

Compression à bas débit de séquences vidéo couleurs par l'algorithme Mask Motion JPEG2000

David FAURA Patrick GARDA

LISIF (Laboratoire des Instruments et Systèmes d'Ile de France)
Université Pierre et Marie Curie 4, place Jussieu 75005 Paris
(David.Faura, Patrick.Garda)@lis.jussieu.fr

Résumé – L'algorithme MMJPEG2000 améliore la qualité visuelle de MJPEG2000. Dans cet article, nous nous sommes intéressés à 3 méthodes de binarisation qui rendent l'algorithme de détection de mouvement automatique. Nous avons étudié l'influence du sous-échantillonnage des chrominances sur MMJPEG2000 et évalué la complexité de l'algorithme.

Abstract – MMJPEG200 improves the visual quality of MJPEG2000. In this paper, we study 3 thresholding methods to automate the motion detection. We studied the influence of the subsampling of the chrominances on MMJPEG2000 and evaluated the complexity of the algorithm.

1 Introduction

Nous avons proposé un algorithme original de compression de séquences d'image basé sur une détection de mouvement par champs de Markov et une compression des images suivant la norme JPEG 2000 [1].

Dans ce contexte, nous avons mis au point un algorithme de segmentation d'image couleur et réalisé un démonstrateur de codec combinant cet algorithme de segmentation et un codec JPEG2000 [2].

Ces premiers travaux ont montré l'intérêt de notre approche pour la transmission de séquence vidéo sur des canaux à bas débit par des systèmes embarqués.

Néanmoins, trois problèmes se posent pour rendre notre approche exploitable dans ce contexte :

- Automatiser entièrement l'algorithme pour l'embarquer dans des systèmes autonomes.
- Diminuer le débit pour atteindre ceux disponibles sur les réseaux GPRS ou UMTS.
- Diminuer le débit de calculs pour pouvoir le porter sur des systèmes embarqués.

L'étude de ces trois problèmes fait l'objet de cet article.

2 Description de l'algorithme

Suite à de nombreux travaux sur la détection de mouvement à base de champs Markoviens[3], nous avons développé un algorithme original de segmentation du mouvement sur des images couleurs.

Dans un premier temps, la séquence d'images est convertie dans le format de couleur, YCbCr. Les images de différence sont utilisées pour construire trois cartes binaires de mouvement. Ces images binaires sont obtenues grâce à un processus markovien et une relaxation

déterministe, l'ICM. Ces cartes binaires servent ensuite à masquer les images de différences – Figure 1. Cette technique nous permet de débruiter les images et de reconstruire certaines zones du mouvement masquées par l'effet de recouvrement.

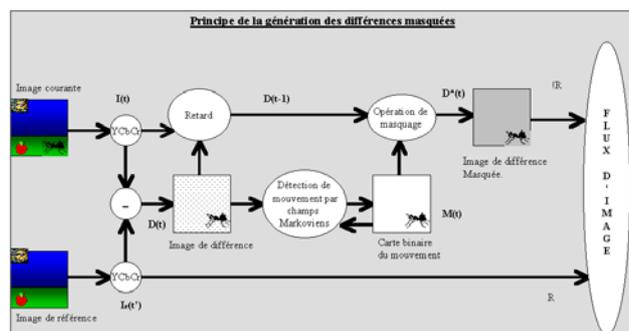


Figure 1- Génération des images de différences masquées

L'intérêt des images de différences est de pouvoir les compresser plus fortement que les images de Référence. Les résultats montrent une amélioration visuelle et statistique significative par rapport à MJPEG2000 (Figure 2).

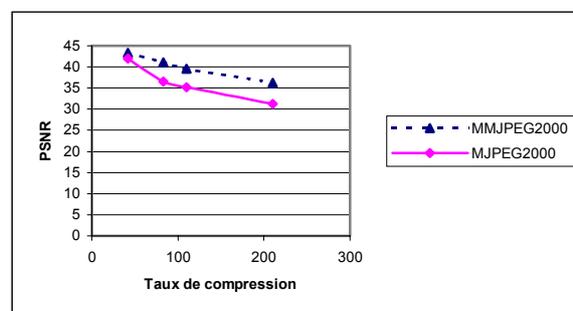


Figure 2- MJPEG2000 vs MMJPEG2000

3 Seuillage automatique

3.1 Le seuillage automatique

La rapidité de la génération des images, nous permet d'introduire des techniques de binarisation automatique pour rendre le système de détection de mouvement autonome. Pour cela, nous recherchons une méthode de binarisation qui adapte le seuil de détection en fonction de la quantité de mouvement dans la séquence

Nous avons étudié trois techniques, une par seuillage entropique, le seuil de binarisation était égal à l'énergie entropique de l'image, une autre par une approximation du seuillage entropique, où l'énergie entropique est approximée par l'écart type[5] de l'image et pour finir le seuillage par la méthode d'Otsu.

3.1.1 Seuillage par l'écart type

Comme dans notre modèle les images de différence sont perturbées par un bruit additif gaussien de variance σ^2 . L'écart type standard σ peut être choisi seuil.

3.1.2 Seuillage Entropique

Shannon a introduit la notion d'énergie entropique $E(X)$ pour une source X d'entropie $H(X)$:

$$E(X) = \frac{1}{2\pi e} e^{2H(X)}$$

$$H(X) = -\sum_{x=0}^{x=n} P(X) \log P(X)$$

$P(X)$ étant la probabilité du symbole X , X étant l'intensité d'un pixel.

Si on considère que les images de différences sont perturbées par un bruit additif gaussien alors la valeur du seuil de binarisation sera égale à l'écart type entropique[7] :

$$K_i = \sqrt{H(X)} = \frac{1}{\sqrt{2\pi e}} e^{H(x)}$$

3.1.3 Seuillage par la méthode d'Otsu

Dans cette méthode, on considère que la puissance du signal utile suit une loi gaussienne, on peut alors utiliser le critère d'Otsu pour séparer les deux classes en optimisant la variance interclasse. soit k le seuil recherché.

La variance interclasse est donnée par :

$$\sigma^2(k) = w(k)(1-w(k))(\mu_0(k) - \mu_1(k))^2$$

$$k = \arg \max \sigma^2(k)$$

avec
$$w(k) = \frac{1}{S} \sum_1^k \text{Hist}(i)$$

$$\mu_0 = \frac{1}{S} \sum_{i=1}^k i \text{Hist}(i) \quad \text{et} \quad \mu_1 = \frac{1}{S} \sum_{i=k}^M i \text{Hist}(i)$$

où Hist est l'histogramme de l'image de différence, S la Taille de l'image.

3.2 Évaluation du seuillage automatique

Pour évaluer leurs performances, nous avons mesuré la valeur des seuils, le PSNR des images reconstruites sans compression, l'entropie des images de différences après masquage et la vitesse de génération des seuils.

3.2.1 Les séquences tests.

La base de test comporte plusieurs types de séquences vidéo au format Qcif.

Table 1- Description des séquences de test

Nom	Description de la séquence
Claire	Présentatrice, pas de changement de plan
ErikAkiyo	Le sujet a un fort mouvement de tête, un changement de plan puis akiyo la présentatrice.
VideoS	Vidéo surveillance avec plusieurs changements de plans, changement important de la luminosité dans certaines scènes.
VideoF	Enfants en train de jouer : très forts mouvements, nombreux changements de plan, balayage de la caméra, zoom.

3.2.2 Evolution des seuils

On constate que la méthode d'Otsu et la méthode du seuillage entropique ont quasiment le même comportement. Mis à part pour les séquences de téléphonie, le seuillage d'Otsu n'est pas adaptatif puisque constant. Par le seuillage entropique, lors de forts mouvements les seuils ont tendance à être trop importants contrairement à Otsu.

Table 2 – Moyenne et écart type des seuils en fonction de la séquence

	Claire		Erikakiyo		VideoF		VideoS	
	m	σ	m	σ	m	σ	m	σ
Entropique	3	0.4	3	9.6	11	5.3	2	4
Ecart type	10	3.6	11	12	28	27	14	12
Otsu	1	0	2	3.	7	5.3	5	5.2

3.2.3 L'entropie

Nous évaluons l'entropie des images de différence car l'entropie $H(x)$ est une mesure du désordre dans une image. La diminution de l'entropie indique une uniformisation de l'image ce qui signifie que le codage de l'image sera simplifié. Mais comme, il s'agit de reconstruire une image à partir d'une image de référence et d'une image de différence, si l'entropie de l'image de différence est trop faible alors la reconstruction sera de mauvaise qualité.

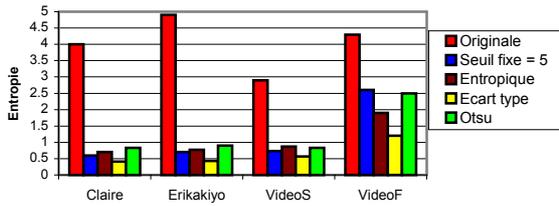


Figure 3- Évaluation de l'entropie en fonction du type de seuillage

3.2.4 Le PSNR

Pour évaluer la qualité de la reconstruction, nous utilisons le PSNR, critère très utilisé en compression, mais ne donnant qu'une mesure statistique des dégradations, calculé de la manière suivante :

$$PSNR = 10 \log \frac{255^2}{MSE(Y)}$$

Y étant la luminance de l'image

La mesure du PSNR, sans compression, sera la borne supérieure des mesures de qualités ultérieures pour les images reconstruites.

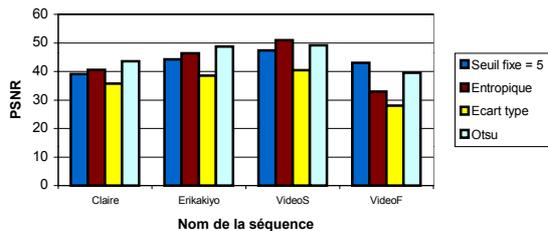


Figure 4- Evaluation du PSNR en fonction du type de seuillage

3.2.5 Mesure de vitesse

Les vitesses des différents modules ont été mesurées sur un Athlon de fréquence 1,5 GHZ avec 256 Mo de RAM, sous Windows 98.

Table 3- vitesse des différentes méthode de binarisation

Vitesse ms	Claire	Erikakiyo	VideoS	VideoF
Seuil fixe = 5	0,006	0,006	0,006	0,006
Entropique	0,6	0,7	0,6	0,6
Ecart type	0,8	0,6	0,5	0,5
Otsu	1,4	1,5	1,5	1,4

3.2.6 Analyse des résultats

En testant ces méthodes, nous avons constaté qu'il fallait instaurer un seuillage minimum, pour les séquences de type vidéophonie, journal, au moins égal à 2, pour que l'ICM puisse éliminer le bruit impulsionnel. Un seuillage trop faible a pour incidence un masque dit "passant" : tous les pixels sont à 1 et l'entropie de l'image de différence n'est pas diminuée mais cela n'a pas de conséquence sur la qualité de l'image reconstruite. Par contre, un seuillage trop fort introduit de fortes erreurs dans la reconstruction, on a alors une superposition de l'image de référence et de l'image courante. Le seuillage par écart type n'est pas

adapté pour notre méthode de compression, par contre, il donne de très bon résultat pour les applications de détection de mouvements dans un environnement très bruité. La méthode d'Otsu génère des seuils constants pour des séquences de type visiophonie mais grâce à la limite inférieure que l'on impose l'entropie des masques est fortement diminuée et le PSNR est meilleur de 3 dB, alors que le seuillage entopique adapte parfaitement ses seuils. En présence de fort mouvement les seuils générés sont plus bas que ceux du seuillage entropique provoquant des dégradations moins importantes. Par contre, la méthode d'Otsu est deux fois plus lente que la méthode du seuillage entropique.

4 Vers les bas débits

Pour diminuer ces débits, nous allons utiliser la méthode classique du sous échantillonnage des chrominances. La diminution de l'information contenue dans les plans de chrominances par la technique de masquage nous laisse envisager que le sous-échantillonnage aura moins d'influence sur la qualité de l'image qu'un sous échantillonnage classique.

Les résultats obtenus pour un sous échantillonnage en 4 :2 :0 montrent que l'on peut obtenir la séquence Claire au format Qcif a un débit de 15 Ko/s (à 12 images secondes) pour un taux de compression de 16 pour l'image de référence et de 32 pour les images de différence, 50 étant le grand maximum (PSNR 33,8, 10 Ko/s). Sans sous échantillonnage, on obtient le même débit, à taux de compression de l'image de référence identique, pour des images de différence comprimé à un taux de compression de 90, Le PSNR obtenu est statistiquement meilleur mais localement les défauts visuels sont plus gênant que lors du sous échantillonnage par contre la restitution des couleurs est meilleure table 4 et figure 5.

Table 4- Comparaison des PSNR pour un débit donné.

Pour le même débit 15 Ko/s	PSNR Y	PSNR Cb	PSNR Cr
Avec Sous/échantillonnage	34.8	38	32
Sans sous/échantillonnage	35	42	38



Figure 5 –Comparaison d'image a 15 Ko/s

Même si les deux méthodes ont des résultats quasi équivalents, le sous échantillonnage permet d'abaisser la complexité au niveau du codeur, quelque soit le format de l'image, puisque l'algorithme de régularisation est appliqué aux 3 plans de chrominances. Le gain en vitesse de génération des images envoyées au codeur est estimé à 20 % pour le format en 4:2:2 et 40 % pour le format en 4:2:0.

5 Diminution de la Complexité de l'algorithme

5.1 Occupation Mémoire.

L'algorithme de segmentation des masques a besoin de 7 emplacements mémoires : deux pour les images d'origine, deux pour travailler en nombres signé, une image pour stoker l'image de différence courante et trois images binaires pour les cartes du mouvement.

On évalue pour une image couleur de taille mxn (ligne*colonne) la mémoire nécessaire pour l'algorithme table 5.

Table 5- Occupation mémoire en fonction de m x n

	PC (octet)	HARD octet
4:4:4	33 mxn	19 m xn
4:2:0	19.5 mxn	11.5 mxn

Table 6- Occupation mémoire pour les formats courants

	PC Mo		HARD Mo	
	4:4:4	4:2:0	4:4:4	4:2:0
Qcif	0,8	0,5	0.5	0.3
Cif	3,4	2.0	2.0	1.2
720 x 576	13,6	8	7,8	4.7

La mémoire étant un facteur très limitant pour l'intégration d'un algorithme sur une architecture embarquée, la table 6 peut justifier l'emploi du sous échantillonnage.

5.2 Nombre d'opérations.

Pour évaluer le volume de calculs nous utilisons le nombre d'opérations par seconde, sans compter les accès mémoires table 7.

Exemple pour une addition du type $X = a + b$ le nombre d'opérations est 1. Les opérations d'entrées sorties ne sont donc pas prises en comptes.

Table 7- Nombre d'opérations pour un plan de chrominance m x n

Transformation de couleurs	6 mxn
Soustraction	1 m x n
Binarisation (fixe)	2 m x n
Calcul statistique	4 m x n
Init énergie de données	2040
ICM	40 m x n
Masque	1 m x n
Total	54 m x n + 2040

Le volume de calcul nécessaire pour embarquer l'algorithme de segmentation est estimé en nombre d'opérations par seconde. Le nombre d'opérations nécessaire, pour traiter une image couleur, est estimé à :

$$\text{Nbr_Operations} = 162 * \text{Taille_Image} + 6120.$$

Les résultats suivants sont donnés pour des images couleurs sous la contrainte de **25 images par secondes** :

Table 7- Nombre d'opérations par seconde (en Millions)

Format	4:4:4	4:2:2	4:2:0
Qcif	105	70	52
Cif	420	280	210
CCIR601	1680	1120	820

Comme précédemment pour l'occupation mémoire, la table 7 montre que le sous échantillonnage permet de diviser d'un facteur 2 le nombre d'opérations par seconde.

6 Conclusion.

Les résultats obtenus sur le module de binarisation automatique montrent que le seuillage adaptatif améliore la qualité de l'image quant les mouvements dans la séquence ne sont pas trop rapide dans ce cas il semble nécessaire de définir une borne supérieur pour le seuillage. La diminution du débit peut être faite soit en augmentant le taux de compression soit en sous-échantillonnant les chrominances. Bien que dégradant les qualités statistiques et la couleur de l'image, visuellement, l'image présente moins de défauts locaux que l'image «sur compressé». De plus, le sous-échantillonnage permet de réduire de façon significative le volume de calcul et l'occupation mémoire de l'algorithme de segmentation.

D'après, l'évaluation de l'algorithme, nous pouvons réaliser une intégration de l'algorithme sur un système embarqué pour des images Qcif, voire Cif et satisfaire les contraintes des réseaux GPRS et UMTS.

Références

- [1] F.Lohier, P.Garda, L.Lacassagne, Procédé et dispositif de traitement de séquences d'images avec masquage *Brevet UPMC, Français en cours d'extension internationale, Demande FR 62060 L du 3 février 2000*
- [2] D.Faura, P.Garda "Segmentation d'images couleurs pour la compression de séquences vidéo par l'algorithme Mask Motion JPEG2000", CORESA2003.
- [3] A.Caplier, Thèse "algorithme Markovien de détection de mouvement dans les séquences d'images", 1995.
- [5] M.Lievin, Thèse, "Analyse entropico-logarithmique de séquences vidéo couleur", 2000.
- [6] David S.Taubman, JPEG2000, image compression, fundamentals, standards and practice.
- [7] M.Liévin, F.Luthon, E. Keeve Entropic Estimation of noise for Medical Volume Restoratoin, ICPR 2002.