



Schéma de compression vidéo

G. Laroche 04/01/2006

Plan



- Introduction
- Étapes de compression
 - Codage du résiduel
 - Calcul du résiduel
 - Autres étapes
- Conclusion
- Questions

Plan



- Introduction
- Étapes de compression
 - Codage du résiduel
 - Calcul du résiduel
 - Autres étapes
- Conclusion
- Questions

Introduction – Qu'est ce qu'un codec vidéo?



- Ensemble d'outils.
- Entrée : séquences vidéo.
- Sortie : bitstream.
- Seul le décodeur est standardisé :
 - Doit pouvoir décoder tout bitstream standard.
 - Et doit rendre une qualité équivalente aux autres décodeurs standardisés.

Introduction – A quoi sert un codec vidéo?

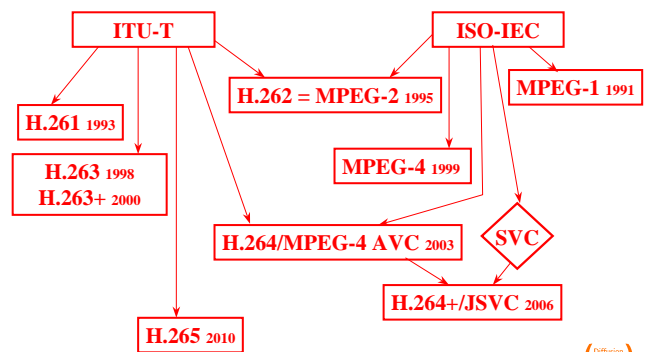


- Compresser un signal dont on connaît les caractéristiques, les redondances spatiales et temporelles (≠ de ZIP).
- Applications vidéo de plus en plus nombreuses (TV, Cinéma, DVD, Visio).
- Permet une augmentation des qualités et résolutions.

Introduction – Instituts et organisation



- Deux instituts de normalisation: ITU – ISO

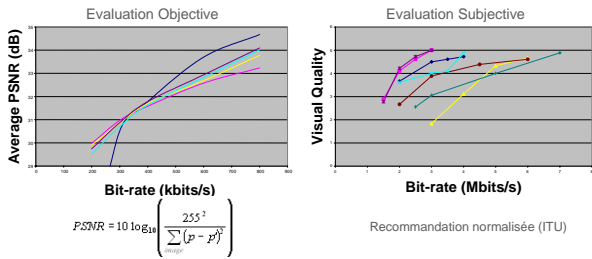


Introduction – Évaluation



→ Évaluation d'une technique de compression

- Codage avec perte: courbe débit-distorsion (R-D)



⇒ Evaluation Objective moins fiable/Subjective très coûteuse

Profil



→ Profil :

- Ensemble d'outils liés à une application.

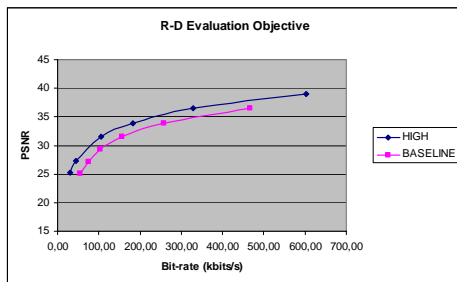
→ Exemple : H.264 7 profils différents :

- Baseline : Application mobile, visiophonie.
- Main : Streaming/broadcast.
- Extended : Mobile video services.

- High :
- High 10
- High 4:2:2 :
- High 4:4:4 :

Fidelity Range Extension,
TVHD, Digital Cinema,
Blue Ray/HDDVD.

Profil



Plan



→ Introduction

→ Étapes de compression

- Codage du résiduel
- Calcul du résiduel
- Autres étapes

→ Conclusion

→ Questions

Schéma Global

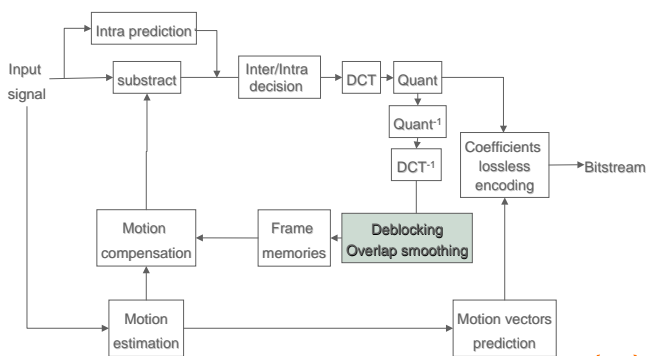
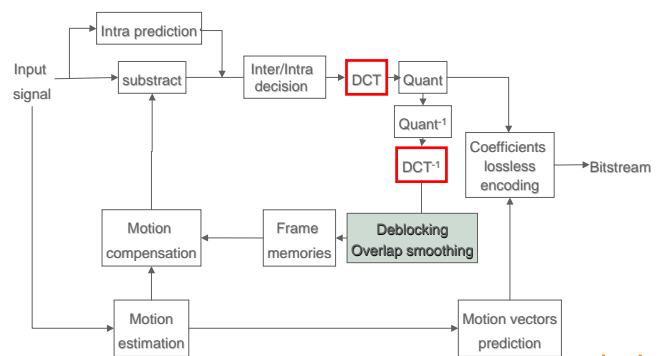


Schéma Global



Transformations - Généralités



- Entropie : valeur moyenne de la quantité d'information associée à une variable aléatoire.
- Objectif : dé-corréler le signal.
- Pourquoi ? Codage entropique + efficace.

Domaine pixel

Équiprobabilité des valeurs sur [0 ; 255]



Énergie concentrée sur les basses fréquences, proba faible sur les grandes amplitudes dans les hautes fréquences

Domaine transformé

Entropie en chute libre !

Transformations - Généralités



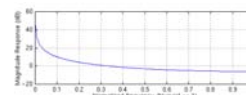
- Pourquoi ça marche ?

Corrélation spatiale dans l'image

« un pixel ressemble souvent à son (ses) voisin(s) »



+ d'énergie dans les basses fréquences



Transformations - DCT



- Transformée fréquentielle
- Implémentation simple
- Algorithmes rapides
- Efficace

⇒ Utilisée dans JPEG, MPEG-1, MPEG-2...

Transformations - DCT



- Formule :

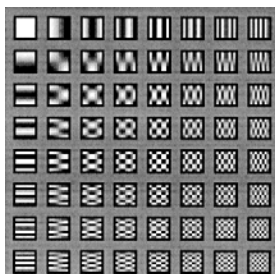
$$F(u,v) = \frac{2}{N} c(u)c(v) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} img(x,y) \cos\left[\frac{\pi}{N}u\left(x + \frac{1}{2}\right)\right] \cos\left[\frac{\pi}{N}v\left(y + \frac{1}{2}\right)\right]$$

- Dans la pratique : DCT(I) = P⁻¹IP

Transformations - DCT



- I est une combinaison linéaire des éléments de la base suivante :



Exemple DCT



- I :

| | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|
| 139 | 144 | 149 | 153 | 155 | 155 | 155 | 155 |
| 144 | 151 | 153 | 156 | 159 | 156 | 156 | 156 |
| 150 | 155 | 160 | 163 | 158 | 156 | 156 | 156 |
| 159 | 161 | 162 | 160 | 160 | 159 | 159 | 159 |
| 159 | 160 | 161 | 162 | 162 | 155 | 155 | 155 |
| 161 | 161 | 161 | 161 | 160 | 157 | 157 | 157 |
| 162 | 162 | 161 | 163 | 162 | 157 | 157 | 157 |
| 162 | 162 | 161 | 161 | 163 | 158 | 158 | 158 |

- DCT(i) :

| | | | | | | | |
|------|-----|-----|----|----|----|----|----|
| 1260 | -1 | -12 | -5 | 2 | -2 | -3 | 1 |
| -23 | -17 | -6 | -3 | -3 | 0 | 0 | -1 |
| -11 | -9 | -2 | 2 | 0 | -1 | -1 | 0 |
| -7 | -2 | 0 | 1 | 1 | 0 | 0 | 0 |
| -1 | -1 | 1 | 2 | 0 | -1 | 1 | 1 |
| 2 | 0 | 2 | 0 | -1 | 1 | 1 | -1 |
| -1 | 0 | 0 | -1 | 0 | 2 | 1 | -1 |
| -3 | 2 | -4 | -2 | 2 | 1 | -1 | 0 |

Transformation DCT



→ En pratique :

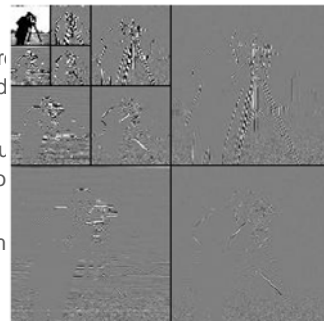
$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & -1 & -2 \\ 1 & -1 & -1 & 1 \\ 1 & -2 & 2 & -1 \end{bmatrix} \begin{bmatrix} x_{00} & x_{01} & x_{02} & x_{03} \\ x_{10} & x_{11} & x_{12} & x_{13} \\ x_{20} & x_{21} & x_{22} & x_{23} \\ x_{30} & x_{31} & x_{32} & x_{33} \end{bmatrix} \begin{bmatrix} 1 & 2 & 1 & 1 \\ 1 & 1 & -1 & -2 \\ 1 & -1 & -1 & 2 \\ 1 & -2 & 1 & -1 \end{bmatrix}$$

⇒ DCT H.264 décalages et additions

Transformation DWT



- Discr stand
- Pas u
- Perm

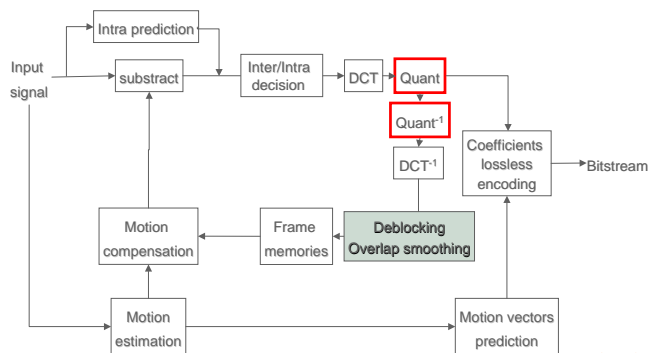


Transformation DWT



- Discrete Wavelet Transform : utilisé dans le standard JPEG2000.
 - Pas utilisé (pour l'instant) dans les codeurs vidéo STANDARD.
 - Permet un scalabilité en résolution.
- ⇒ Outil pour la scalabilité

Schéma Global



Introduction de pertes - quantification



- Matrice de quantification en fonction du "pas de quantification".
 - Division entière des coefficients de la matrice DCT par les coefficients de la matrice de quantification.
 - Objectif : réduire l'alphabet de la matrice DCT.
- ⇒ Entraîne une perte de données.

Exemple quantification



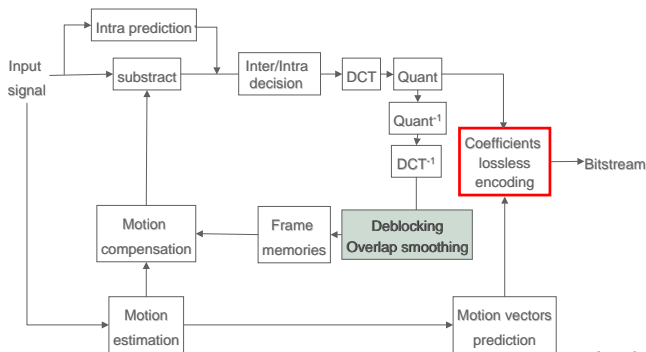
→ DCT(i) :

$$\begin{bmatrix} 1260 & -1 & -12 & -5 & 2 & -2 & -3 & 1 \\ -23 & -17 & -6 & -3 & -3 & 0 & 0 & -1 \\ -11 & -9 & -2 & 2 & 0 & -1 & -1 & 0 \\ -7 & -2 & 0 & 1 & 1 & 0 & 0 & 0 \\ -1 & -1 & 1 & 2 & 0 & -1 & 1 & 1 \\ 2 & 0 & 2 & 0 & -1 & 1 & 1 & -1 \\ -1 & 0 & 0 & -1 & 0 & 2 & 1 & -1 \\ -3 & 2 & -4 & -2 & 2 & 1 & -1 & 0 \end{bmatrix}$$

→ Q(DCT(i)) :

$$\begin{bmatrix} 79 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ -2 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Schéma Global



Codage entropique



- Suite à la quantification on a beaucoup de valeurs proches de 0.
- Le codage entropique va permettre de réduire la quantité d'information.
- 2 types de codage:
 - Variable Length Coding.
 - Codage Arithmétique.

Codage entropique – VLC (Variable Length Coding)



- Longueur d'une séquence donnée
- $$L = Nb_symboles * \sum_{i=0}^{N-1} p_i \cdot l_i$$

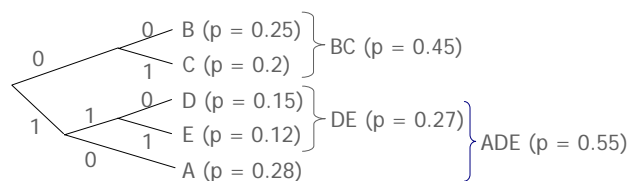
Nb de valeurs prises par les symboles

Nb de bits attribués à la valeur i

Probabilité de la valeur i
- Entropie de cette séquence (bits/symbole):
- $$h = -\sum_i p_i \cdot \log_2(p_i)$$
- Codage de Huffman (1952) = optimal au sens de l'entropie(bits/symbole):

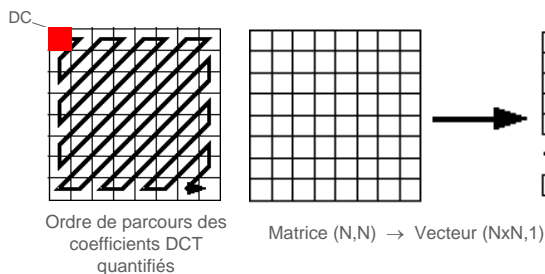
⇒ On attribue $-\lceil \log_2(p_i) \rceil$ bits au symbole i

Codage entropique - Huffman (1952)



A = 10 B = 00 C = 01 D = 110 E = 111

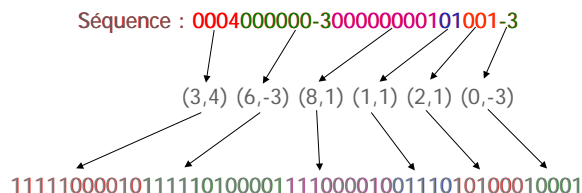
Le RLC ... Zig-zag scanning



- Balayage par ordre croissant des fréquences spatiales.

⇒ RLC efficace

Codage entropique - RLC+VLC



- Chaque couple codé par l'algorithme de Huffman.
- ⇒ plus la fréquence d'apparition d'un symbole est forte, plus la chaîne est courte.

Codage Arithmétique - IBM, début 80 's



- Limite du codage de Huffman : nb de bits entier par symbole.
- Ex: $p_A = 90\% \Rightarrow A$ codé sur 1 bit au lieu de 0.15 bit (6 fois trop).
- Codage arithmétique : on code une suite de symboles par un intervalle en virgule flottante sur $[0,1[$.

Exemple

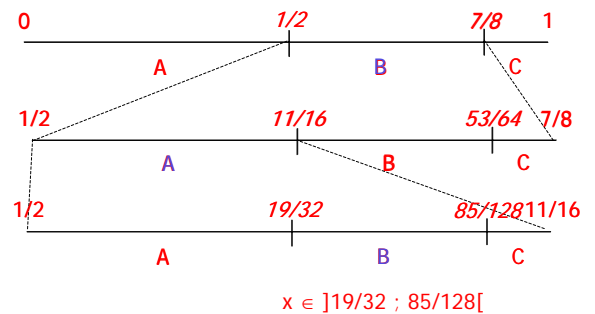
- A ($p = 0.500$) • On veut transmettre BAB
 Symboles : B ($p = 0.375$) • Probabilités connues du codeur
 C ($p = 0.125$) et du décodeur

Recherche & Développement



Septembre 2005

Codage Arithmétique - IBM, début 80 's



Recherche & Développement



Septembre 2005

Codage Arithmétique - IBM, début 80 's



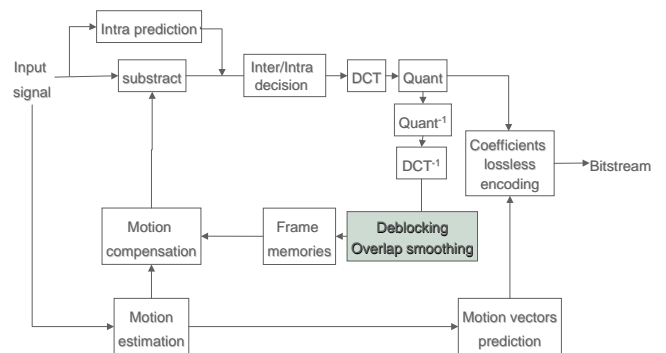
- On code $19/32 \Rightarrow 5$ bits.
- Au décodage on connaît la table de probabilité.
- On a :
 - $1/2 > 19/32 > 7/8 \Rightarrow$ premier caractère est B
 - $(19/32 - 1/2) / (\text{intervalle de B}) = 3/32 / (7/8 - 1/2) = 0.25$
 - $0 > 0.25 > 1/2 \Rightarrow$ 2ième caractère est A
 - Etc...
 - La chaîne est donc BAB

Recherche & Développement



Septembre 2005

Schéma Global

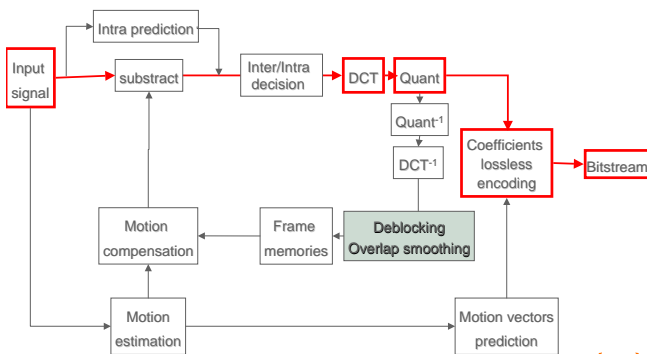


Recherche & Développement



Septembre 2005

Schéma Global JPEG



Recherche & Développement



Septembre 2005

Plan



- Introduction
- Étapes de compression
 - Codage du résiduel
 - Calcul du résiduel
 - Autres étapes
- Conclusion
- Questions

Recherche & Développement



Septembre 2005

Résiduel ?



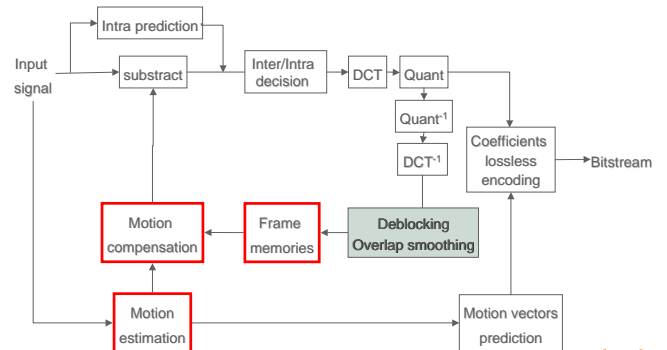
- Recherche d'un prédicteur à partir de l'information déjà connue.
- Puis codage de la différence entre le prédicteur et le bloc de départ.

$$e(x, y) = p(x, y) - \hat{p}(x, y)$$

- Et inversion au décodage.

$$p(x, y) = \hat{p}(x, y) + e(x, y)$$

Schéma Global



Redondances temporelles



- Compression d'image fixe : redondance spatiale.
- Compression d'image vidéo : redondance spatiale et temporelle.
- Entre 2 images successives, peu de différences.
⇒ il doit y avoir moyen de déduire l'une de l'autre...

Redondances temporelles

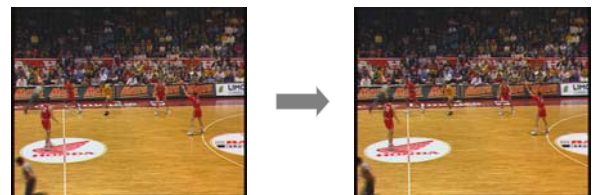


Image 1 (t_0)

Image 2 ($t_0 + 40 \text{ ms}$)

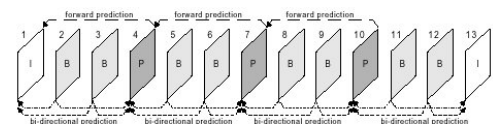
Image Type



- I : (INTRA) Codage uniquement spatial. Équivalent à JPEG.
- P : (PREDICTED) Codage temporel, en fonction d'une image précédente.
- B : (BIDIRECTIONAL) Codage temporel, en fonction de 2 images.

$$\Rightarrow \text{Débit(I)} > \text{Débit(P)} > \text{Débit(B)}$$

Temporel - Schémas prédictifs



- On prédit une image par rapport à une (plusieurs) autre(s) de référence ; on code la \neq .
- Ex: images P et B.
- Suppose une boucle de décodage dans le codeur.
- Et qu'une image soit codée en Intra (I).

Estimation de mouvement – nécessaire ?



- Zone immobile :
 - Très faible erreur de prédiction
 - Skipping
- Zone en mouvement :
 - Mouvement trop rapide pour l'œil
 - Forte quantification

Le block matching



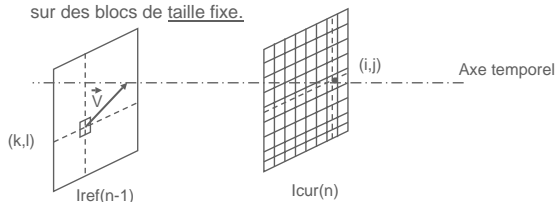
- Principe :
 - Mise en correspondance par minimisation de la MSE sur des blocs de taille fixe.
- Avantages :
 - Pas de segmentation à transmettre.
 - Simplicité.
 - Bonne compatibilité avec le codage spatial (block based).

Block matching - full search



→ Principe :

- Mise en correspondance par minimisation d'une fonction coût sur des blocs de taille fixe.

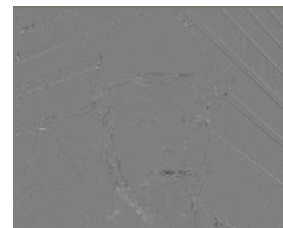


- Pour chaque bloc $B[i][j]$ centré en (i,j) , recherche dans l'image de référence du bloc $B[k][l]$ lui ressemblant le plus.
- Le vecteur est $V(B[i][j]) = (i-k, j-l)$.

Compensation de mouvement - Intérêt



t_0-t_1



Compensation de mouvement entre t_0 & t_1

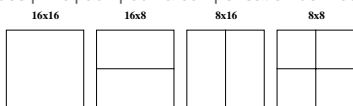
Motion Block Shapes



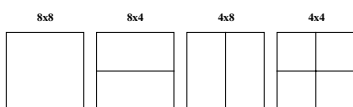
→ Anciens codeurs : bloc 8x8

→ H.264 :

- 4 modes principaux pour la compensation de mouvement.



- En mode 8x8, un bloc peut être divisé.



⇒ mise en compétition

Sub-Pixel Motion Prediction



→ Pourquoi ?:

- Mouvement non pixelique.
- Engendre plus de détails.

→ Estimation au pixel entier, au 1/2 et au 1/4 de pixel

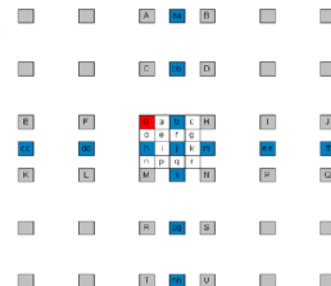
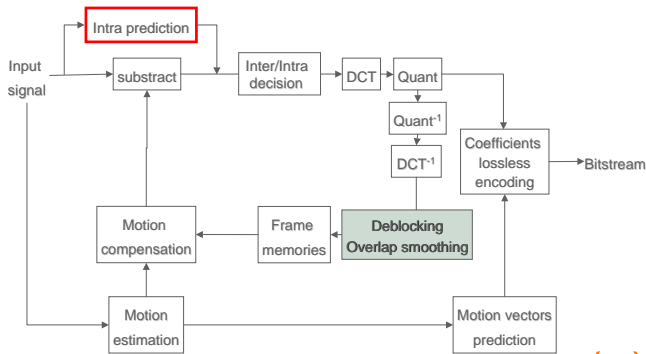


Schéma Global



Recherche & Développement

Diffusion Libre
Septembre 2005

Prédiction Intra



BUT : Exploiter la redondance spatiale.

- H.263 : Prédiction domaine DCT
 - Bloc 8x8 : 3 modes
 - H.264 : Prédiction domaine pixel
 - Bloc 4x4 : 9 modes
 - Bloc 16x16 : 4 modes
 - Bloc 8x8 (high) : 9 modes
- ⇒ H.264 plus efficace que H.263 et JPEG2000

Recherche & Développement

Diffusion Libre
Septembre 2005

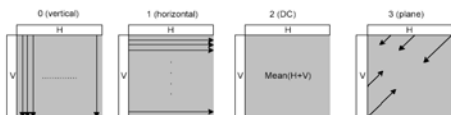
Prediction Intra 16x16



→ Le macrobloc est prédit en fonction de la ligne du haut et la colonne de gauche.

→ 4 modes de prédiction :

- DC prediction
- Horizontal prediction
- Vertical prediction
- Plane prediction



⇒ mise en compétition

Recherche & Développement

Diffusion Libre
Septembre 2005

Plan

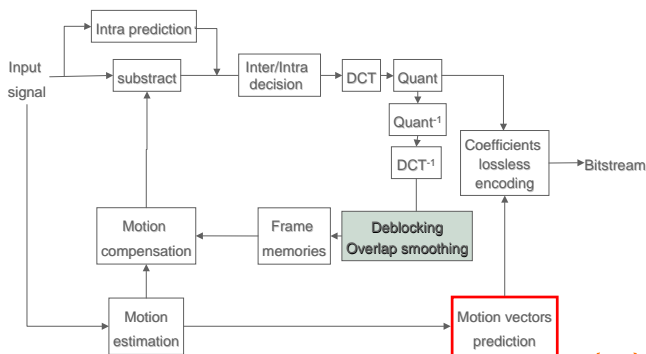


- Introduction
- Étapes de compression
 - Codage du résiduel
 - Calcul du résiduel
 - Autres étapes
- Conclusion
- Questions

Recherche & Développement

Diffusion Libre
Septembre 2005

Schéma Global



Recherche & Développement

Diffusion Libre
Septembre 2005

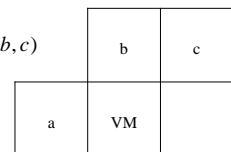
Temporel - Codage des vecteurs !



→ En général:

- Les vecteurs sont codés sans perte
- MV_x et MV_y séparément
- MV_x et MV_y sont codés à l'aide d'un prédicteur qui est la médiane des vecteurs voisins déjà enregistrés.

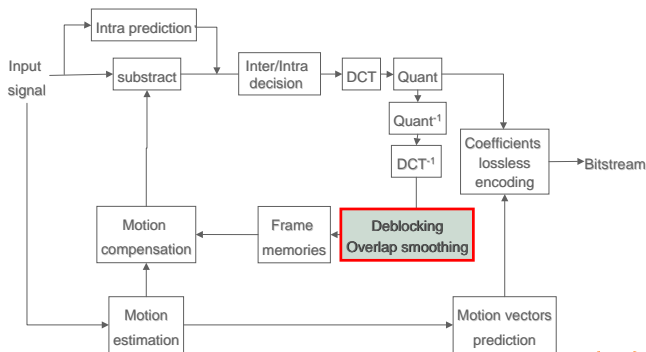
$$\text{Résiduel}_{VM} = VM - \text{Médian}(a, b, c)$$



Recherche & Développement

Diffusion Libre
Septembre 2005

Schéma Global



Recherche & Développement

Diffusion Libre
Septembre 2005

Deblocking filter

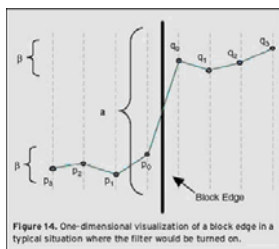


- Pourquoi?
 - Amélioration de la qualité visuelle.
 - Gommage des effets de bloc.
 - Compensation de mouvement plus efficace.
- Comment?
 - Dans la boucle de compensation de mouvement.
 - "Moyenne pondérée" des pixels en bordure de bloc.
 - Organisé par filtre.
 - Choix des filtres en fonction des caractéristiques des blocs.

Recherche & Développement

Diffusion Libre
Septembre 2005

Deblocking filter



1. $|p_0 - q_0| < \alpha(QP)$
2. $|p_1 - p_0| < \beta(QP)$
3. $|q_1 - q_0| < \beta(QP)$

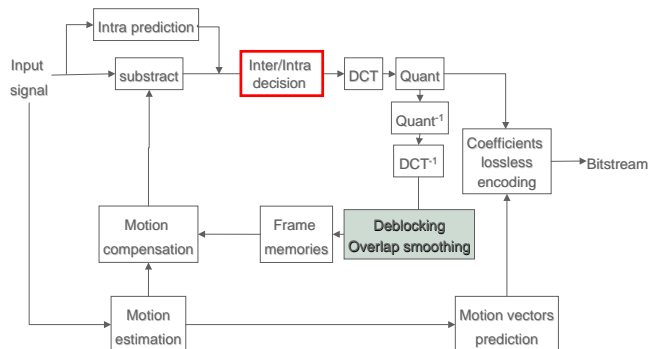
Figure 14. One-dimensional visualization of a block edge in a typical situation where the filter would be turned on.

⇒ Beaucoup de tests, clips, très coûteux, très efficace.

Recherche & Développement

Diffusion Libre
Septembre 2005

Schéma Global



Recherche & Développement

Diffusion Libre
Septembre 2005

Inter/Intra decision



- Mise en compétition de tous les modes possibles.
- Minimisation d'une fonction coût respectant :
 - La qualité
 - Le bit-rate

$$J = D + \lambda R$$

Distortion (qualité)

Rate (bit-rate)

Non normatif

Recherche & Développement

Diffusion Libre
Septembre 2005

Conclusion



- Compression toujours nécessaire.
- Exploitation des redondances spatiales et temporelles.
- Schéma de codage vidéo calqué sur celui du codage d'image.
- L'enjeu des prochaines années : HD, consommation mobile, la vidéo 3D.

Recherche & Développement

Diffusion Libre
Septembre 2005

La bibliographie



- Jean-Alain Hernandez, René Joly, **Les normes et les standards du multimédia**, Dunod 2000.
- **Compression d'images fixes et animées JPEG et MPEG**, conférence ENST Bretagne 2003.
- T. Wiegand, G.J. Sullivan, **Overview of the H.264/AVC Video Coding Standard**, IEEE 2003.

Chiffres et Résultats



- Un film de 1h30 :
 - 25 Go non compressé.
 - 650 Mo en MPEG4.
 - 350 Mo en H.264.
 - 175 Mo en H.265 soit 4 films sur un CD de 700 Mo