

Compression Fractale basée- région des séquences vidéo segmentées¹

K. BELLOULATA² et F. ROSSIGNOL

² Département de Génie électrique et de Génie Informatique
Université de Sherbrooke, Sherbrooke, QC, Canada, J1K 2R1
Tel/Fax: +1 (819) 821-8000/7937

kamel.belloulata@courrier.usherb.ca

Résumé – Nous proposons une nouvelle méthode de compression basée région, des séquences vidéo, basée sur le codage fractal. Chaque objet de la séquence est individuellement codé par un système de fonctions itératives locales (SFIL) approprié. L'algorithme utilise une prédiction circulaire entre les quatre (4) premières trames de la séquence vidéo (Circular Prediction Mapping -CPM-) et des transformations affines non contractantes (Non Contractive Interframe Mapping -NCIM-) pour les trames suivantes. Le CPM et NCIM combinent le codage fractal avec l'algorithme d'estimation et de compensation de mouvement (ME/MC), en exploitant la corrélation temporelle entre les trames adjacentes. Pour introduire la fonctionnalité basée - région, chaque région de la séquence doit être codée indépendamment en accord avec une carte de segmentation. Ceci permet une transmission/stockage indépendant et, par conséquent, un décodage individuel des régions. Les blocs (source/cibles) peuvent donc contenir plus qu'un segment. Dans ce scénario, chaque segment dans un bloc cible peut être modélisé par un segment d'un bloc source du même objet. Notre codeur fractal basé - région des images fixes pourra générer des résultats plus intéressants dans le contexte vidéo. Cependant, cet algorithme sera utilisé pour coder des objets définis par une carte de segmentation, exactement comme dans MPEG-4. Il a été prouvé par simulation sur une station de travail que l'algorithme proposé produit des images de qualité visuels intéressantes pour la vidéo conférence à un débit très faible, inférieure à 64 kbps.

Abstract – We propose a new fractal video compression scheme with region-based functionalities. Each object in the video sequence is individually encoded using a local system of iterated functions (fractals -LIFS-). Our algorithm encodes the first four (4) frames of a video sequence as a "set" using the Circular Prediction Mapping (CPM) method then it encodes the remaining frames using the Non Contractive Interframe Mapping (NCIM) method. The CPM and NCIM methods combine fractal coding with the Motion-Estimation/Motion Compensation (ME/MC) algorithm by exploiting the high temporal correlation between adjacent frames. To introduce the region-based functionalities, we encode each region independently using an Alpha Plane segmentation chart exactly the same as MPEG-4. This allows independent transmission/storage and hence independent decoding of each region in the sequence. Domain and range blocs may now contain more than one segment. In this scenario, each segment of a range block may be match using a segment of a domain block belonging to the same region (object).

1. Codage fractal des séquences vidéo

Nous proposons une nouvelle méthode de compression basée région, des séquences vidéo, basée sur le codage fractal. Cette approche tient compte de l'auto similarité naturelle existante dans chaque objet dans une scène naturelle. Cette nouvelle approche produira la fonctionnalité basée - région, pour des séquences vidéo, en accord avec une carte de segmentation prédéfinie [7]. La séquence originale est segmentée en plusieurs régions selon la forme des objets qui y sont retrouvés. Chaque objet est codé par un système de fonctions itératives locales (SFIL) approprié en utilisant un partitionnement en quadtree [10] pour les blocs sources et les blocs cibles. Cet algorithme utilise une prédiction circulaire entre les trames de la séquence vidéo (Circular Prediction Mapping -CPM-) et des transformations affine non contractantes (NonContractive Interframe Mapping -NCIM-) [3]. Le CPM et le NCIM combinent le codage fractal avec l'estimation et la compensation de mouvement (ME/MC), en exploitant la corrélation temporelle entre les trames

adjacentes (Fig. 1); la majeure différence par rapport aux autres schémas de codage vidéo par fractale [11], [12]. Dans les deux schémas CPM et NCIM, chaque bloc cible est compensé par un bloc source, de même taille, à partir de la trame précédente [3].

Pour produire la fonctionnalité basée - région, les images doivent être codées région par région en accord avec une carte de segmentation [6]. En utilisant des blocs sources et des blocs cibles carrés, nous pouvons partager les blocs de contours d'objets en segments qui appartiennent à des différentes régions (objets).

Cette approche basée - région est appliquée dans le domaine spatial (pixel) pour coder les séquences d'images en utilisant le schéma hybride CPM et NCIM. Elle peut être aussi appliquée dans le domaine fréquentielle [5].

Au niveau du décodeur, l'ensemble des transformations affines est appliqué itérativement sur une image arbitraire pour converger vers un point fixe qui s'approchera le plus de l'image originale [1]. Si les transformations affines peuvent

¹ Ce travail est supporté par le CRSNG (Conseil de Recherches en Sciences Naturelles et en génie du Canada) sous la subvention numéro 238966-01.

être représenté par moins de bits que l'image originale, un gain en taux de compression sera obtenu [1], [9], [10]. Dans le schéma CPM, les quatre premières trames sont considérées comme un groupe initial et chaque trame est prédite, d'une manière circulaire, à partir de celle qui précède.

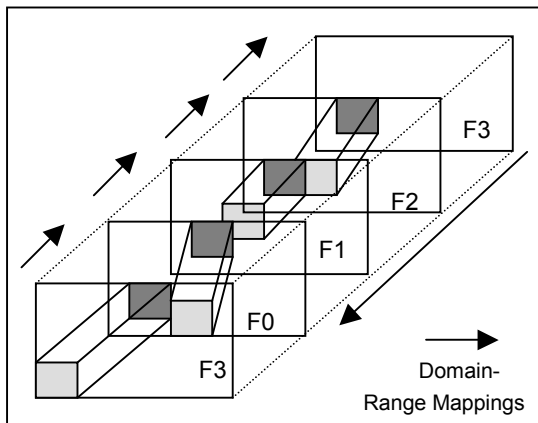


FIG.1 : structure du CPM

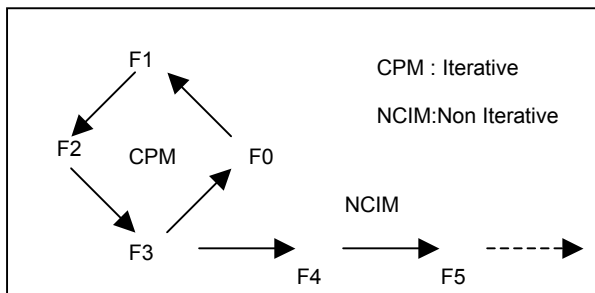


FIG.2 : schéma hybride CPM et NCIM

Dans un autre terme, les 4 trames T_k sont partitionnés en blocs cibles R_i , et chaque bloc cible R_i dans T_k est modélisé par un bloc source $D_{a(i)}$ à partir de la trame précédente $T_{[k-1]4}$, tel que $[k]4$ est le modulo (k modulo 4). Par exemple, la trame T_0 est codée à partir de la trame T_3 et la trame T_1 est codée à partir de la trame T_0 . La taille des blocs sources est la même que la taille des blocs cibles, ce qui est différent par rapport à l'algorithme de Fisher [2]. Vu que les séquences d'images vidéo contiennent plus de corrélation, l'appariement entre blocs sources et blocs cibles deviendra plus efficace et plus naturel si la taille des blocs sources est la même que la taille des blocs cibles.

Dans ce cas, l'appariement source-cible est semblable à la technique de compensation de mouvement. Par conséquent, $a_{(i)}$, le vecteur mouvement, décrit la translation spatiale entre le bloc source et le bloc cible. En plus du mouvement, le changement en contraste s_i (facteur d'échelle) et en décalage de niveau de gris o_i sont aussi compensés. Dans ce contexte, l'avantage majeur de cet appariement source-cible est que dans un mouvement réel de séquences d'images, les petits déplacements sont plus probables que les grands déplacements. Par conséquent, la région de recherche pour le vecteur de mouvement $a_{(i)}$ peut être localisée au voisinage du bloc cible (à l'intérieur du même objet à coder par exemple). En plus, ce vecteur de mouvement peut être codé avec moins

de bits que dans le cas d'une recherche globale (sur toute la trame). Au niveau du décodeur, le CPM est appliqué itérativement sur une séquence arbitraire pour converger vers une séquence « attracteur ». Donc, le CPM doit être contractant pour que le processus itératif converge. Le système à fonctions itératives est contractant lorsque le facteur d'échelle s_i est forcément maintenue entre -1 et 1.

Les quatre premières trames de la séquence sont codées en appliquant le CPM, et le reste des trames est codé par le NCIM. La structure du NCIM est semblable à celle du CPM, sauf qu'il n'y a pas de contrainte sur le facteur d'échelle s_i . Par conséquent, nous pouvons exploiter la corrélation temporelle plus efficacement avec le NCIM, ce qui donnera un gain important en codage. Au niveau du décodeur, les quatre premières trames sont reconstruites en appliquant itérativement le CPM. Alors que le reste de la séquence est reconstruit par l'application du NCIM à partir de la trame reconstruite précédemment, et ce avec une seule itération; le NCIM n'est donc pas contractant. Les quatre premières trames, codées par le CPM, représentent un ensemble minimal décodable [4], dans le sens où elles ne dépendent pas d'autres trames pour se coder/décoder (Fig. 2). Donc, seulement le CPM affectera le processus global de convergence du décodeur fractal. Le NCIM ne dépendra que de la dernière trame décodée de la séquence initiale de 4 trames du CPM.

2. Fonctionnalités basées région

Dans un schéma classique de codage par fractal, le partitionnement de l'image est indépendant de son contenu, et les blocs cibles peuvent contenir des pixels qui font partie de différents objets (régions). Notre approche, pour le codage fractal basé – région des objets vidéo, nécessite une carte de segmentation (α -paln) définie en aval (méthodes de deuxième générations [8]) pour coder chaque région indépendamment. Ceci permet une transmission/stockage indépendant et, par conséquent, un décodage individuel des régions, tout en offrant une nouvelle fonctionnalité [6]. Les blocs (source/cibles) peuvent contenir plus d'un segment; et chaque segment dans un bloc cible peut être modélisé par un segment d'un bloc source du même objet. Pour le bloc source, les pixels indéfinis (pixels n'appartenant pas au segment à coder) sont extrapolés (Padding). Ces pixels sont remplacés par la valeur moyenne des pixels définis du segment (Fig. 3). La distorsion entre le bloc cible et le bloc source, correspondent, est évaluée seulement sur les pixels du segment du bloc cible concerné. Notons que, lorsque les deux blocs, cible et source, font partie de l'intérieur de l'objet à coder, la distorsion se simplifiera au cas standard (en tenant compte de tous les pixels du bloc cible).

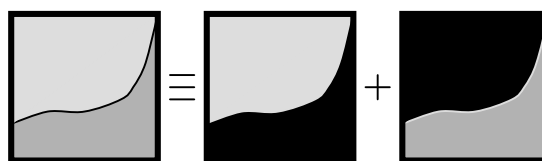
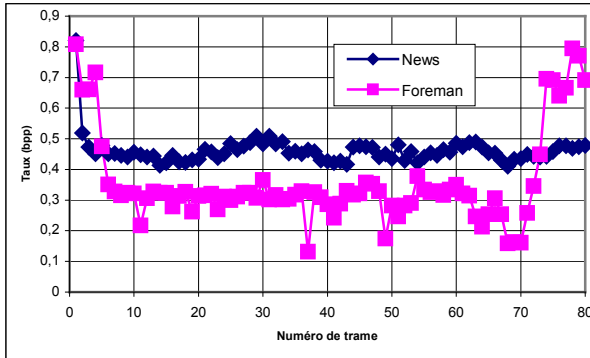


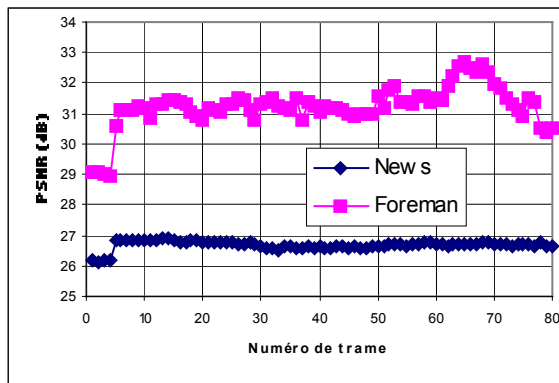
FIG. 3 : bloc avec deux segments S^1 et S^2 , appartenant à différentes régions, et sa décomposition en deux blocs avec des pixels actifs (gris) dans S^1 ou dans S^2 , et inactifs (noir).

3. Résultats

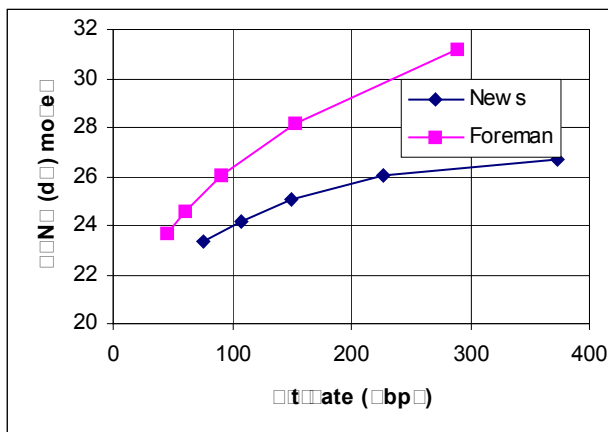
Le codeur fractal hybride CPM+NCIM (basé-région et non basé-région) a été testé sur des séquences vidéo réelles de format CIF (352 x 288) à niveaux de gris; avec un taux de rafraîchissement original de 25 trames/s. Mais ce taux de rafraîchissement a été réduit à 8.33 trames/s pour démontrer la performance à des débits très faibles, pour des applications en visiophonie. En d'autres termes, chaque troisième trame est codée et les deux autres sont ignorés.



a) Débit binaire pour les différentes trames



b) Le Rapport Signal/Bruit pour les différentes trames



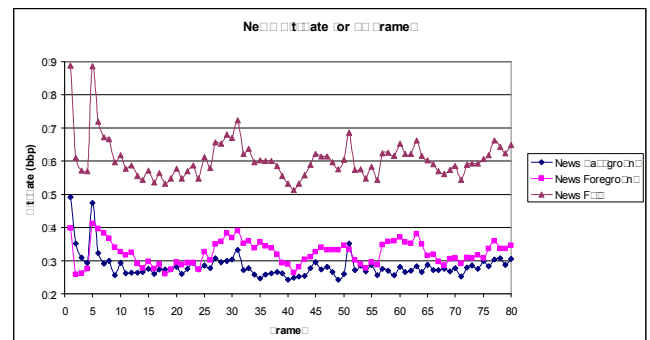
c) Le R.S.B moyen pour les différents débits binaires sur les 80 trames (En fait, 238 trames)



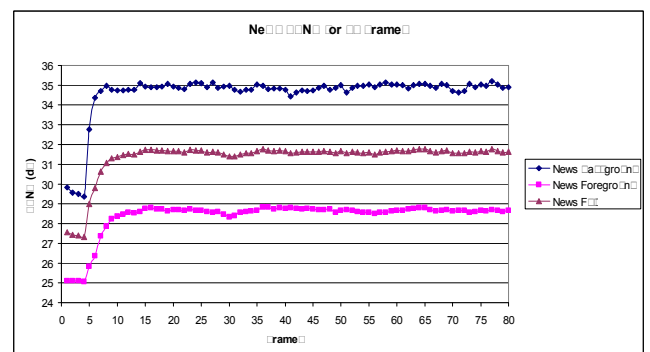
d) La trame n° : 20 reconstruite ;0.4 bpp et R.S.B=26.76dB

FIG. 4 : différents résultats obtenus pour le codeur hybride (CPM+NCIM [10]) non basé-région.

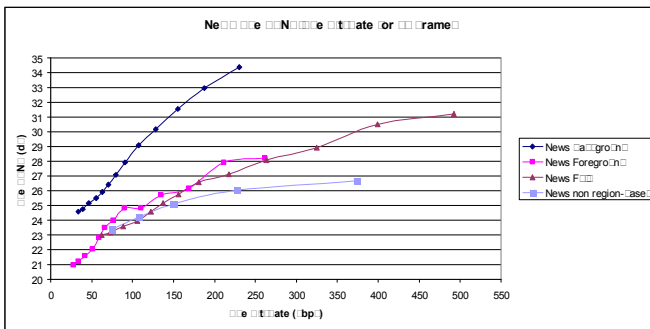
La figure. 4 montre les performances de l'algorithme hybride CPM+NCIM non basé-région en termes débit binaire/RSB sur les séquences News et Foreman. Vu que les deux séquences ne contiennent pas de changement dans les 1^{er} - 80^{ème} trames (originellement, 1^{er} - 238^{ème} trames), les quatre premières trames sont codées avec le CPM et le reste est codée avec le NCIM. Le RSB moyen et le débit moyen dans c) est obtenu en calculant la moyenne sur l'ensemble des 80 trames utilisées (équivalent aux 238 trames).



a) Débit binaire pour les différentes trames



b) Le Rapport Signal/Bruit (R.S.B) pour les différentes trames



f) Le fonds (Background) ; 0.26 bpp et R.S.B 34.86 dB

FIG. 5 : différents résultats obtenus pour le codeur hybride (CPM+NCIM) basé-région.

La Fig. 5 montre aussi les performances de l'algorithme hybride CPM+NCIM basé-région, en termes de débit binaire/RSB sur la séquence News. Notre concept basé-région est bien prouvé dans cet ensemble de résultats. On peut envoyer les deux objets vidéo, codé séparément, en utilisant deux largeurs de bandes différentes sur un réseau hétérogène de télécommunications. Cette fonctionnalité basé-région nous permet donc d'aller vers des débits binaires très faibles.

Le RSB moyen et le débit binaire moyen dans c) sont obtenus en calculant la moyenne sur l'ensemble des 80 trames utilisées (équivalent aux 238 trames).

c) Le R.S.B moyen pour les différents débits binaires



d) La trame n° : 20 : 0.55 bpp et R.S.B 31.66 dB



e) L'objet (Foreground) ; 0.29 bpp et R.S.B 28.71 dB



Références

- [1] A. Jacquin, 'Image coding based on a fractal theory of iterated contractive image transformations,' IEEE Trans. Image Process., vol., pp. 18-30, Jan. 1992.
- [2] J. Fisher, T. P. Shen and D. Rogovin, 'Fractal (Self-VQ) Encoding of Video Sequences,' in Proc. SPIE: VCIP, vol. 2308, pp. 1359-1370, 1994.
- [3] C. S. Kim, R. C. Kim and S. U. Lee, 'Fractal coding of video sequence using circular prediction mapping and noncontractive interframe mapping,' IEEE Trans. Image Processing, vol. 7, Apr. 1998.
- [4] J. Domas and V. A. Vaishampayan, 'Graph-theoretical analysis of the fractal transform,' in Proc. ICASSP, 1995, vol. 4, pp. 2559-2562.
- [5] K. Belloulata, R. Stasinski, and J. Konrad, 'Region-based image compression using fractals and shape-adaptive DCT,' in Proc. IEEE Int. Conf. Image Processing, vol. II, pp. 815-819, Oct. 1999.
- [6] K. Belloulata and J. Konrad, 'Fractal Image Compression with Region-Based Functionality,' IEEE Trans. Image Process., vol. 11, pp. 351-362, April 2002.
- [7] Special issue on MPEG-4, IEEE Trans. Circuit Syst. Video Technol., vol. 7, Feb. 1997.
- [8] M. Kunt, A. Ikononopoulos and M. Kocher, 'Second-generation image-coding techniques,' Proc. IEEE, vol. 73, pp. 549-574, Apr. 1985.
- [9] B. Ohlberg and G. de Jager, 'A review of the fractal image coding literature,' IEEE Trans. Image Process., vol. 8, pp. 1716-1729, Dec. 1999.
- [10] J. Fisher, 'Fractal encoding with quadrees,' in Fractal Image Compression: Theory and Applications (J. Fisher, ed.), pp. 55-77, Springer-Verlag, 1995.
- [11] M. S. LaBar and L. T. Bruton, 'Fractal Block Coding of Digital Video,' IEEE Trans. Circuits and Systems for Video Technology, vol. 4, no. 3, pp. 297-308, Jun. 1994.
- [12] J. Li, M. Novak and R. Forchheimer, 'Fractal-based Image Sequence Compression Scheme,' Optical Engineering, vol. 32, no. 7, pp. 1588-1595, 1993.