

# Codage multirésolution sans perte et insertion de données : une approche intégrée

Jean MOTSCH<sup>1</sup> Olivier DUTHEIL<sup>2</sup> Olivier DÉFORGES<sup>2</sup> Marie BABEL<sup>2</sup>

<sup>1</sup> CREC Saint-Cyr/LESTP

IETR UMR CNRS 6164

Groupe Image et Télédétection

jean.motsch@st-cyr.terre.defense.gouv.fr, {olivier.deforges, marie.babel}@insa-rennes.fr

<sup>2</sup> INSA RENNES

IETR UMR CNRS 6164

Groupe Image et Télédétection

## Résumé

*Pour éviter les artefacts de codage et les erreurs d'interprétation, l'imagerie médicale nécessite une compression sans perte. Le suivi des examens et des images, ainsi que l'ajout d'informations utiles, a conduit au développement de méthodes d'insertion de métadonnées. Les méthodes usuelles, celles du standard de fichiers DICOM<sup>1</sup> par exemple, reposent sur des entêtes pour y parvenir. Cet article présente l'utilisation d'une méthode réversible d'insertion de données rapide et efficace intégrée dans un schéma de compression multirésolution sans perte d'images fixes. Des résultats expérimentaux confirment la pertinence d'une telle adéquation en termes de qualité visuelle et de capacité d'insertion de données. En outre, cette insertion ne se fait pas au détriment des performances de compression qui restent des plus intéressantes.*

## Mots clefs

insertion réversible de données, codage multirésolution sans perte

## 1 Introduction

Les schémas de compression d'images médicales doivent remplir plusieurs missions : gérer une énorme quantité de données, fournir de bonnes performances en compression sans perte et intégrer de nouveaux services. Ces services incluent la protection du contenu, une transmission sécurisée [1] et l'insertion de métadonnées. Dans la plupart des applications, les services sont obtenus indépendamment de la méthode de compression, pour un surcoût important.

L'insertion de données vise à cacher des informations, appelée *charge utile*, dans une image numérique. Cela doit se faire de la manière la plus transparente possible. Ainsi, la qualité visuelle de l'image doit-elle rester très bonne après l'insertion des données. Les performances des algorithmes

d'insertion de données sont mesurées par trois principaux critères [2] :

1. la capacité d'insertion de données, c'est-à-dire la quantité maximale d'information pouvant être insérée ;
2. la qualité visuelle, dépendant des artefacts et de la distorsion introduits par l'algorithme ;
3. la complexité, représentant le coût calculatoire de la méthode d'insertion utilisée.

Ainsi, les méthodes spatiales reposent sur la modification des pixels suivant un schéma spécifique. Des exemples concernent l'altération des bits de poids faibles [3], des méthodes statistiques comme le *patchwork* [3] et l'usage de fractales [4].

Les méthodes spectrales utilisent l'invariance du spectre par rapport à des classes de transformations géométriques [5]. L'étalement de spectre est souvent usité pour différentes transformations, tels la transformée de Fourier-Mellin [5] [6], la transformée de Laguerre [7], et celles de Fresnel [8] et Haar [9].

Les méthodes jointes de compression et insertion de données sont classiquement des méthodes spectrales utilisant la transformation calculée par le codeur d'images fixes. Des méthodes reposant sur la transformée en cosinus discrète (TCD) sont ainsi intégrées dans le codeur JPEG, utilisant par exemple l'inversion des coefficients de la TCD [10], l'addition [11] et l'étalement de spectre [12]. De la même façon, les coefficients d'ondelette de JPEG-2000 peuvent être marqués [13].

Cet article présente l'utilisation conjointe d'un schéma de compression, le codeur LAR et d'une méthode d'insertion de données, l'expansion de la différence (ED). La section 2 introduit le codeur LAR et la section 3 l'expansion de la différence. La section 4 décrit ensuite un schéma d'insertion de données dans le flux binaire LAR dont certains résultats expérimentaux sont présentés dans la section 5 avec quelques éléments de discussion. La section 6 conclut notre propos.

<sup>1</sup><http://dicom.nema.org>

## 2 LAR Interleaved S+P

En 1999, Said[14] résume les propriétés souhaitables pour un codeur : scalabilité, adaptabilité, ajustement automatique du débit binaire, unicité de l'algorithme pour du codage avec et sans perte, décodage des zones d'intérêt, faible complexité, fort taux de compression, bonne robustesse aux erreurs et bonne qualité visuelle.

Les principales normes de compression se concentrent seulement sur une ou deux des propriétés énoncées. Le schéma de compression LAR suit une démarche différente en cherchant à les obtenir toutes [15]. Ainsi, le LAR tend à être une solution de codage complète, le « couteau suisse » de la compression des images fixes. Le codeur LAR fournit une scalabilité tant en résolution qu'en qualité, des flux binaires emboîtés, du codage avec pertes vers du codage sans perte, de faible complexité, pour des images en niveaux de gris et en couleurs, le tout avec des performances au-delà de l'état de l'art en termes de courbes débit-distorsion.

La famille des codeurs LAR repose sur l'idée que la résolution doit être adaptée à l'activité locale présente dans l'image. Une résolution basse correspond donc aux zones uniformes tandis qu'une résolution élevée est adoptée pour les zones texturées (contours). L'image est ainsi constituée de la superposition d'une image basse résolution, basse qualité, bas débit, et d'une image de texture.

Pour séparer les différentes zones, une partition est créée à partir d'un découpage quadtree par blocs (de taille 2x2 jusqu'à 16x16 pixels) et d'un critère de gradient morphologique. L'image basse qualité est obtenue en remplissant chaque bloc par sa valeur moyenne.

Le LAR Interleaved S+P (LAR-iSP) [16] utilise une décomposition hiérarchique de l'image, comme présentée sur la figure 1. Cette décomposition autorise à la fois une scalabilité en résolution (déplacement vertical dans la pyramide) et une scalabilité en qualité (déplacement horizontal) pour un codage allant jusqu'au sans perte.

Le codeur LAR-iSP repose sur la construction ascendante d'une pyramide utilisant la transformée en S 1-D des pixels adjacents diagonalement. Le codage est réalisé en deux descentes de la pyramide, avec une prédiction à contexte enrichi (360 degrés) selon le concept introduit par Wu[17] et un codage de l'erreur de prédiction. La première descente fournit l'image basse résolution tandis que la seconde amène la texture. L'efficacité provient d'une prédiction à la fois inter- et intra-niveaux. Le contexte implicite issu de la connaissance de la partition quadtree permet de réduire le coût entropique. Ainsi, le codeur LAR-iSP est-il meilleur que l'état de l'art de la compression sans perte, en particulier pour les images médicales[18].

## 3 Expansion de la différence

Cette section présente l'expansion de la différence (ED), une méthode introduite par Tian [2] qui insère un bit par paire de pixels S-transformés.

Les pixels sont d'abord appariés suivant un schéma prédéterminé et la transformée en S est appliquée sur



Figure 1: Scalabilités en résolution et en qualité pour le LAR-iSP

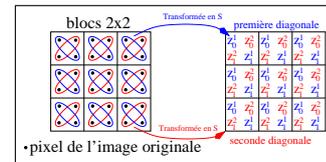


Figure 2: La transformation Interleaved S+P.  $z_0^i$  et  $z_1^i$  sont respectivement la moyenne et le gradient de la  $i^{\text{e}}$  diagonale.

chaque paire obtenue. Un bit  $b$  de donnée est inséré dans le gradient  $h$  tel que  $h' = 2 \times h + b$ . Le gradient modifié  $h'$  est ensuite utilisé (à la place de  $h$ ) dans la transformée en S inverse produisant les deux pixels marqués. Cette opération est appelée ED. Pour éviter les problèmes de dépassement de capacité dans la transformation inverse, et pour augmenter la capacité d'insertion, les gradients sont répartis en trois ensembles :

1. les gradients *expansibles* peuvent être marqués par l'ED ;
2. les gradients *modifiables* qui ne peuvent être modifiés par l'ED mais peuvent incorporer un bit en sauvegardant leur bit de poids faible et en le remplaçant par  $h' = 2 \times \lfloor \frac{h}{2} \rfloor + b$  ;
3. les gradients *non modifiables* qui ne sont pas utilisables du tout.

De plus, afin d'avoir un système amenant un minimum de distorsions, certains gradients *expansibles*, qui sont de facto *changeables* sont changés plutôt qu'expansés. Pour obtenir la charge utile désirée, un algorithme itératif

détermine les paires de pixels qui seront expansées ou changées. Cet algorithme utilise une carte binaire, la carte des positions, créée pour connaître les positions des gradients expansibles, modifiables ou non modifiables.

Cette technique fournit parmi les meilleures performances en terme de capacité-distorsion présentes dans la littérature, avec le bénéfice d'être complètement réversible. De fait, l'image originale est récupérée sans distorsion grâce à la carte de positions incluse dans les données insérées. L'ED permet d'insérer au plus 0.5 bpp (1 bit par paire de pixels) et pour augmenter cette valeur, une extension de la méthode concerne l'ED hiérarchique, qui procède par insertion successive avec l'ED ordinaire.

## 4 Insertion de données dans le LAR-iSP

Le LAR-iSP et l'ED utilisent tous deux la transformée en S. L'idée majeure est de réaliser l'insertion de données sans dégrader les performances de compression. L'ajustement de l'ED au LAR-iSP se traduit par l'introduction de quelques modifications mineures de la méthode ED originale présentée à la section 3.

D'abord la transformée en S est appliquée sur la première et la seconde diagonale des blocs 2x2 partitionnant l'image, comme présenté sur la figure 2 ; l'ED est effectuée sur les valeurs  $z_1^i$  résultantes. Les coefficients sont choisis pour leur appartenance à des blocs de taille donnée, conformément à la partition quadtree et dépendent donc du contenu de l'image même. La première diagonale est d'abord marquée, puis la seconde, si la charge nécessite une capacité accrue.

En outre, dans l'ED originale, la séparation des coefficients expansibles se fait de manière itérative. Pour accélérer cette étape, tous les gradients expansibles sont expansés, et, par l'usage de marqueurs terminaux insérés dans le flux binaire, le décodeur détermine la position des gradients expansés.

## 5 Résultats expérimentaux

Toutes les expériences menées consistent en l'insertion soit d'une charge utile de 1024 bits, soit d'une charge utile aussi grande que possible. La charge est une séquence binaire pseudo-aléatoire générée par la fonction  $C_{rand}()$  pour différentes graines. La section 5.1 se concentre sur l'influence de la taille des blocs marqués dans l'image usuelle *lena* 8 bpp, 512x512. L'évaluation des performances en terme de capacité-distorsion est également faite. La section 5.2 évalue les performances en terme de qualité visuelle tandis que la section 5.3 s'intéresse au coût supplémentaire de codage induit par l'insertion de données.

## 5.1 Performances de capacité-distorsion

Le tableau 5.1 présente les capacités disponibles pour différentes tailles de blocs marqués. Il faut noter que la charge utile maximale, en utilisant les pixels des blocs 2x2, est plus petite que celle obtenue pour les autres tailles de blocs, alors que la capacité est plus grande. Ce paradoxe apparent s'explique par les différences de taille des cartes de positions compressées. La mesure de cette différence constitue le surcoût lié à l'ED.

taille bloc	2	4	8	16	4-16
capacité max (bits)	34816	30848	32640	32768	96256
surcoût ED (bits)	7676	994	1042	1170	3042
surcoût ED (bpp)	0.03	0.004	0.004	0.004	0.01
Charge utile max (bits)	27142	29854	31598	31710	93214
Charge utile max (bpp)	0.10	0.11	0.12	0.12	0.36
PSNR (dB)	32.93	43.41	46.58	48.60	40.91

Tableau 1: Charge utile insérée, capacité et PSNR en fonction de la taille des blocs marqués pour *lena*.

La figure 3 présente les courbes capacité-distorsion pour le LAR+ED, l'ED originale (équivalente au LAR+ED effectué pour tous les blocs), l'ED hiérarchique et la méthode G-LSB [19]. LAR+ED est considéré pour plusieurs tailles de blocs. On peut remarquer qu'il y a davantage de distorsions lorsque seuls les blocs 2x2 sont marqués. Ceci est confirmé par des expérimentations sur d'autres images. Pour les autres tailles de blocs, LAR+ED est meilleur que G-LSB et parfois meilleur que l'ED hiérarchique (itérative), principalement pour les petites charges utiles. Dans cette configuration, l'ED originale ne fournit pas de bons résultats.

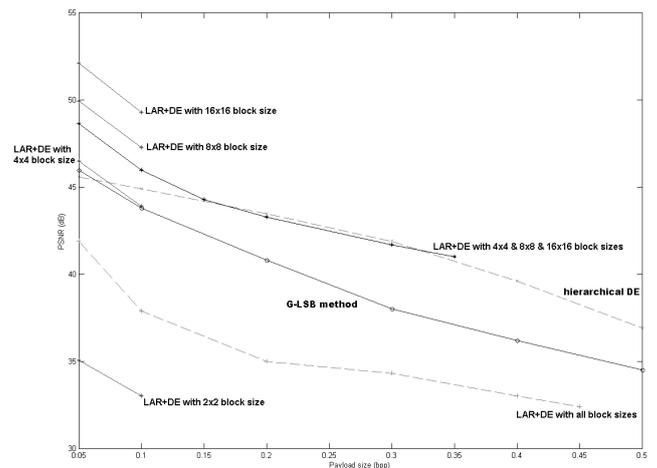


Figure 3: Distorsion en fonction de la capacité pour *lena*.

## 5.2 Qualité visuelle

La figure 4 montre les images obtenues après insertion de la charge utile maximale, en utilisant des paires de pixels diagonaux dans les blocs de différentes tailles. La qualité visuelle est bonne dans l'ensemble. Cependant, un examen minutieux des images indique que les artefacts sont davantage visibles lorsque l'insertion se fait sur les blocs 2x2, comme le montre la figure 5b. Voici quelques éléments d'explication.



Figure 4: Évaluation de la qualité visuelle après insertion réversible de la charge utile maximale dans l'image *leba* pour différentes tailles de blocs

Le système visuel humain est moins sensible aux basses et hautes fréquences spatiales. La modification des composantes hautes fréquences peut donc se faire dans de bonnes conditions d'invisibilité. Cependant, des hautes fréquences

impliquent des contours, dont les pixels sont très pertinents dans la définition du contenu sémantique de l'image. Ceci explique la présence de davantage d'artefacts dans les blocs 2x2 que dans les autres : l'ED modifie les contours de manière inappropriée, comme le montre la figure 5d. Cet artefact introduit une incohérence inter-blocs. Son effet est d'inverser localement l'orientation du gradient, et d'introduire ainsi de nouveaux contours visibles. Néanmoins, les blocs 2x2 peuvent continuer à être marqués, pour atteindre de bonnes performances de capacité-distorsion, à condition de sélectionner les blocs à marquer en fonction de leur voisinage.

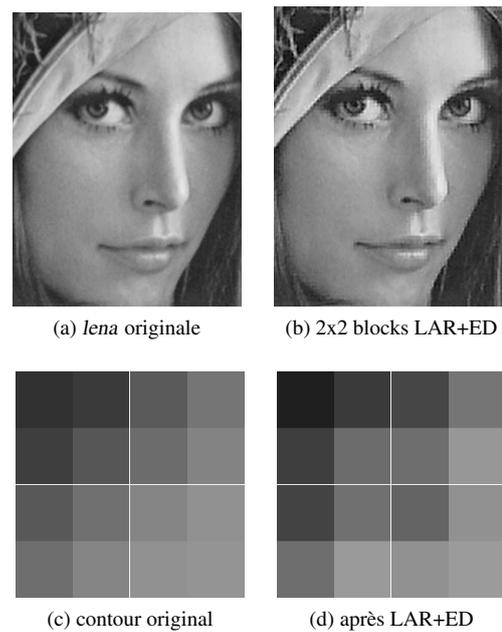


Figure 5: Artefact d'incohérence inter-blocs

## 5.3 Coût de codage

L'évaluation de l'impact de l'insertion de données sur les performances de compression passe par le calcul de la taille du flux binaire nécessaire à coder sans perte les différentes images. Le tableau 2 résume les résultats obtenus. La charge utile étant issue d'une séquence binaire pseudo-aléatoire, son coût de codage minimal correspond à sa longueur. De même, il faut également prendre en compte le surcoût de l'ED lié à la carte des positions.

La caractéristique majeure du tableau 2 concerne les surcoûts liés au codage, c'est-à-dire le coût engendré par l'insertion de données. Ajouté au surcoût lié à l'ED, il donne le surcoût total du LAR+ED. Le surcoût de l'ED ayant été vu à la section 5.1, seul le surcoût lié au codage est ici discuté.

Le surcoût de codage reflète principalement l'adaptation (ou non) du prédicteur aux nouvelles valeurs des gradients modifiés par l'insertion de données ED. Le tableau 2 montre que le surcoût de codage est directement lié à la

taille bloc	charge utile [bpp]	surcoût ED [bpp]	coût codage [bpp]	surcoût codage [bpp]	surcoût codage [%]
2	0.10	0.03	4.50	0.06	60
4	0.11	0.004	4.51	0.08	73
8	0.12	0.004	4.51	0.08	67
16	0.12	0.004	4.49	0.07	58
4–16	0.36	0.01	4.73	0.05	14

Tableau 2: Coût de codage pour le LAR+ED. L'entropie brute de l'image originale compressée par le LAR-iSP s'élève à 4.31 bpp.

charge utile. Plus nombreux sont les pixels marqués, ou, de manière équivalente, plus nombreux sont les gradients modifiés, meilleur sera le prédicteur. De fait, la distribution des gradients s'uniformise et permet au prédicteur de mieux fonctionner. Ainsi, l'insertion de 0.36 bpp de données n'induit-elle un surcoût de codage de 0.05 bpp, soit seulement 14 % de la charge utile.

## 6 CONCLUSION

Cet article a présenté l'adaptation d'une méthode réversible d'insertion de données rapide et efficace (l'ED) dans le schéma de compression scalable LAR+iSP. Les deux méthodes exploitent la redondance présente dans les images numériques pour obtenir de meilleurs résultats que l'état de l'art tant en compression sans perte qu'en insertion de données. Les résultats obtenus en termes de capacité-distorsion pour l'insertion de données sont parmi les meilleurs présentés dans la littérature concernant l'insertion sans perte, tandis que parallèlement, le surcoût de codage de cette insertion se maintient dans des limites très raisonnables.

## Références

- [1] M. Babel, B. Parrein, O. Déforges, N. Normand, J. P. Guédon, et J. Ronsin. Secure and progressive transmission of compressed images on the internet : application to telemedicine. Dans *SPIE*, pages 126–136, 2005.
- [2] J. Tian. Reversible data embedding using a difference expansion. Dans *IEEE Transactions on Circuits and Systems for Video Technology*, pages 890–896, august 2003.
- [3] W. Bender, D. Gruhl, N. Morimoto, et A. Lu. Techniques for data hiding. *IBM Systems*, 1996.
- [4] P. Bas, J-M. Chassery, et F. Davoine. Using the fractal code to watermark images. Dans *Proc. of the Int. Conf. on Image Processing*, volume 1, pages 469–473, 1998.
- [5] J. J.K. Ó Ruanaidh et T. Pun. Rotation, scale and translation invariant digital image watermarking. Dans *Proc. of the Int. Conf. on Image Processing*, pages 536–539, 1997.
- [6] M. H. Hayes. The reconstruction of a multidimensional sequence from the phase or magnitude of the FFT. Dans *IEEE Transactions on Acoustics, Speech and Signal Processing*, pages 140–154, april 1992.
- [7] S. Asif Mahmood Gilani et A. N. Skodras. Dlt-based digital image watermarking. Dans *IEEE Balkan Conference on Signal Processing, Communications, Circuits and Systems*, 2000.
- [8] S. Kang et Y. Aoki. Digital image watermarking by Fresnel transform and its robustness. Dans *Proc. of the Int. Conf. on Image Processing*, volume 2, pages 221–225, 1999.
- [9] J. Tian et R. O. Wells. Reversible data-embedding with a hierarchical structure. Dans *Proc. of the Int. Conf. on Image Processing*, pages 3419–3422, october 2004.
- [10] J. Zhao et E. Koch. Embedding robust labels into images for copyright protection. Dans *Proc. of the Int. Congress on Intellectual Property Rights for Specialized Information, Knowledge and new Technologies*, Vienne, august 1995.
- [11] M. Barni, F. Barolini, V. Cappellini, et A. Piva. A DCT-domain system for robust image watermarking. *Signal Processing*, pages 357–372, may 1998.
- [12] I. J. Cox, J. Kilian, T. Leighton, et T. Shamoan. Secure spread-spectrum watermarking for multimedia. Rapport technique, NEC Research Institute, 1995.
- [13] X-G. Xia, C. G. Boncelet, et G. R. Arce. A multiresolution watermark for digital images. Dans *Proc. of the Int. Conf. on Image Processing*, 1997.
- [14] A. Said. Wavelet-based image compression. Dans *Imaging Technology Department*. Hewlett-Packard Laboratories, 1999.
- [15] M. Babel, O. Déforges, et J. Ronsin. Adaptive Multi-Resolution Scheme for Efficient Image Compression. Dans *Proc. of Picture Coding Symposium*, april 23-25 2003.
- [16] M. Babel, O. Déforges, et J. Ronsin. Interleaved S+P pyramidal decomposition with refined prediction model. Dans *Proc. of the Int. Conf. on Image Processing*, 2005.
- [17] X. Wu. Lossless compression of continuous-tone images via context selection, quantization and modeling. Dans *IEEE Transactions on Image Processing*, pages 656–664, 1996.
- [18] M. Babel. *Compression d'images avec et sans perte par la méthode LAR Locally Adaptive Resolution*. Thèse de doctorat, INSA Rennes, september 2005.
- [19] M. U. Celik, G. Sharma, A. M. Tekalp, et E. Saber. Reversible data hiding. Dans *Proc. of the Int. Conf. on Image Processing*, pages 157–160, 2002.