

# Compression vidéo couleur basée sur le codage entropique du signe des coefficients en ondelettes en utilisant la compensation de mouvement

(1) *Wissal Hassen*, <sup>(1)</sup> *Mbainabeye Jérôme*, <sup>(1)</sup> *Nouredine Ellouze* et <sup>(2)</sup> *Christian Olivier*

<sup>(1)</sup> Unité de Recherche en Traitement du Signal, Traitement de l'Image et Reconnaissance de Formes  
Ecole Nationale d'Ingénieurs de Tunis, BP37, Le Belvédère, Tunis 1002, Tunisie  
{wissal\_hassen, mbai\_jerome }@yahoo.fr, n.ellouze@enit.rnu.tn

<sup>(2)</sup> Laboratoire XLIM, Département Signal, Images et Communications, Université de Poitiers, Bât. SP2MI,  
Téléport 2, Bvd Marie et Pierre Curie, BP 30179, 86962 Futuroscope Chasseneuil Cedex France.  
olivier@sic.sp2mi.univ-poitiers.fr

**Résumé :** La TOD<sup>1</sup> est un outil mathématique récent qui a stimulé beaucoup de développements dans plusieurs domaines scientifique et technique, particulièrement en Signal et Images. Son caractère multi-résolution en est une force majeure. La TOD utilisée en compression d'images fixes a permis la réalisation de la norme JPEG2000. Cependant, la gestion de mouvement dans la compression vidéo utilisant la TOD est une difficulté majeure due aux algorithmes de décomposition utilisée par la TOD, et ceci est une des raisons qui expliquent l'absence à l'heure actuelle d'une norme de compression vidéo basée sur la TOD. Cet article présente notre contribution en compression vidéo couleur utilisant la TOD et propose un nouveau schéma basé sur le codage séparé du signe et de l'amplitude des coefficients en ondelettes appelé SSC-MC<sup>2</sup>. Une évaluation du schéma proposé a été opérée sur un certain nombre de séquences vidéo standard en termes de mesure objective de PSNR<sup>3</sup>, de SSIM<sup>4</sup> et de VQM<sup>5</sup>. Les résultats obtenus sont comparés aux codeurs SSC sans compensation de mouvement, Motion JPEG2000 et MPEG2. Il est montré que le codeur proposé est compétitif aux codeurs ci-dessus et en améliore leur performance en termes de mesures objectives et subjectives.

**Mots clefs :** Compression vidéo, SSC-MC, compensation de mouvement, block-matching, ondelettes.

## 1. Introduction

L'objectif de la compression vidéo est de réduire autant que possible les redondances spatiale et temporelle et de la représenter avec les informations essentielles sans que l'oeil humain ne détecte les dégradations engendrées. La réduction de la redondance spatiale est obtenue généralement par une transformation orthogonale qui exploite la corrélation intra image ; c'est le cas de la TCD<sup>6</sup> utilisée dans les schémas de compression d'images fixes (norme JPEG) et vidéo (norme MPEG), la TOD utilisée en compression d'images fixes (norme JPEG2000). La réduction de la redondance temporelle ne concerne que la vidéo et est obtenue par différentes techniques permettant de prédire une image dans une séquence vidéo en utilisant les informations passées, présentes ou futures. Dans les normes MPEG (MPEG1 et MPEG2), l'utilisation de la TCD permet de gérer sans difficulté majeure le mouvement dans les séquences vidéo car chaque image de la séquence est divisée en blocs et une compensation en mouvement est opérée. Pour les hauts débits, ces normes donnent des qualités vidéo

---

<sup>1</sup> Transformée en Ondelettes Discrète

<sup>2</sup> Separated Sign Coding and Motion Compensation

<sup>3</sup> Peak Signal to Noise Ratio

<sup>4</sup> Structural SIMilarity

<sup>5</sup> Video Quality Metric

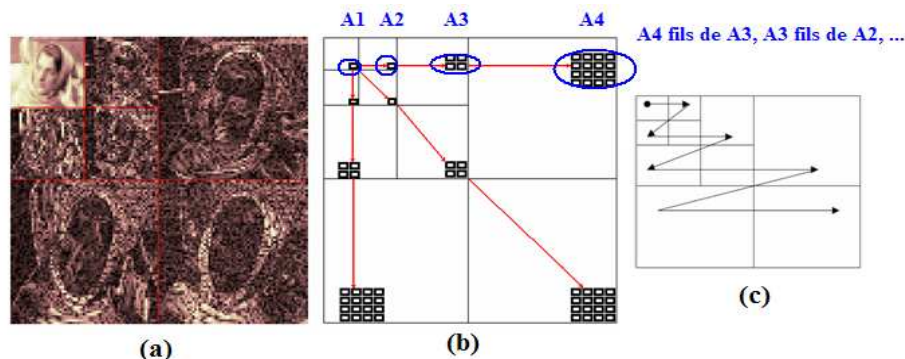
<sup>6</sup> Transformée en Cosinus Discret

satisfaisantes ; cependant les effets de blocs apparaissent lorsque la compression à bas débits est envisagée. La TOD a montré ses performances par rapport à la TCD et est adoptée dans la norme JPEG 2000 [TAU00, TZ94, ZA05, Wu01]. Si les schémas de compression vidéo utilisant la TOD sont proposés ça et là, le déficit majeur est la gestion de mouvement dans les séquences vidéo car non seulement l'image est décomposée dans sa taille initiale (pas de division en blocs) mais le sous échantillonnage appliqué dans la décomposition en ondelettes ne facilite pas la tâche et ceci explique l'absence à l'heure actuelle d'une norme de compression vidéo basée sur la TOD.

Dans [JER 06], l'auteur a proposé un schéma de compression d'images fixes basé sur le codage séparé du signe et l'amplitude des coefficients en ondelettes appelé SSC<sup>1</sup>. Ce codeur a donné des résultats compétitifs en termes de qualités objective et subjective avec la norme JPEG2000 et en améliore cette dernière. Compte tenu de cette performance, nous proposons dans cet article un schéma de compression vidéo couleur utilisant le SSC. En effet pour chaque GOP<sup>2</sup>, l'image de type I (image en mode intra) ainsi que les images de type P (images en mode prédiction ou images résiduelles) sont codés par le SSC sans compensation de mouvement. Les résultats obtenus sur un certain nombre de séquences vidéo standard montrent une bonne qualité de la séquence codée mais celle-ci présente une saccade lorsqu'on la visualise à l'écran. Afin de résoudre ce problème, nous avons opéré une compensation de mouvement dans le SSC en utilisant les algorithmes de compensation de mouvement tels que ceux intégrés dans la norme MPEG [DEB05, BOL05]. Le reste de l'article est organisé comme suit : dans la section 2, nous présentons une description du codeur SSC. La section 3 présente brièvement le concept de base de la compensation de mouvement. La section 4 présente le schéma de compression de la vidéo couleur proposé. La section 5 présente les résultats obtenus en termes de qualités objective et subjective. Enfin la section 6 présente la conclusion et les perspectives de ce travail.

## 2 Le codeur Separated Sign Coding (SSC)

L'idée du codeur SSC [JER 06] est d'exploiter le principe du codage en arbre de zéros (Zero Tree Coding) similaire au codeur des coefficients EZW<sup>3</sup> [SHO93]. La spécificité du codeur SSC est le codage séparé du signe des coefficients en ondelettes. Il utilise quatre symboles S, ZT, UZT et Z pour le codage des coefficients en ondelettes. Le symbole S est utilisé pour coder les coefficients significatifs c'est-à-dire supérieurs ou égaux à un seuil prédéfini ; ce symbole peut être assimilé à deux symboles POS et NEG utilisés par le EZW. Le symbole UZT est utilisé pour coder les coefficients inférieurs au seuil mais dont l'arbre contient au moins un coefficient significatif ; ce symbole peut être assimilé au symbole IZ utilisé par le EZW. Le symbole ZT est utilisé pour coder un arbre dont la racine et tous les descendants sont inférieurs au seuil ; ce symbole peut être assimilé au symbole ZTR utilisé par le EZW. Enfin le symbole Z est utilisé pour coder les coefficients non significatifs appartenant aux sous-bandes de détails au premier niveau de décomposition. Une carte est générée de manière progressive indiquant la présence ou l'absence du symbole S décrite par le symbole '1' ou par le symbole '0'.



**Figure 1.** (a) Décomposition par ondelettes de l'image Barbara sur deux niveaux de résolution, (b) arbres de coefficients, (c) ordre de balayage des coefficients dans le processus de codage des coefficients

<sup>1</sup> Separated Sign Coding

<sup>2</sup> Group Of Picture

<sup>3</sup> Embedded Zerotree of Wavelet



La figure 3 montre l'architecture du schéma de compression vidéo proposé où la figure 3A décrit le codeur et la figure 3B décrit le décodeur.

Le codeur (figure 3A) est décomposé en deux boucles dont chacune opère le traitement en fonction du type d'image dans la séquence vidéo : I, P ou B. La première boucle correspond au codage spatial de l'entête de chaque GOP (a) suivi de la reconstruction des images de référence I à partir du flux transmis. La deuxième boucle exploite la redondance temporelle en comparant une image à son prédécesseur (image P, ou Prédite) ou en comparant une image à la fois à son prédécesseur et à son successeur (image B, ou Bidirectionnelle). Le signal d'erreur de prédiction est ensuite codé par le SSC et transmis avec le vecteur mouvement (b). Une régulation du débit de sortie est gérée par le codeur proposé en utilisant un indicateur de débit calculé à partir de la taille souhaitée du fichier résultant et de la fréquence de la vidéo. Ainsi, on peut allouer plus de bits pour le codage des images de type I et moins de bits pour le codage des images de type P ou B.

Le décodeur (figure 3B) effectue les opérations inverses du codeur pour restituer la séquence vidéo. En effet le signal d'erreur et le vecteur de mouvement sont décodés et l'image actuelle de la séquence vidéo est restituée à partir de l'image précédente.

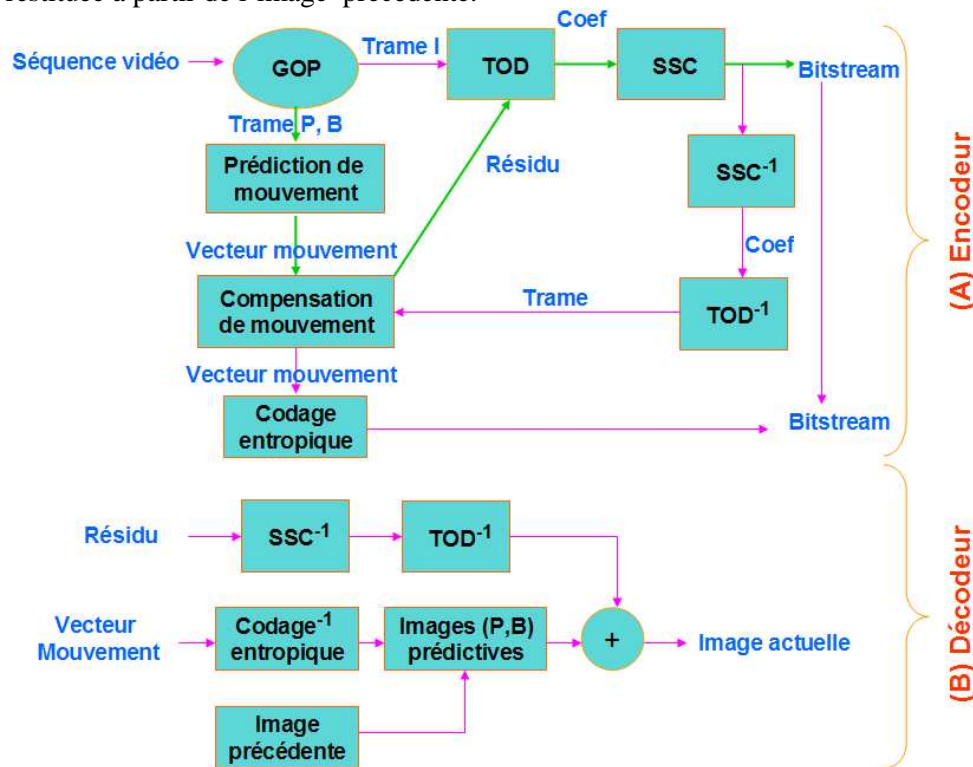


Figure 3. Architecture du schéma compression vidéo proposé : (A) Codeur, (B) Décodeur

#### 4 Résultats expérimentaux et discussions

Afin d'évaluer le schéma de compression proposé, une évaluation est effectuée sur trois séquences vidéo de test standard : la séquence Flower en format CIF, 240 x 352 x 3 pixels par image, 30Hz, 100 images, qui est caractérisée par un mouvement important de la caméra ; la séquence Akiyo en format QCIF, 176 x 144 x 3 pixels par image, 30Hz, 100 images, qui est caractérisée par une caméra plus ou moins stationnaire et la séquence Foreman en format CIF, 288 x 352 x 3 pixels par image, 30Hz, 300 images, qui est caractérisée par un mouvement brusque et rapide de la caméra . Ces séquences sont transformées de l'espace RVB (Rouge Vert et Bleu) vers l'espace YCbCr puis décomposées en ondelettes sur trois échelles en utilisant l'ondelette biorthogonale 9/7 de Daubechies et enfin elles sont codées par le schéma proposé. Les résultats sont comparés par rapport à ceux obtenus par les normes MPEG1, MPEG2et le MJPEG2000.

#### 4.1 Evaluation de la qualité objective

Pour évaluer la qualité objective du schéma proposé, nous avons utilisé le rapport signal sur bruit ou PSNR donné par l'équation (1) :

$$PSNR = 10 \cdot \log_{10} \left( \frac{255^2}{EQM} \right) \quad (1)$$

Où EQM est l'erreur quadratique moyenne calculée par l'équation (2) :

$$EQM = \frac{1}{HL} \sum_{i=1}^H \sum_{j=1}^L (I(i, j) - \hat{I}(i, j))^2 \quad (2)$$

Où H et L sont respectivement la hauteur et la largeur de chaque image de la séquence vidéo, I et  $\hat{I}$  sont respectivement l'image originale et l'image reconstruite. Il est à noter que les équations (1) et (2) sont appliquées séparément sur les composantes R, G et B de l'espace RGB. Le PSNR global relatif aux composantes R, G et B est calculé comme la moyenne des PSNR-R, PSNR-G et PSNR-B, soit  $(PSNR-R + PSNR-G + PSNR-B)/3$ .

#### 4.2 Evaluation de la qualité subjective

Pour la qualité subjective, nous avons utilisé deux outils de mesure : SSIM [ZCL 04] et VQM [ZHO 03].

Le SSIM permet de mesurer la qualité visuelle d'une image compressée ; en effet, l'idée de SSIM est de mesurer la similarité structurale entre les deux images, plutôt qu'une différence pixel à pixel comme le fait le PSNR. La métrique SSIM est calculée sur plusieurs fenêtres de la luminance (composante Y) de l'image. La mesure entre deux fenêtres x et y de taille N x N est donnée par l'équation (3) :

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + c_1)(2cov_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (3)$$

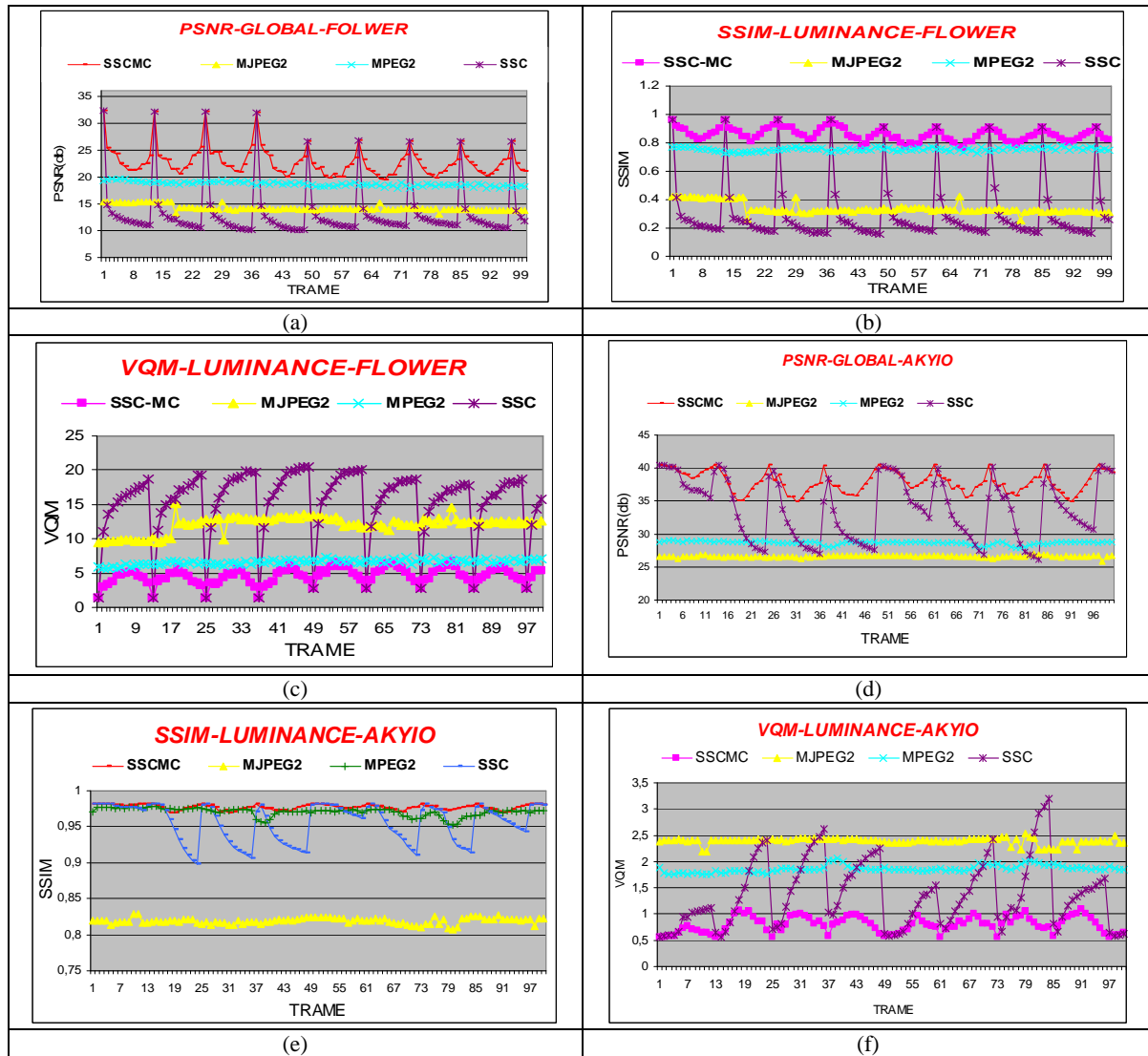
Où  $\mu_x$  et  $\mu_y$  sont les moyennes de x et y,  $\sigma_x^2$  et  $\sigma_y^2$  sont les variances de x et y,  $cov_{xy}$  est la covariance de x et y,  $c_1 = (K_1L)^1$  et  $c_2 = (K_2L)^2$  sont deux variables destinées à stabiliser la division quand le dénominateur est très faible, L est la dynamique des valeurs des pixels, soit 255 pour des images codées sur 8 bits et enfin  $K_1 = 0,01$  et  $K_2 = 0,03$  par défaut. Ainsi, plus la valeur de SSIM est voisine de 1, meilleur est le codeur.

Le VQM est basé sur la fonction spatio-temporelle humaine de sensibilité au contraste ; il exploite le fait que la perception humaine est moins sensible aux modèles spatio-temporels à haute fréquence en utilisant la TCD. Le VQM exploite également le fait que l'oeil humain ne perçoit pas une image entière en même temps mais plutôt les secteurs d'activité temporelle élevée où des déformations masquent les autres secteurs. Le VQM est estimé en mesurant la déformation moyenne et maximum d'une seule trame ; la valeur de VQM la plus proche de zéro indique alors une déformation minimum. Ainsi, plus faible est de la valeur de VQM, meilleur est le codeur.

Les figures 4 et 5 ainsi les tableaux 1 et 2 montrent les résultats en termes d'évaluation des qualités objective et subjective sur les séquences Flower et Akiyo.

La figure 4 et les tableaux 1 et 2 montrent que le schéma de compression vidéo donne des résultats quantitatifs meilleurs que ceux des codeurs SSC, MPEG2 et MJPEG2000. En terme de PSNR global, les gains obtenus par le codeur SSC-MC sont très significatifs ; en terme de SSIM, le codeur SSC-MC donne des valeurs proches de 1 ; en terme de VQM, le minimum est obtenu par le codeur SSC-MC.

Le zoom montré sur la figure 5 met en évidence les effets de blocs dus à la TCD utilisée dans les normes MPEG1 et MPEG2 ; ainsi, pour un même débit, la meilleure qualité visuelle est obtenue par le schéma proposé. La figure 6 montrent le décodage de quelques trames de la séquence Akiyo par le codeur proposé et les codeurs SSC, MJPEG et MPEG2 ; la trame N°1 est une trame de référence de type I, la trame N°17 est une trame de type B prédite à partir de I1 (image numéro 1 du GOP) et P4 (trame numéro 4 du GOP) du deuxième GOP de la séquence, la trame N°34 est la trame numéro 4 du troisième GOP de la séquence et elle est prédite à partir de la trame numéro 1 du troisième GOP de la séquence.



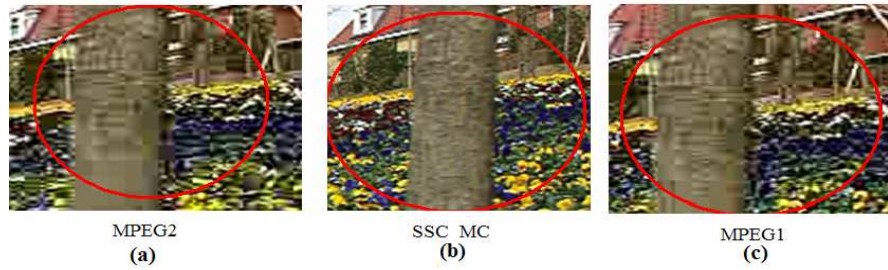
**Figure 4.** (a) PSNR-GLOBAL de la séquence Flower codées à 512 kps, (b) le SSIM pour la séquence Flower codées à 512 kps, (c) le VQM pour la séquence Flower codées à 512 kps, (d) PSNR-GLOBAL pour la séquence Akiyo codées à 128 kps, (e) le SSIM pour la séquence Akiyo codées à 128 kps, (f) le VQM pour la séquence Akiyo codées à 128 kps.

**Table 1.** Les valeurs moyennes de PSNR-GLOBAL, SSIM-Luminance et VQM-Luminance pour les codeurs SSC-MC, SSC, MPEG2, MJPEG2000 de la séquence Flower codée à 512kps.

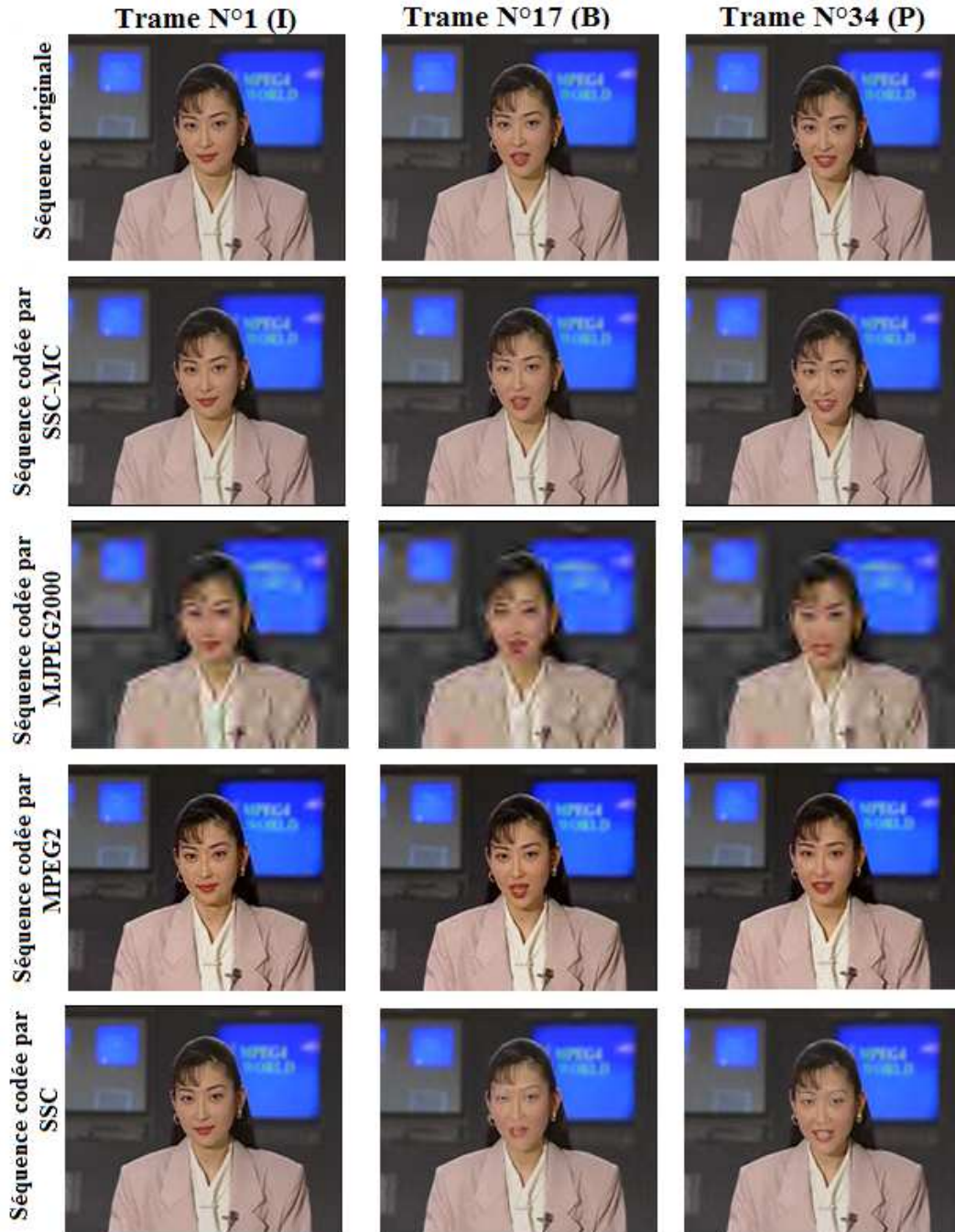
| Flower         | SSC        | SSC-MC            | MPEG2       | MJPEG2000  |
|----------------|------------|-------------------|-------------|------------|
| PSNR-GLOBAL    | 13.2801418 | <b>22.7780286</b> | 18.66086837 | 14.2054419 |
| SSIM-Luminance | 0.2955607  | <b>0.8575347</b>  | 0.7534581   | 0.3387478  |
| VQM-Luminance  | 15.5316475 | <b>4.5123322</b>  | 6.6058059   | 12.0658589 |

**Table 2.** Les valeurs moyennes de PSNR-GLOBAL, SSIM-Luminance et VQM-Luminance pour les codeurs SSC-MC, SSC, MPEG2 de la séquence Akiyo codée à 128kps

| Akiyo          | SSC         | SSC-MC             | MPEG2     | MJPEG2000   |
|----------------|-------------|--------------------|-----------|-------------|
| PSNR-GLOBAL    | 33.85289273 | <b>37,91280607</b> | 28.674637 | 26.60943517 |
| SSIM-Luminance | 0.9557758   | <b>0,9769988</b>   | 0.9706007 | 0.8189987   |
| VQM-Luminance  | 1.3712504   | <b>0,7934762</b>   | 1.8630239 | 2.390635    |



**Figure 5.** Zoom sur la trame 16 de la séquence Flower codées à 512 kps : (a) codée par MPEG2, (b) codée par SSC-MC, (c) codées par MPEG1.



**Figure 6.** Evaluation subjective : Résultats de codage de la séquence Akiyo à 128kps

## 5 Conclusion et perspectives

Nous avons présenté une approche de compression vidéo couleur basée sur le codage séparé des signes des coefficients en ondelettes et la compensation de mouvement. L'architecture du codeur et décodeur proposé est présenté. Une évaluation du schéma proposé est opérée sur deux séquences de test standard. Les résultats obtenus sont comparés à la norme MPEG2 et le codeur MJPEG2000. Ces résultats montrent que des gains significatifs peuvent être obtenus en terme de PSNR. Par ailleurs, nous avons montré par deux mesures subjectives que le codeur proposé améliore de manière significative la qualité visuelle.

Comme perspectives de ce travail, nous envisageons d'intégrer un indicateur qui permet de rendre compte de la dynamique de la séquence vidéo à coder ; ainsi il sera possible d'améliorer la performance du schéma proposé en terme de qualité visuelle de la séquence et en terme de débit de codage. On pourra alors allouer plus de bits pour les vidéos complexes (de forte dynamique) à coder et moins de bits dans le cas contraire. Il faut pour cela opérer deux passages, comme c'est le cas du codeur H264 : un premier passage pour analyser la dynamique de la séquence vidéo à coder (plus ou moins de mouvements) puis un deuxième passage pour coder la séquence vidéo en tenant compte de sa dynamique. Nous envisageons aussi d'augmenter le taux de compression en codant uniquement les vecteurs de mouvement des luminances des images P et B et d'améliorer la qualité en allouant plus de bits au codage de la luminance et moins de bits au codage des chrominances. Par ailleurs, nous envisageons d'utiliser d'autres méthodes de compensation de mouvement et aussi de comparer les résultats obtenus avec la norme H264.

## Références

- [JER05] M. Jérôme, N. Ellouze, "Optimal Image Compression Based on Sign and Magnitude Coding of Wavelet Coefficients", *International Journal of Signal Processing* Vol. 3, No 4, pp.243-251, 2006.
- [ZHO03] Y.Zhong, I. Richardson, A.Saharie et P.McGeorge, "Qualitative and quantitative assessment in video compression", 12<sup>th</sup> European Conference on Eye Movement, 20-24 August 2003, Dundee, Scotland.
- [MAR00] M.W. Marcellin, M.J. Gormis, A. Bilgin and M.P. Boliek, "An overview JPEG2000", *Proc. IEEE Data Compression Conference*, pp.523-541, 2000.
- [ISO93] ISO/IEC JTC1/SC22: Generic coding of moving pictures and associated audio- Séoul, Novembre 1993.
- [ZCL 04] Zhou Wang, Alan Conrad Bovik, Ligang Lu, "Video Quality Assessment Based on Structural Distortion Measurement", *IEEE Signal Processing: Image Communication*, Vol 19, No 2. pp. 121-132, February 2004.
- [SHO93] J.M.Shapiro, "Embedded Image Coding Using Zerotrees of Wavelet Coefficients", *IEEE trans. on Signal Processing*, Vol.41, No.12, pp.3445-3462, Dec. 1993.
- [RAI91] C. Raimondo, C. Galand, E. Goubault, E.Lançon et J.Menez, "Low bit-rate hybrid coding using hierarchical motion compensation and low complexity vector quantization", *Conf. Icassp*, Toronto, Mai 1991
- [TAU00] D. Taubman. "High performance scalable image compression with EBCOT", *IEEE Trans. on Image Processing*, Vol.9, No.7, pp.1158-1170, July 2000.
- [TZ94] D. Taubman and A. Zakhor. "Multirate 3-d subband coding of video", *IEEE Trans. on Image Processing*, Vol.3, No.5, pp.572-588, Sept., 1994.
- [DEB05] E. Debreuve, M. Gastaud, M. Barlaud, and G. Aubert, "A region-based joint motion computation and segmentation on a set of frames," in *Proceedings of European Workshop on Image Analysis for Multimedia International Services (WIAMIS)*, Montreux, Swiss, Apr. 2005.
- [BOL05] S. Boltz, E. Debreuve, and M. Barlaud, "A joint motion segmentation algorithm for video coding", in *Proceedings of EUSIPCO*, Antalya, Turkey, 2005.
- [SAI96] A. Said and W.A. Pearlman, "A new fast and efficient image codec based on Set Partitioning in Hierarchical Trees", *IEEE Trans. on Circuits and Systems for Video Technology*, No.6, pp. 243-250, June 1996.
- [BAR04] Barjatya A., "Block matching algorithms for motion estimation", 2004.
- [BHA97] Bhaskaran V., Konstantinides K., *Image and video compression standards*, USA, 1997
- [BAU99] Bauer S., Kneip J., Mlasko T., Schmale B., Vollmer J., "The MPEG-4 multimedia coding standard: algorithms, architectures and applications", Germany, 1999.
- [ZA05] A. Ouled Zaid, C. Olivier C., F. Marmiton, "Wavelet image coding with parametric thresholding: Application to JPEG2000", *SPIE'03- Electronic Imaging*, Vol. 5014, Santa Clara, California, USA, Jan. 2003.
- [Wu01] J.Wu, Christian Olivier and Christian Chatellier, "Embedded Zerotree runlength wavelet image coding", 8<sup>th</sup> ICIP-Signal Processing, Thessaloniki, Greece, pp.784-787, Oct.2001.