 1, j), Y_{l+1}(i, j + 1), Y_{l+1}(i + 1, j + 1)) \right). \quad (9)$$

Deuxième pyramide en S. Le contexte intra-niveau disponible pour l'estimation des coefficients transformés $z_0^{l,3}$ et $z_1^{l,3}$ s'avère suffisant pour une prédiction simple linéaire. Ainsi, les valeurs de moyennes $z_0^{l,3}$ sont traitées par l'application de la relation 7, où $(\alpha_0, \beta_0) = (0.37, 0.63)$. Quant aux coefficients $z_1^{l,3}$, leur estimation est obtenue après application de l'équation 8, avec le couple de valeurs $(\alpha_1, \beta_1) = (\frac{1}{4}, 0)$.

5 Résultats - Modélisation de contexte implicite

L'état de l'art des méthodes de compression d'images sans perte désigne le codeur CALIC [1] comme étant l'un des plus efficaces en termes d'entropie du premier ordre (taux de compression) ainsi que de complexité. L'algorithme S+P [2] constitue quant à lui une référence en matière de schéma de compression progressif. Nous avons donc procédé à des tests comparatifs avec ces précédents schémas de codage. Dans cette étude, l'encodage entropique (par codage arithmétique adaptatif) n'est pas implanté. Le tableau 1 expose les entropies du premier ordre obtenues après application de l'Interleaved S+P, CALIC et S+P. Notre méthode dépasse largement le simple S+P. De plus, les résultats de l'Interleaved S+P équivalent en moyenne à ceux de CALIC ; pour la majorité des images de cette série de mesures, notre algorithme surpasse même ce schéma de codage.

Grâce à notre sous-échantillonnage non uniforme, l'entropie est considérablement réduite, comme l'indique le tableau 1. En effet, l'entropie globale peut être diminuée lorsque plusieurs classes de symboles suivant la même loi peuvent être isolées. Notre méthode implique une séparation immédiate des lois : pour chaque niveau de la pyramide, nous distinguons les erreurs de prédiction issues de l'image LAR des blocs de l'image de texture, la passe 2 de la passe 3, les coefficients z_0 des coefficients z_1 . De ce fait, grâce à notre représentation basée contenu, une modélisation de contexte implicite est réalisée.

Image	Entropie (bpp)			
	Interleaved S+P - Sans partition	Interleaved S+P	CALIC	S+P
Barb2	5.21	4.96	4.93	5.04
Café	5.66	5.43	5.37	5.42
Gold	4.82	4.64	4.65	4.73
Hotel	4.91	4.66	4.57	4.97
Lena	4.46	4.30	4.33	4.33
Peppers	4.74	4.57	4.58	4.67
tools	5.78	5.52	5.53	5.73
us	4.01	3.55	3.60	3.78
Moyenne	4.95	4.70	4.70	4.83

TAB. 1 – Entropie du premier ordre de la méthode Interleaved S+P (avec partition ou non), CALIC et S+P.

6 Conclusion

Cet article a présenté une technique originale de compression conjointe avec et sans perte. Fondée sur la méthode LAR, cette solution de codage utilise la transformée en S de telle sorte qu'un contexte de prédiction enrichi soit disponible pour chaque étape d'estimation. La connaissance de la structure Quadtree est suffisante pour améliorer de façon significative le taux de compression sans perte, par le regroupement et le codage séparé des informations de même nature (coefficients transformés, image LAR basse résolution, contours...). Les performances de cet algorithme Interleaved S+P dépassent significativement celle de la transformée classique S+P, et sont globalement équivalentes à CALIC. En supplément, la technique proposée permet non seulement la reconstruction progressive d'images, mais aussi fournit une carte de segmentation gratuite à la fois au codeur et au décodeur : une application immédiate en est la compression localement sans perte d'une image, par la définition d'une zone d'intérêt.

Nos recherches actuelles portent sur l'adaptation du codage LAR à la vidéo. D'autres travaux concernent l'élaboration d'une approche conjointe source-canal, en tant que solution pour la transmission sécurisée d'images médicales à travers les réseaux à faibles bandes passantes du type Internet.

Références

- [1] X. Wu and N. Memon and, "A Context-based, Adaptive, Lossless/Nearly-Lossless Coding Scheme for Continuous-Tone Images (CALIC)," *International Standards Organization working document, ISO/IEC SC29/WG 1/N256*, vol. 202, 1995.
- [2] A. Said and W. Pearlman, "Reversible Image Compression via Multiresolution Representation and Predictive Coding," in *Visual Communication and Image Processing*. SPIE, 1993, vol. 269, pp. 664-374.
- [3] O. Deforges and J. Ronsin, "Region of Interest Coding for Low Bit-Rate Image Transmission," in *Proc. ICME'2000*, July 2000, vol. 1, pp. 107 - 110.
- [4] M. Babel, O. Déforges, and J. Ronsin, "Lossless and Lossy Minimal Redundancy Pyramidal Decomposition for Scalable Image Compression Technique," in *Proc. ICASSP'03*, Hong Kong, April 6-10 2003, Conference cancelled - Invited paper in ICME 2003, vol. III, pp. 249-252.
- [5] Marie Babel, Olivier Déforges, and Joseph Ronsin, "Décomposition pyramidale à redondance minimale pour compression d'images sans perte," in *GRETSI'03*, Paris, France, Sept. 8-11 2003.
- [6] O. Deforges and J. Ronsin, "Supervised Segmentation at Low Bit Rates for Region Representation and Color Image Compression," in *Proc. ICME'2002*, Aug. 26-29 2002, vol. 1, pp. 665 - 668.
- [7] X. Wu, "Lossless Compression of Continuous-Tone Images, via Context Selection, Quantization and Modelling," *IEEE Trans. on Image Processing*, vol. 6, no. 5, pp. 656-664, May 1997.
- [8] M.J. Weinberger, G. Seroussi, and G. Sapiro, "LOCO-I : a low complexity, context-based, lossless image compression algorithm," *Proc. DCT'96*, pp. 141-150, 1996.