

Généralisation du JPEG standard à travers la méthode LAR (Locally Adaptive Resolution)

Olivier DEFORGES, Joseph RONSIN

Laboratoire ARTIST, INSA de Rennes
20, Avenue des Buttes de Coësmes, CS 14315, 35043 Rennes, France

olivier.deforges@insa-rennes.fr, joseph.ronsin@insa-rennes.fr

Résumé - Cet article présente un nouveau schéma de codage pour les images fixes, combinant à la fois un codeur spatial, et un codeur spectral. Le premier est basé sur une nouvelle forme de sous-échantillonnage non uniforme simple, adaptant la résolution locale à l'activité dans l'image, et ayant pour effet d'éliminer la texture locale. Il peut être utilisé seul pour atteindre de fort taux de compression. Afin d'améliorer la qualité, l'image d'erreur peut être transmise à travers un second décodeur basé JPEG, mais pour lequel la taille des blocs est variable, et fournie par le premier codeur.

Abstract - This paper presents a still image coding scheme combining both a spatial and a spectral encoders. The first one is based on a new fast and simple nonuniform sub-sampling method adapting the resolution to the local image activity, and leading to the local texture removal. This encoder can be used as a stand-alone technique for data compression at low bit rate. To improve the image quality, the error image can be passed through a second encoder, JPEG based, but with variable blocks sizes provided by the first encoder.

1. Introduction

De manière synthétique, on peut dire qu'il existe deux voies principales pour le codage d'image fixe : les méthodes spatiales et les méthodes spectrales. Les premières agissent en sélectionnant directement dans l'image des échantillons. Elles peuvent parfois atteindre de bonnes performances à bas débit [1, 2], mais elles sont le plus souvent lourdes et complexes à mettre en œuvre. De plus, elles ne donnent que de résultats médiocres à moyen/haut débit. A l'inverse, les méthodes spectrales basées sur des transformations d'espace de représentation sont relativement simples d'emploi, offrent de bonnes performances à moyen et haut débits, mais dans une moindre mesure à bas débit.

Les techniques basées transformation ont prouvé leur efficacité dans le domaine du codage, plus spécifiquement lorsque les données présentes dans l'image sont fortement corrélées. Dans le cas contraire, une quantification trop importante dans l'espace transformé aboutit à la création d'artefacts dans l'image reconstruite. Ainsi, le principal défaut de techniques de transformation utilisant des blocs de taille fixe, tel que la DCT 8x8, est que la qualité visuelle à bas débit est fortement liée à la stationnarité de chaque bloc, et n'est donc pas constante dans toute l'image.

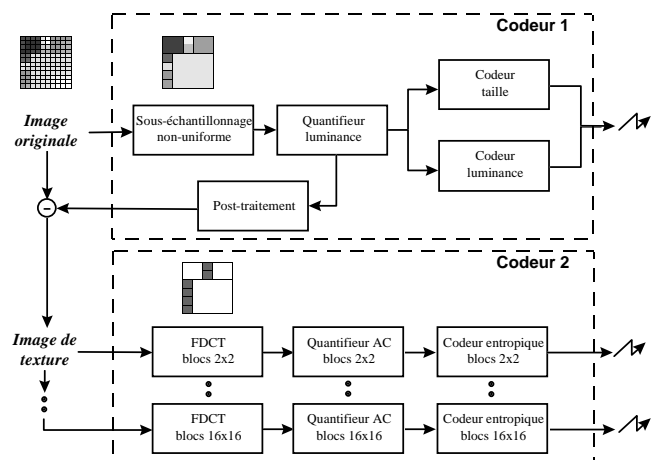
Quelques méthodes originales ont proposé de construire une représentation spatiale initiale de l'image de type bin-tree ou quad-tree, en fonction de sa stationnarité locale. La transformée est alors appliquée mais sur des blocs de taille variable [3]. Dans [4], une DCT à taille de bloc variable est introduite. La taille des blocs y est transcrite à travers un quad-tree. Un critère basé sur une différence moyenne décide de la taille qui peut être 4, 8 ou 16.

La méthode présentée ici est proche de la précédente à la différence essentielle que la représentation spatiale est fournie par un premier schéma de codage, et ne constitue

donc pas un coût supplémentaire. Ce premier codeur est destiné à atteindre de fort taux de compression. Il peut être rapproché d'une segmentation de type quad-tree, mais en réalité nous verrons qu'il présente des singularités qui lui valent d'atteindre de meilleures performances tout en restant de faible complexité. Le second codeur peut s'apparenter à un codeur de type DCT à taille de bloc variable. L'ensemble offre un schéma de codage progressif.

La figure 1 présente le bloc diagramme général.

L'idée à la base de la méthode LAR (Locally Adaptive Resolution) est que la résolution locale (taille d'un pixel) peut être adaptée en fonction de l'activité dans l'image : lorsque la luminance est localement uniforme, la résolution peut être abaissée (pixel de grande taille). A l'inverse, si l'activité est localement élevée, alors la résolution doit être plus importante (pixel de petite taille).



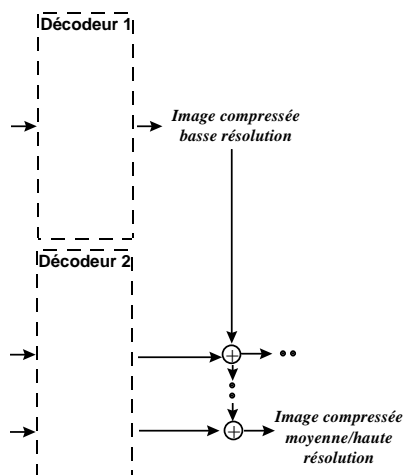


FIG. 1 : bloc diagramme général

2. Codeur spatial

2.1. Sous-échantillonnage non uniforme par maillage carré

En fonction d'une famille de taille de bloc donnée à croissance dyadique (typiquement on utilise les carrés 2x2, 4x4, 8x8, 16x16), l'algorithme sous-échantillonne l'image sur une grille carrée non-uniforme sans recouvrement, en adaptant la taille du carré à l'activité. L'estimation de cette activité est réalisée à travers la mesure d'un gradient morphologique dans une taille de carré considéré. Ce critère d'absence de contour comme caractéristique de stationnarité sera exploité par la suite. Ainsi, les carrés de petite taille se retrouvent concentrés sur les frontières, alors que les zones homogènes sont décrites par des carrés de grande taille. Une structure de représentation de type quad-tree peut à ce niveau être employée. Toutefois nous avons développé un autre type de maillage appelé half quad-tree, pour lequel seule une contrainte de pavage est imposée dans le sens vertical. Dans la direction horizontale au contraire, il n'existe aucune limitation, ce qui aboutit à une adéquation complète dans cette direction entre les tailles maximales estimées et les tailles réelles utilisées. Comparé au quad-tree, ce pavage induit en moyenne 20% de carrés en moins pour la description de l'image.

Un exemple d'image des tailles de carré est fourni figure 2.a. Ce pavage de l'image fournit à lui seul une carte de segmentation de l'image. A noter le nombre importants de carrés engendrés, dû notamment à l'utilisation de carrés de petite taille, qui peut apparaître prohibitif comparé aux autres méthodes de sous-échantillonnage non uniforme. La suite montre qu'en fait ces carrés peuvent être codés très efficacement.

2.2. Quantification de la luminance

A chaque carré obtenu est associée la luminance moyenne de la zone correspondante dans l'image source. La

quantification utilisée ensuite est une simple quantification linéaire, mais s'appuie sur la connaissance a priori du contenu de l'image fournie par la taille des carrés : les grands carrés requièrent une quantification fine car ils sont placés sur des zones homogènes (forte sensibilité de l'œil humain à des variations de contraste). Les petits carrés eux, situés sur des zones frontières, supportent une quantification très grossière sans toutefois être perçue. Le tableau 1 présente le nombre de niveaux de quantification utilisé par taille, pour une quantification non perceptible.

TAB. 1 : quantification des luminances

Taille carré	2x2	4x4	8x8	16x16
Niveaux de quantif.	16	32	64	128

2.3. Codage sans perte des tailles et des luminances

Les luminances quantifiées sont ensuite codées sans perte sous leur forme différentielle. Afin d'optimiser les prédictions des luminances, obtenues sous une forme de type MICD, on utilise une nouvelle fois l'information fournie par la taille : aux grands carrés correspondent des estimateurs de type moyenne en fonction du voisinage, alors que pour les petits carrés, on utilise des estimateurs de type non linéaires.

Pour le codage entropique final, un codeur arithmétique adaptatif d'ordre 2 est employé pour à la fois les tailles et les luminances. Le coût moyen final par carré est inférieur à 4 bits (environ 0.8 bit pour la taille, 3 bits pour la luminance).

2.4. Post-traitement

De par son essence, la méthode crée des effets de blocs dans les zones homogènes. Toutefois, l'image résultante reste de bonne qualité visuelle car les contours principaux des objets sont conservés. De plus, un post-traitement simple mais aussi très efficace peut être utilisé. Son principe est que plus un carré est grand, plus sa luminance peut être corrigée (ici par une interpolation linéaire des luminances de son voisinage). Un exemple d'image résultante bas débit est donné figure 4.a.

Plus de détails sur ce codeur pourront être trouvés dans [5]. Nous retiendrons simplement ses principales caractéristiques :

- une technique très rapide et efficace à bas débit,
- la méthode simplifie l'image en éliminant localement la texture, mais préserve les contours,
- la représentation est multi-échelle.

3. Codeur spectral

Le codeur précédent peut être utilisé seul pour de fort taux de compression (grossoirement entre 80 et 30). Lorsque l'on désire augmenter la qualité, il se révèle toutefois moins efficace. En effet, cela suppose représenter l'image par des carrés plus petits. Si la quantification des luminances présentée ci-dessus est conservée, elle devient alors de plus en plus visible. L'autre solution proposée ici est de compresser l'image d'erreur (\approx texture locale) issue du

premier codeur par un second codeur de type spectral, basé sur l'approche DCT à taille de bloc variable. Ce codeur est donc composé d'une batterie de codeurs de type JPEG avec pertes [6], chacun étant adapté à une taille de carré. (voir fig. 1). Seuls les coefficients AC nécessitent d'être transmis puisque la luminance moyenne (coefficient DC) est déjà fournie par le premier codeur.

Ce choix de codeurs JPEG induit que le JPEG standard apparaît comme un cas particulier de la méthode LAR, lorsque le sous-échantillonnage initial est uniforme sur des carrés 8x8.

3.1. Intérêt des blocs à taille variable

A l'exception des blocs 2x2 qui correspondent à la résolution minimale, par construction, les autres blocs ne contiennent pas d'élément frontière. Seule la texture locale est présente. Cela entraîne que l'énergie moyenne des coefficients AC est très inférieure à celle trouvée pour du JPEG standard (voir fig. 3). Après quantification, le nombre de coefficients à transmettre est donc limité.

3.2. Résolution « sémantiquement » adaptative et codage progressif

Les blocs d'une taille donnée sont traités indépendamment des blocs d'une taille différente. Cela suppose que pour chaque taille de bloc, une quantification spécifique est appliquée. Une grande flexibilité dans le choix de la qualité finale est alors introduite : pour améliorer la finesse des contours, on privilégiera la quantification des blocs 2x2. Au contraire, pour rehausser la qualité du fond des objets, on accordera plus d'importance à la quantification des grands blocs. Ceci constitue une forme de multirésolution non pas globale sur toute l'image, mais adaptée au contenu sémantique (frontière forte, frontière faible, zone faiblement homogène, zone fortement homogène).

Ce même aspect peut être exploité pour rendre le codage progressif. A partir de l'image bas débit issue du premier codeur, la qualité augmente progressivement par ajout des blocs d'erreur compressés en considérant successivement les différentes tailles. Le processus illustré fig.1 conduit à améliorer en premier les frontières, puis en final à restaurer le fond. Le procédé inverse peut bien sûr être implanté.

4. Résultats

La figure 4 montre les images successives obtenues, et donne les performances correspondantes.

Les résultats comparatifs des rapports débit/distorsion avec le JPEG standard dans le mode séquentiel (non progressif) montre que la méthode LAR aboutit à des PSNR toujours au dessus d'au moins 1db, et ce quel que soit les débit. Il faut pourtant noter que ce critère global de mesure d'erreur défavorise la méthode qui utilise une quantification psychovisuelle des luminances, autrement dit engendrant des erreurs mais non perceptibles.

Plus que sur le PSNR, les performances de la méthode peuvent se juger sur la qualité visuelle des images compressées. A bas débit, l'image est simplifiée en diminuant la résolution d'une part (si la taille de carré 1x1

n'est pas employée), et en lissant les zones homogènes, mais conserve globalement les contours des objets (fig. 4.a). A moyen débit, si l'on rajoute la texture des petits blocs, la résolution sur les contours est améliorée, conduisant à une amélioration significative de l'image (fig. 4.b). Enfin la texture des grands blocs permet un rendu complet de l'image (fig. 4.c).

5. Conclusion et perspectives

Nous avons présenté une nouvelle méthode de codage pour les images fixes baptisée LAR pour « Locally Adaptive Resolution ». Le premier codeur pour les bas débits est basé sur une représentation par un maillage carré de l'image en fonction de son activité. De faible complexité, il permet d'atteindre de forts taux de compression tout est conservant une image de bonne qualité visuelle. En outre, il fournit à un second codeur la représentation spatiale nécessaire pour accomplir une compression de type JPEG pour des tailles de bloc variables.

L'ensemble du schéma de codage est globalement progressif, et permet d'adapter la résolution en fonction du contenu sémantique de l'image.

L'extension actuelle de la méthode concerne le codage de zone d'intérêt, où l'idée est de coder grossièrement l'image sauf la zone d'intérêt qui elle est codée plus finement. L'image est tout d'abord compressée à bas débit par le premier codeur. En considérant une zone d'intérêt comme une composante connexe, une segmentation effectuée à la fois au niveau du codeur et du décodeur, construit les composantes comme union de carrés connexes. Ainsi, une zone d'intérêt est simplement désignée par le numéro de la ou des composantes connexes qui la composent, sans coût de description supplémentaire. Seule la texture locale des carrés appartenant à la zone d'intérêt est ensuite transmise. Un avantage majeur ici de cette approche est que la description d'une zone suit exactement les contours des objets.

Les travaux futurs vont tenter d'étendre la méthode au codage de séquences d'image. Si le problème de la représentation en quad-tree pour une telle application est qu'il ne présente pas d'invariance en translation, ce problème est en partie résolu par le pavage « half quad-tree » qui lui possède cette propriété mais uniquement suivant la direction horizontale. Aussi avons-nous développé un nouveau pavage baptisé « object quad-tree » possédant l'invariance en translation à la fois en horizontal et en vertical.

Références

- [1] G. Wolberg, I. Gertner, *Non-uniform image reconstruction using multilevel surface interpolation*. SAMPTA-97, pp 187-192, 1997.
- [2] S. Saryazdi, V. Haese-Coat, J. Ronsin. *Lossy image Compression by a high performance non-uniform morphological scheme*. SAMPTA-97, pp 37-42, 1997.
- [3] R. Nohre. *On Image Coding Using fragmented Blocks*. PCS'96, Melbourne, pp 645-646, 1996.

- [4] C. Chen. *Adaptive transform coding via quad-tree based variable block-size DCT*. ICASSP'89, Glasgow, pp 1854-1856, 1989.
- [5] O. Déforges, J. Ronsin. *Nonuniform sub-sampling using squares elements : a fast still image coding at low bit*

- rate*. International Picture Coding Symposium'99, Portland, 21-23 april 1999.
- [6] W. B. Pennebaker, J. Mitchell. *JPEG: still image data compression standard*. V. N. Reinhold, 1993.



FIG. 2 : Lena 512x512 compressée par la méthode LAR : a) image des tailles (16x16 en blanc, 2x2 en noir), 13888 carrés b) image reconstruite par la luminance moyenne quantifiée, c) erreur de quantification (multipliée par 15)

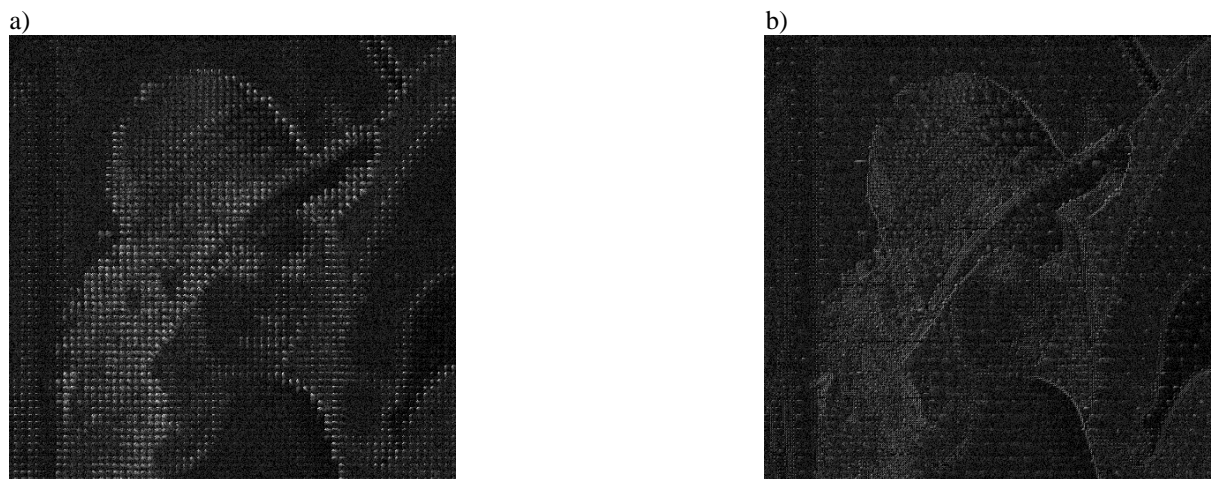


FIG. 3 : image des coefficients AC : a) pour le JPEG standard, b) pour LAR



FIG. 4 : résultats du codage progressif a) image basse résolution issue du premier codeur : 0.2 bpp - 29.1db b) + blocs 2x2 et 4x4 : 0.39 bpp - 32 db, c) + blocs 8x8 et 16x16 : 0.5 bpp - 33.5 db