

DÉCOMPOSITION PYRAMIDALE A REDONDANCE MINIMALE POUR COMPRESSION D'IMAGES SANS PERTE

Marie BABEL, Olivier DÉFORGES, Joseph RONSIN

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

Résumé – Afin de répondre au besoin de progressivité des méthodes de compression, une nouvelle méthode de décomposition pyramidale associée à l'utilisation d'une forme originale de MICD à contexte enrichi a été mise en œuvre. La méthode LAR (Locally Adaptive Resolution) pour du codage avec pertes a donc été adaptée à cet effet. L'image d'erreur résultante est ensuite codée sans perte par prédiction.

Abstract – In order to answer the need of progressive compression schemes, a new method of pyramidal decomposition jointly used with an original context-refined DPCM scheme is presented. Thus the LAR (Locally Adaptive Resolution) method designed first for lossy compression has been adapted. The error image is subsequently losslessly encoded using prediction.

1 Introduction

De nombreuses applications mettent en jeu une compression sans perte des images. En général, les techniques mises en œuvre peuvent se classer en deux groupes : l'état-de-l'art fait en effet mention de l'utilisation de prédicteurs (comme la méthode CALIC [1]), ou celle des transformées réversibles. Une solution mixte est quant à elle proposée par l'algorithme S+P [2].

Si on considère l'image comme étant la superposition d'une information globale et de la texture, ainsi que le propose TMW [3], il est possible de décider à partir d'un même codeur de fournir une image compressée avec perte (première partie du message), ou sans perte (deuxième partie, correspondant à l'image d'erreur).

Le LAR (*Locally Adaptive Resolution*), très efficace pour de la compression avec pertes, se base sur une décomposition à taille de blocs variable, proposant ainsi une technique de compression avec pertes d'images fixes très efficace. Le plan s'articule autour des axes suivants : dans un premier temps, la méthode est brièvement décrite. La section 3 définit alors une décomposition pyramidale particulière apportant la progressivité spatiale dans le schéma de compression avec pertes du LAR. Enfin, la dernière partie traite du codage sans perte qui tire parti de la modélisation de contexte implicite inhérente à notre décomposition.

2 La méthode LAR

La méthode de compression LAR (*Locally Adaptive Resolution*), a été initialement conçue pour réaliser un *codage graduel avec pertes des images en niveaux de gris* [4]. L'originalité de la technique proposée réside dans un schéma de codage progressif à deux couches (codeurs spatial et spectral). Le LAR

a de plus été étendu au cas du codage basé région des images couleur. Enfin, la qualité des images basse résolution obtenues a été évaluée par tests psychovisuels qui démontrent la supériorité de notre technique sur Jpeg-2000 [5].

2.1 Le codeur spatial.

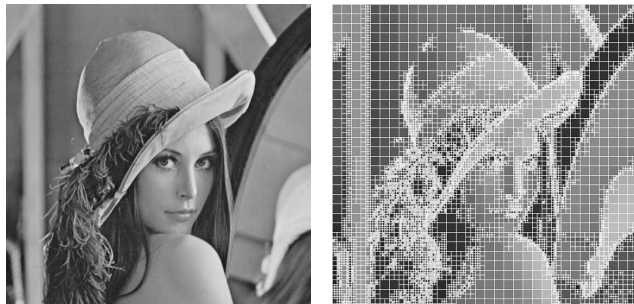
Partant du principe que la résolution de l'image (la taille des pixels) varie en fonction de l'activité locale - estimée au moyen d'un gradient morphologique -, le codeur spatial sous-échantillonne l'image sur une grille carrée non uniforme (figure 1). Une image basse résolution est obtenue en reconstruisant chaque bloc par la valeur moyenne de la luminance. La taille des blocs donne la nature du pixel : les plus petits blocs (2×2) correspondent aux pixels de contours alors que les plus grands blocs (16×16) sont situés dans les zones homogènes. De ce fait, une carte de segmentation grossière est disponible à la fois au codeur et au décodeur.

Une prédiction de type MICD (via le Gradient Adjusted Predictor) est utilisée pour coder la luminance moyenne de chaque bloc. La quantification de l'erreur s'adapte en fonction de la taille des pixels, respectant la sensibilité de l'œil humain : elle est donc grossière sur les petits blocs et fine sur les gros blocs. Enfin, l'erreur de prédiction ainsi que la taille des blocs sont compressées par un codeur arithmétique adaptatif. Les images obtenues sont post-traitées afin d'éliminer les effets de blocs présents dans les zones uniformes (interpolation adaptative basée sur la théorie de l'"optimal recovery"), définie par Darian Muresan [6].

En résumé, deux caractéristiques essentielles du codeur peuvent être dégagées :

- la technique est très efficace et très rapide pour les forts taux de compression,

- la méthode revient à simplifier l'image source (suppression de la texture locale, conservation des frontières des objets).



(a) (b)

FIG. 1 – (a) Image originale. (b) Grille de sous-échantillonnage.

2.2 Le codeur spectral.

Celui-ci constitue le deuxième étage du codeur global, et permet la reconstruction progressive d'une image de meilleure qualité. L'image des erreurs (texture locale) est traitée à travers une approche de type DCT à taille de bloc variable. La taille de ces blocs et la composante DC sont fournies par le codeur spatial. Ce codeur de base a également été étendu pour de la compression avec zone d'intérêt [4] et du codage d'image couleur basé régions [5].

3 Approche pyramidale prédictive : compression avec pertes

L'image des blocs du LAR respecte une topologie de type quad-tree. Cependant, une approche de type MICD classique s'était révélée plus performante qu'une décomposition quad-tree traditionnelle. Afin de permettre une reconstitution progressive de l'image des blocs, nous avons défini un nouveau schéma prédictif par l'intermédiaire d'une décomposition pyramidale.

3.1 L'algorithme de prédiction

Le codeur spatial du LAR initial encodait la luminance des blocs au moyen du GAP (Gradient Adjusted Predictor) [1]. La nouvelle version s'appuie désormais sur le prédicteur décrit par Wu dans [7]. Le principe est le suivant : pour la pleine résolution, l'image d'erreurs est codée à travers trois sous-échantillonnages entrelacés. Un pixel donné dispose de ce fait d'un contexte enrichi, dont la configuration spatiale est dite à 360°.

Le principe général est donc le suivant (figure 2). Dans un premier temps, on cherche à encoder l'image sous-échantillonnée formée par la moyenne de deux pixels

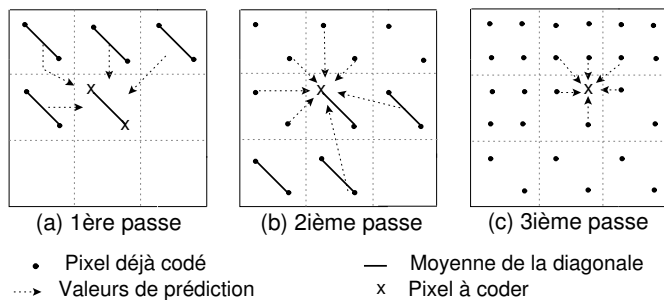


FIG. 2 – Algorithme de Wu : 3 passes and voisinage de prédiction

diagonalement adjacents à travers un MICD classique. La deuxième passe encode les valeurs des deux pixels ayant servi au calcul de la moyenne précédente. Enfin, la troisième passe permet la reconstruction de la deuxième moitié de l'image originale. La prédiction au cours des deuxième et troisième passes s'appuie sur un contexte adjacent entourant entièrement le pixel à modéliser.

3.2 Décomposition pyramidale à redondance minimale

Afin de réaliser une représentation pyramidale de l'image, l'algorithme de prédiction précédent a été adapté.

3.2.1 Construction du type Bottom-Up.

Quatre blocs $\frac{N}{2} \times \frac{N}{2}$ sont regroupés pour former un bloc $N \times N$ (figure 3), dont la valeur de la luminance est égale à la moyenne de la première diagonale. Ainsi, la valeur d'un bloc quelconque de la pyramide correspond à la moyenne de la diagonale entière de l'image pleine résolution. Le plus haut niveau de la pyramide (niveau 4) correspond aux blocs 16×16 (taille la plus grande autorisée dans la méthode LAR).

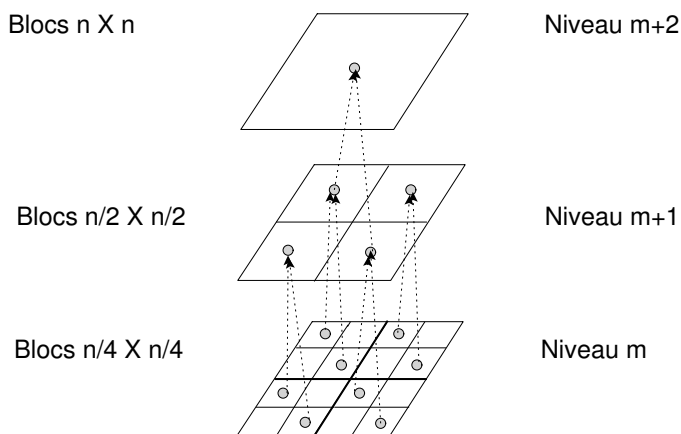


FIG. 3 – Construction *bottom-up*.

3.2.2 Décomposition Top-Down.

Considérons l'image originale de largeur W et de hauteur H . Le premier niveau traité de notre pyramide concerne

l'image de taille $W/16 \times H/16$. La première passe de l'algorithme de Wu est alors appliquée (MICD simple). Ensuite, chaque carré est divisé jusqu'à ce que sa taille maximale (donnée par le codeur spatial) soit atteinte. Pour un niveau donné (taille de bloc) de la pyramide, les blocs de taille inférieure ou égale sont traités via les passes 2 et 3 de l'algorithme. Les valeurs des blocs de taille supérieure sont recopiées afin de raffiner le contexte (figure 4).

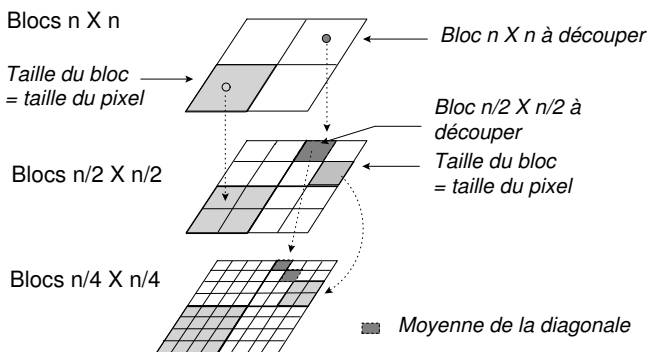


FIG. 4 – Valeur des blocs à prédire

Il est à noter que le schéma proposé développe de cette manière une méthode progressive *basée contenu* : si un contexte restreint est suffisant pour les grands blocs, on dispose d'un contexte plus riche pour les petits blocs (frontières des objets). En termes d'entropie, les résultats obtenus sont quasi équivalents à ceux que l'algorithme initial présentait, en termes de taux de compression, pour la compression sans perte (sans quantification) et avec pertes (avec quantification) de l'image des carrés.

La propriété principale de cette méthode est la minimisation de la redondance. En effet, partant du fait que la moyenne sur la diagonale est connue à un niveau donné, un seul bit est nécessaire à la reconstruction de la deuxième valeur issue de la diagonale au niveau inférieur, c'est à dire 0.25 bit par bloc et par niveau. Si une quantification est appliquée, ce bit disparaît et le nombre de symboles à transmettre est exactement égal au nombre de blocs du LAR. Notre approche conserve les propriétés d'un quad-tree, à savoir une progressivité de la résolution spatiale, mais améliore fortement son coût de codage. Un exemple est présenté figure 5.

4 Codage sans perte progressif

L'étage précédent du codeur produit une image basse résolution, et permet de séparer les blocs en deux groupes :

- les zones uniformes pour tous les blocs $N \times N$ où $N \in \{4, 8, 16\}$, caractérisées par une entropie faible,
- les contours (blocs 2×2), caractérisés par une forte entropie.

L'extension de la méthode à la compression sans perte consiste au codage de la texture issue de la différence entre l'image LAR basse résolution et l'image originale. Chaque bloc $N \times N$ de la représentation LAR est décomposé jusqu'à l'obtention

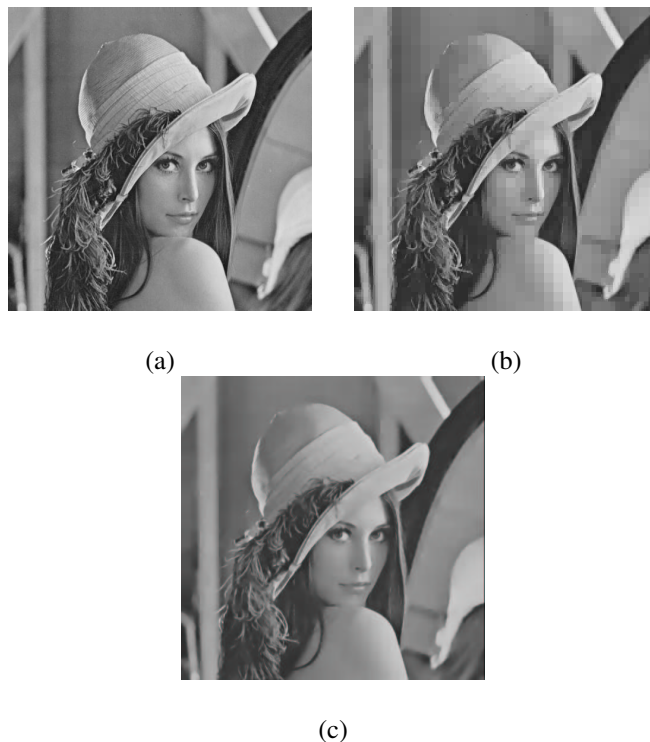


FIG. 5 – (a) Image originale. (b) Image LAR. MICD classique : 6.17 bit par bloc (0.32 bpb), approche pyramidale : 5.47 bpb (0.33 bpb). (c) Image LAR post-traitée.

de la pleine résolution, selon notre principe de décomposition pyramidale et de prédiction par contexte enrichi. Le codage sans perte implique l'absence de quantification des erreurs de prédiction à travers la pyramide.

4.1 Lois d'entropie et modélisation de contexte implicite

L'entropie globale est réduite lorsque l'on peut isoler différentes classes de symboles qui respectent la même loi. Ce principe est souvent utilisé dans le domaine de la compression sans perte, par le biais de la modélisation de contexte [1]. Une estimation de l'activité locale d'un pixel conduit alors à sa classification dans une loi a priori.

Notre méthode introduit quant à elle une séparation, issue d'une analyse préalable, des lois à deux niveaux :

- par la taille des blocs : les blocs de petites taille (typiquement 2×2), situés sur les contours induisent une erreur moyenne de forte entropie, alors que les autres blocs sont homogènes du point de vue de la texture ;
- par les différentes passes pour chaque niveau de décomposition : les valeurs issues de la passe 2 diffèrent de façon significative de celles de la passe 3. Le bit supplémentaire provenant de la moyenne de la diagonale est de même transmis séparément.

Le tableau 1 donne un exemple d'entropie finale obtenue lorsque les données issues de la représentation LAR et de la texture sont séparées, ainsi que celles provenant des différentes passes de l'algorithme de prédiction. Le niveau correspondant

TAB. 1 – Codage sans perte de Lena (4.42 bpp).

Niveau	Taille des blocs	Entropie (bpp)	
		Passé 1 : 6.07	
4	Tous blocs	Passé 2	Passé 3
3	Texture : 16×16	4.07	3.69
	LAR : $8 \times 8,4 \times 4,2 \times 2$	5.66	5.72
2	Texture : $16 \times 16,8 \times 8$	3.69	3.61
	LAR : $4 \times 4,2 \times 2$	5.46	5.71
1	Texture : $16 \times 16,8 \times 8,4 \times 4$	3.49	3.75
	LAR : 2×2	5.47	5.70
0	Texture : $16 \times 16,8 \times 8,4 \times 4$	3.37	3.72
	Contours : 2×2	4.81	4.98

à la pleine résolution distingue les blocs 2×2 des autres. Il est clair que la division du bitstream donne de meilleurs résultats. En fait, une modélisation implicite du contexte est fait et résulte directement de la nature de la structure quad-tree (basée contenu).

4.2 Résultats de la compression sans perte

Les résultats sont exprimés en termes d'entropie : le codeur arithmétique adaptatif n'a pour le moment pas été implanté. Il nous est donc difficile de comparer objectivement les performances du LAR avec celles des autres méthodes de l'état de l'art, qui présentent toutes un étage dédié au codage entropique.

Actuellement, les performances sont légèrement en dessous de l'état de l'art (ex. TMW : 4.30 bpp pour l'image correspondant à la composante de luminance de Lena, d'entropie brute 7,45 bpp), mais la méthode présente une **compression spatialement graduelle** (voir tableau 2). De plus, ce schéma a l'avantage de proposer à partir du même algorithme, une compression avec ou sans perte de l'image. Enfin, notre schéma peut encore être amélioré.

TAB. 2 – Résultats comparatifs : sans codage entropique

Image	Entropie (bpp)		
	Proposé (1 loi)	Proposé (pls lois)	CALIC
Baboon	7.02	6.13	6.14
Barb	5.88	5.02	4.93
Bike	5.40	4.63	4.53
Cafe	6.41	5.50	5.37
Finger	6.56	5.69	5.52
Hotely	5.66	4.82	4.56
Lena	5.07	4.42	4.33
Zelda	4.80	4.16	3.98
Average	5.85	5.03	4.92

5 Conclusion et perspectives

Ce papier a présenté une technique originale de compression progressive avec et sans perte. La seule connaissance de

la structure quad-tree est suffisante pour améliorer de façon significative le taux de compression pour le codage sans perte, en regroupant et encodant de façon séparée les informations de même nature.

Un avantage non négligeable du codeur LAR est de proposer au moyen d'un unique algorithme une compression avec ou sans perte. La configuration du codec est de plus simplifiée par le fait que le paramètre utile à la détermination de l'activité locale n'a pas d'impact significatif. La scalabilité introduite ici élargit le spectre des applications du codec LAR.

Les travaux futurs porteront sur l'adaptation de la prédiction en fonction de la nature des informations locales (contours ou texture). Dans le cas des images couleur, il est de plus possible d'améliorer le schéma en prenant en compte les informations d'orientation de contours dans la composante de luminance afin de coder plus efficacement les valeurs de chrominance.

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*, 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.
- [3] B. Meyer and P. Tisher, "TMW : a new Method for Lossless Image Compression," *Proc. Picture Coding Symposium, Berlin*, 1997.
- [4] O. Déforges and J. Ronsin, "Codage de Région d'Intérêt par la Méthode LAR," *CORESA'00*, Octobre 2000.
- [5] O. Déforges and J. Ronsin, "Supervised Segmentation at Low Bit Rates for Region Representation and Color Image Compression," *IEEE ICME, Lausanne*, Octobre 2000.
- [6] D. D. Muresan and T. W. Parks, "Optimal Recovery Approach to Image Interpolation," *Proc. IEEE ICIP, Greece*, p. Image Interpolation and Spatial Filtering, 2001.
- [7] X. Wu, "Lossless Compression of Continuous-Tone Images, via Context Selection, Quantization and Modeling," *IEEE Trans. on Image Processing*, vol. 6, no. 5, pp. 656–664, 1996.