Décomposition pyramidale par prédiction et transformée pour du codage avec et sans perte

Marie Babel, Olivier Déforges

▶ **To cite this version:**

Marie Babel, Olivier Déforges. Décomposition pyramidale par prédiction et transformée pour du codage avec et sans perte. Jan 2003, pp.80. hal-00132617

HAL Id: hal-00132617
https://hal.science/hal-00132617

Submitted on 22 Feb 2007

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

DÉCOMPOSITION PYRAMIDALE PAR PRÉDICTION ET TRANSFORMÉE POUR DU CODAGE AVEC ET SANS PERTE

Marie Babel[1] Olivier Déforges[1]

[1] UMR CNRS 6164 IETR Groupe Image, INSA de Rennes

20, av. des Buttes de Coësmes, 35043 RENNES Cedex, FRANCE

{mbabel,odeforges}@insa-rennes.fr

Résumé

Cet article présente une technique de compression conjointe avec et sans perte des images en niveaux de gris. En particulier, nous introduisons une forme originale de MICD à contexte enrichi à travers une décomposition pyramidale, afin d'améliorer et de rendre progressive la méthode LAR pour du codage avec pertes. L'image d'erreur peut ensuite être codée sans perte par prédiction ou transformée Hadamard 2×2 *réversible, en fonction de l'activité locale de l'image.*

Mots clefs

Compression avec et sans perte, MICD, transformée Hadamard, pyramide, progressivité.

1 Introduction

Dans le domaine de la compression sans perte des images en niveaux de gris, deux grands groupes de techniques se distinguent : celles qui utilisent des prédicteurs, comme le standard CALIC [1], et celles qui manipulent des transformées réversibles. Une solution mixte, combinant à la fois prédiction et transformée, est proposée par l'algorithme S+P [2].

Par ailleurs, TMW [3] pose le principe de deux étapes d'encodage. Ainsi, l'image peut être vue comme la superposition de l'information globale et de la texture (information locale). Ce concept nous permet de fournir une compression avec pertes (première partie du message) de très bonne qualité visuelle dans le cas du LAR, ou sans perte (ajout de la deuxième partie du message, correspondant à l'image d'erreurs).

La méthode LAR (*Locally Adaptive Resolution*) [4], basée sur une décomposition à taille de blocs variable (section 2), se révèle être très efficace pour de la compression avec pertes. Cet article présente donc une nouvelle technique de compression conjointe avec et sans perte, améliorant de ce fait les performances et les fonctionnalités du codeur LAR. Pour ce faire, un schéma de type MICD et une transformée Hadamard 2x2 réversible sont introduits dans une décomposition pyramidale particulière.

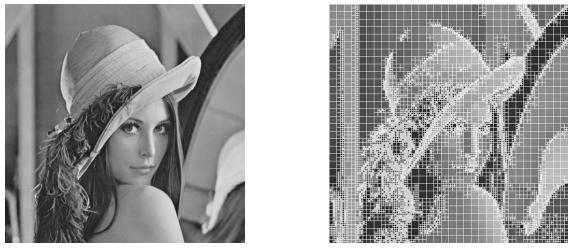
La section 3 définit donc une décomposition pyramidale à redondance minimale permettant une reconstitution progressive de l'image compressée avec pertes, alors que la dernière partie traite du codage sans perte.

2 La méthode LAR

La méthode de compression LAR, initialement conçue pour un codage graduel avec pertes des images en niveaux de gris, met en œuvre deux codeurs (spatial et spectral) [5]. Le codeur spatial fournit une image compressée à bas débit, alors que le codeur spectral encode la texture. De plus, des tests comparatifs ont permis d'évaluer la qualité des images LAR basse résolution et ont ainsi démontré la supériorité de notre codage vis-à-vis de Jpeg-2000 [4].

2.1 Le codeur spatial

Le principe est basé sur le fait que la résolution de l'image (la taille des pixels) peut varier en fonction de l'activité locale, estimée par un gradient morphologique (Fig.1). Le codeur spatial sous-échantillonne l'image sur une grille carrée non uniforme (topologie de type quad-tree), et produit une image basse résolution : chaque bloc est alors reconstruit par la valeur moyenne de la luminance issue de l'image source. La taille des blocs obtenus 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 est automatiquement disponible à la fois au codeur et au décodeur.



(a) (b)

Figure 1 – *(a) Image originale. (b) Représentation à taille de blocs variable (grille sous-échantillonnée : 0.032 bpp).*

Une prédiction de type MICD est utilisée pour coder la luminance moyenne de chaque bloc. La quantification de l'erreur s'adapte en fonction de la taille des pixels : elle est donc grossière sur les petits blocs et fine sur les gros blocs (l'oeil perçoit fortement les variations de luminosité sur les zones uniformes). 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 linéaire).
En résumé, ce codeur possède principalement deux caractéristiques essentielles, à savoir : d'une part, la technique utilisée est très rapide et efficace pour des taux de compression élevés, et d'autre part, la méthode tend à simplifier l'image source en supprimant la texture tout en préservant les contours des objets.

2.2 Le codeur spectral

Le codeur spectral constitue le deuxième étage du codeur, et permet la reconstruction 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. Utiliser des tailles de blocs adaptées permet un codage progressif basé contenu : en effet, le rehaussement de contour par exemple peut être réalisé par la seule transmission des coefficients AC des petits blocs.

Ce codeur de base a également été étendu pour de la compression avec zone d'intérêt [5] et du codage d'image couleur basé régions [4].

3 Approche pyramidale prédictive

L'image des blocs du LAR respecte une topologie de type quad-tree. Cependant, une approche classique MICD 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 Algorithme de prédiction

L'actuel codeur spatial du LAR encode en une seule passe (raster scan) l'intensité lumineuse des blocs à l'aide du prédicteur GAP (Gradient Adjusted Predictor) [1]. La nouvelle version met en place un schéma de type MICD basé sur le prédicteur décrit dans [6]. Pour une image pleine résolution, les erreurs sont codées sans perte au moyen de trois échantillonnages entrelacés de l'image originale. Ainsi, on obtient un contexte dit à 360° (configuration spatiale du voisinage entourant un pixel donné).
Le principe général est le suivant (Fig.2) : la première passe encode l'image sous-échantillonnée formée par la moyenne de deux pixels diagonalement adjacents à travers un MICD classique, puis la seconde passe fournit la valeur des deux pixels qui ont servi au calcul de la diagonale précédente. A ce stade, la prédiction utilise un contexte à 360° constitué des valeurs connues des pixels de la passe courante et des moyennes des diagonales codées à la première passe. Enfin, la troisième passe traite la moitié restante de l'image : de nouveau, un contexte entourant complètement le pixel courant est disponible.

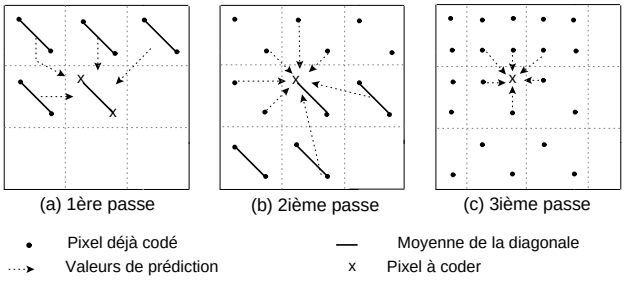


Figure 2 – *Algorithme de Wu : 3 passes et voisinage*

3.2 Décomposition pyramidale à redondance minimale

Le principe général de prédiction précédent traitant de la pleine résolution a été étendu au cas d'une décomposition pyramidale.
La construction de la pyramide est du type *bottom-up*. Ainsi, 4 blocs $\frac{N}{2} \times \frac{N}{2}$ sont regroupés pour former un bloc $N \times N$ (Fig.3). Deux possibilités s'offrent à nous :
– pour un bloc de taille finale 2×2, au niveau 2×2, la moyenne de la diagonale est encodée,
– si la taille du bloc $N \times N$ est supérieure à celle déterminée par le codeur spatial du LAR, alors la valeur du bloc est formée par la moyenne de la première diagonale,
– sinon, la valeur du bloc est la moyenne sur les 4 blocs.
Le plus haut niveau de la pyramide (niveau 4) correspond aux blocs 16×16 (taille la plus grande autorisée dans la méthode LAR).

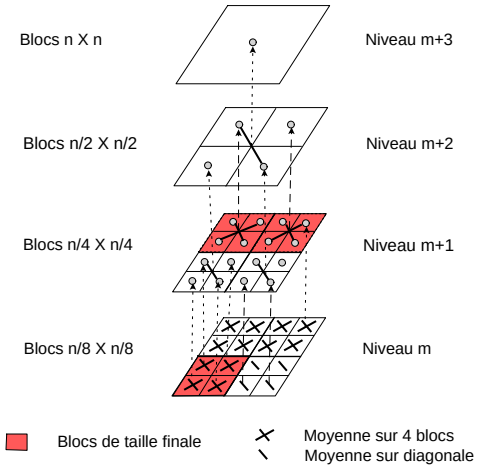


Figure 3 – Construction *bottom-up*.

La décomposition pyramidale va donc se faire par chemin inverse (*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 (Fig.4).

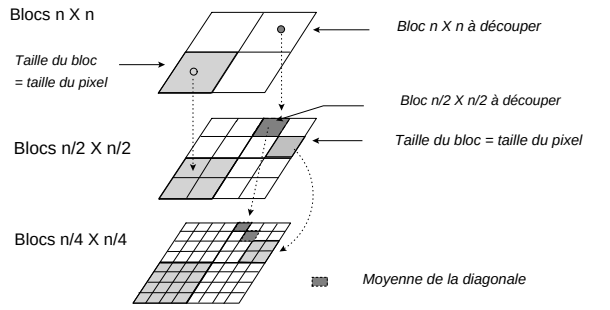


Figure 4 – *Valeurs prédites des blocs*

Ainsi, une méthode *progressive basée contenu* est développée : si un contexte restreint est suffisant pour les grands blocs, un contexte plus riche est disponible pour le codage des petits blocs (frontières des objets). La principale propriété de ce schéma est une *redondance minimale*. Considérons les quatre blocs issus de la division d'un bloc du niveau supérieur. La moyenne de la première diagonale est connue grâce au niveau supérieur de la pyramide. Afin de reconstruire la valeur du deuxième bloc de la diagonale, il nous suffit d'un seul bit supplémentaire, à savoir 0.25 bit en moyenne par bloc et par niveau. Ce bit disparaît si une quantification est mise en œuvre : le nombre de symbole est alors exactement égal au nombre de blocs de la représentation LAR.

	Entropie (bit par bloc)		Entropie globale (bpp)	
Image name	Init.	Prop.	Init.	Prop.
Baboon	6.73	6.79	1.05	1.06
Barb	6.45	6.54	0.67	0.68
Bike	6.68	7.17	0.41	0.44
Cafe	6.80	7.12	0.84	0.88
Finger	6.51	6.46	1.36	1.35
Hotely	6.25	6.61	0.52	0.55
Lena	6.17	5.98	0.32	0.33
Zelda	5.81	6.15	0.17	0.18

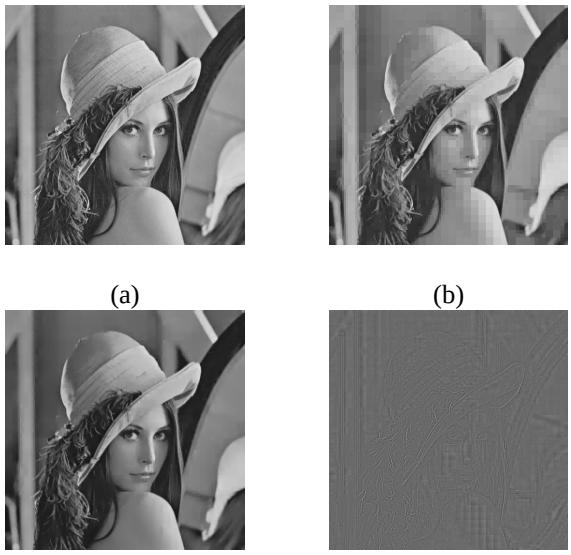
Tableau 1 – *Résultats de compression pour l'image LAR : version initiale et porposeé.*

3.3 Résultats

Le tableau 1 donne les entropies obtenues avec notre approche pyramidale, comparées avec celles résultant du LAR actuel. L'observation principale est la suivante : l'entropie est équivalente pour les deux techniques ; cependant, l'encodage est désormais totalement progressif.

4 Codage sans perte progressif

Le principe général correspond à celui décrit dans l'algorithme TMW [3] : l'image peut être vue comme la superposition de l'information globale (ici l'image des blocs) et du résidu contenant l'information sur la texture locale. L'image d'erreurs issue de l'image LAR et de l'image originale est donc codée afin d'obtenir une compression sans perte.



(a) (b)

(c) (d)

Figure 5 – *(a) Image originale. (b) Image LAR. MICD classique : 6.17 bit par bloc (0.32 bpp), approche pyramidale : 5.47 bpb (0.33 bpp). (c) Image LAR post-traitée. (d) Image d'erreurs. (coût sans perte : 4.38 bpp).*

4.1 Lois d'entropie et codage

L'entropie globale est réduite lorsque différentes classes de symboles appartenant à une même loi peuvent être séparées. La modélisation de contexte [1] tire parti de ce principe par l'estimation de l'activité d'un pixel donné.
L'image des blocs peut donc être codée en tenant compte des différentes passes appliquées, ainsi qu'en transmettant séparément le bit issu de la première diagonale calculée. L'image d'erreurs peut alors être elle aussi divisée en deux types de zones : les zones homogènes (blocs 4×4 à 16×16), pour lesquelles le résidu est faible et fortement aléatoire, et les zones de contours (blocs 2×2) à valeur d'erreur importante mais corrélée au voisinage. Le premier type de zone a été codée par transformée de Hadamard sans

perte 2×2 [7], où la composante DC est ici directement fournie par le niveau supérieur de la pyramide. Les zones de contours restent quant à elles plus efficacement codées par notre principe de prédiction par contexte enrichi.

Niveau	Taille des blocs	Entropie (bpb)
3	Texture : 16×16	Had. : 2.76
	LAR : 8×8,4×4,2×2	Passe 2 : 5.47
		Passe 3 : 5.47
2	Texture : 16×16,8×8	Had. : 2.78
	LAR : 4×4,2×2	Passe 2 : 5.18
		Passe 3 : 5.34
1	Texture : 16×16,8×8,4×4	Had. : 3.13
	LAR : 2×2	Passe 2 : 5.03
		Passe 3 : 5.31
0	Texture : 16×16,8×8,4×4	Had. : 3.42
	Contours : 2×2	Passe 2 : 4.49
		Passe 3 : 4.75
4	Tous blocs	Passe 1 : 6.16

Tableau 2 – *Codage sans perte de Lena*

Le tableau 2 donne un exemple de l'entropie finale sur l'image Lena, lorsque les valeurs de la représentation LAR, de la texture et des passes 2 et 3 sont séparées (Entropie finale : 4.38 bpp). Comparée à un codage global, une nette amélioration de l'entropie est aisément observable. En fait, la décomposition quad-tree étant déjà basée contenu, une modélisation du contexte est implicite dans notre schéma.

	Entropie (bpp)	
Image	Proposé (pls lois)	CALIC
Baboon	6.10	6.14
Barb	5.03	4.93
Bike	4.62	4.53
Cafe	5.45	5.37
Finger	5.78	5.52
Hotely	4.69	4.56
Lena	4.38	4.33
Zelda	4.16	3.98
Moyenne	5.02	4.92

Tableau 3 – *Résultats de compression sans perte*

4.2 Résultats

Le tableau 3 présente les résultats de CALIC et du LAR sans perte sur plusieurs images standard. Les performances sont un peu moins bonnes que l'état de l'art (ex. TMW : 4.30 pour Lena), mais l'approche pyramidale possède l'avantage d'être une technique de compression totalement graduelle.

Par ailleurs, le LAR doit être configuré à l'aide d'un seul paramètre, qui règle le seuil du gradient morphologique. Il détermine la proportion d'information destinée respectivement à la texture et à l'image basse résolution. Nous avons pu constater que pour des valeurs du seuil comprises entre 10 et 20, l'entropie reste sensiblement constante. La méthode est donc très peu sensible à ce paramètre : il est alors possible de fixer le seuil a priori.

5 Conclusion

Cet article décrit donc une technique originale de compression avec et sans perte progressive. La connaissance de la structure quad-tree associée à l'image LAR est suffisante pour améliorer de façon significative le taux de compression du codage sans perte de l'image, en regroupant et en codant séparément les informations de même nature (coefficients transformés, contours...).

A partir d'un unique algorithme, l'utilisateur choisi de coder une image sans perte ou avec pertes. Ses opérations sont de plus simplifiées par le fait que le choix du paramètre d'estimation du gradient n'a pas d'impact significatif sur le taux de compression. La notion de progressivité introduite ici ajoute de nombreuses possibilités aux codec. De plus, les propriétés du précédent codeur LAR sont conservées.

Les travaux de recherche futurs concerneront la prédiction éventuelle des coefficients de la transformation Hadamard, et l'adaptation plus fine de l'utilisation de cette transformée ou du schéma MICD en fonction de la nature des informations locales (contours ou texture).

Références

[1] X. Wu et 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 et W. Pearlman. Reversible image compression via multiresolution representation and predictive coding. Dans *Visual Communication and Image Processing*, volume 269. SPIE, 1993.

[3] B. Meyer et P. Tisher. TMW : a new Method for Lossless Image Compression. *Proc. Picture Coding Symposium, Berlin*, 1997.

[4] O. Déforges et J. Ronsin. Supervised Segmentation at Low Bit Rates for Region Representation and Color Image Compression. *IEEE ICME, Lausanne*, Octobre 2000.

[5] O. Déforges et J. Ronsin. Codage de Région d'Intérêt par la Méthode LAR. *CORESA'00*, Octobre 2000.

[6] X. Wu. Lossless Compression of Continuous-Tone Images, via Context Selection, Quantization and Modelling. *IEEE Trans. on Image Processing*, 6(5) :656–664, 1996.

[7] H.Y. Jung, T.Y. Choi, et R. Prost. Rounding Transform and its Application for Lossless Pyramid Structured Coding. *IEEE Trans. Image Processing*, 7(2) :234–237, Février 1998.