



La compression



Stratégie

- Deux stratégies de réduction de débit
 - **Sans perte** : l'information compressée est intègre \Rightarrow données informatiques
 - **Avec perte** : après compression, on ne peut reconstituer intégralement l'information originale \Rightarrow suffisant pour l'audio et la vidéo. Meilleure réduction de volume en général.
- La compression avec pertes **fragilise** l'information transmise.
 - \Rightarrow La redondance peut servir à retrouver l'information malgré les erreurs. Ex : répétition du même signe 10 fois \Rightarrow 9 pertes de signes tolérées.
- Après compression, les échantillons successifs ne sont plus liés, pas d'interpolation possible \Rightarrow Besoin un **code correcteur d'erreurs**.

Cependant, le gain obtenu par compression est toujours supérieur à la redondance introduite par le code correcteur d'erreurs.



Rapport de compression (ou taux de compression)

- Noté $\eta:1$, ce qui signifie réduction d'un facteur η .

$$\eta = \frac{\text{débit sans compression}}{\text{débit avec compression}} = \frac{\text{volume sans compression}}{\text{volume avec compression}}$$



Application à l'audio et la vidéo

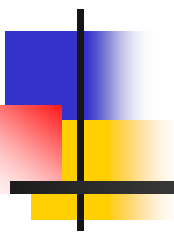
- Beaucoup de redondance dans les phénomènes naturels (sons et images non synthétisés).
- Vidéo = son + image
- Dans une vidéo, où se trouve la redondance ?
 - Dans le son, redondance temporelle.
 - Dans l'image, redondance
 - spatiale (deux pixels voisins sont généralement assez semblables)
 - temporelle (peu de différence entre l'image N et l'image N+1)
 - fréquentielle : l'énergie d'une image est concentrée essentiellement dans les basses fréquences, tandis que les hautes fréquences participent peu au sens.



Comment compresser?

- **Corrélation spatiale** : Dans une image, il y a beaucoup de plages uniformes \Rightarrow exploité par la **Transformée en Cosinus Discrète**.
- **Corrélation temporelle** : D'une image à la suivante, il y a peu de changements ; on transmet uniquement la différence entre images \Rightarrow **Compensation de mouvement**.
- **Redondance subjective** : On tient compte de l'acuité visuelle. Les détails fins n'ont pas besoin d'être codés avec autant de précision que les parties essentielles \Rightarrow **Quantification**.
- **Redondance statistique** : Certains mots reviennent plus fréquemment que d'autres \Rightarrow **Codage entropique**.

Principes de la compression des images basée sur la DCT – Compression JPEG (rappels)



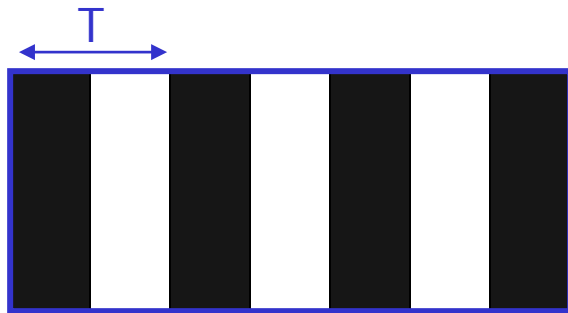


La compression JPEG

- JPEG (*Joint Photographic Expert Group*) est une norme ISO de compression d'images fixes avec pertes, standardisée en 1992.
- JPEG prend en compte les failles du systèmes de perception humaine, de manière à réduire la quantité d'information de l'image qui sera codée.
- L'algorithme de compression repose sur
 - Une transformation mathématique de l'image : la DCT (Discrete Cosine Transform)
 - Un codage à longueur variable : le codage de Huffman
- Formats source
 - R, V, B en 4:4:4
 - Y, R-Y et B-Y en 4:2:2

La notion de fréquences dans une image

- Une image peut présenter une fréquence spatiale horizontale et une fréquence spatiale verticale.

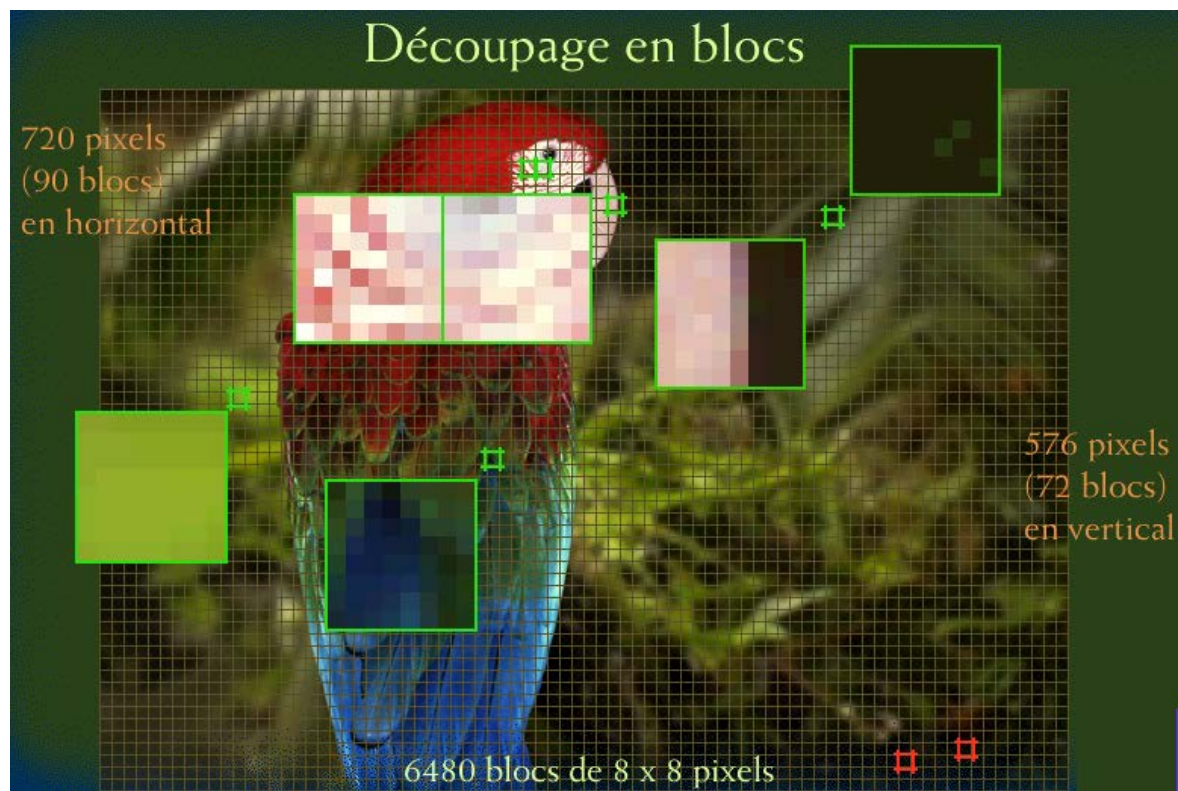


Image

- Si l'œil parcourt de gauche à droite cette image, sensation de variation cyclique de la luminance. Il existe une **période spatiale horizontale**. La fréquence spatiale représente le nombre de cycles par unité de longueur.
 - La période spatiale verticale par contre est nulle.
- Pour les signaux temporels, on utilise la transformée de Fourier pour passer dans le domaine des fréquences.
 - Pour les images, on peut, entre autres, passer dans le domaine des fréquences par la **Transformée en Cosinus Discrète** (TCD, ou DCT – *Discrete Cosin Transform*)

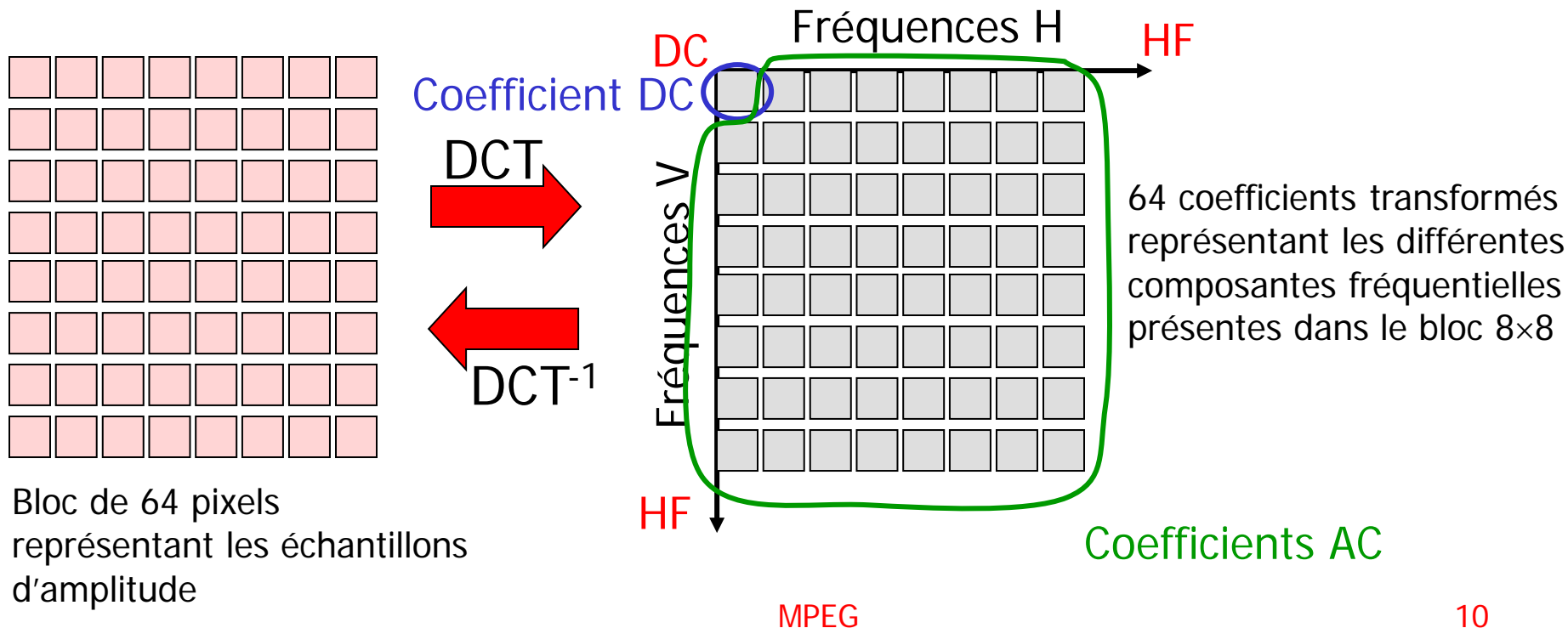
La Transformée en Cosinus Discrète dans la norme JPEG

- La DCT porte sur des blocs de **8×8 pixels**.



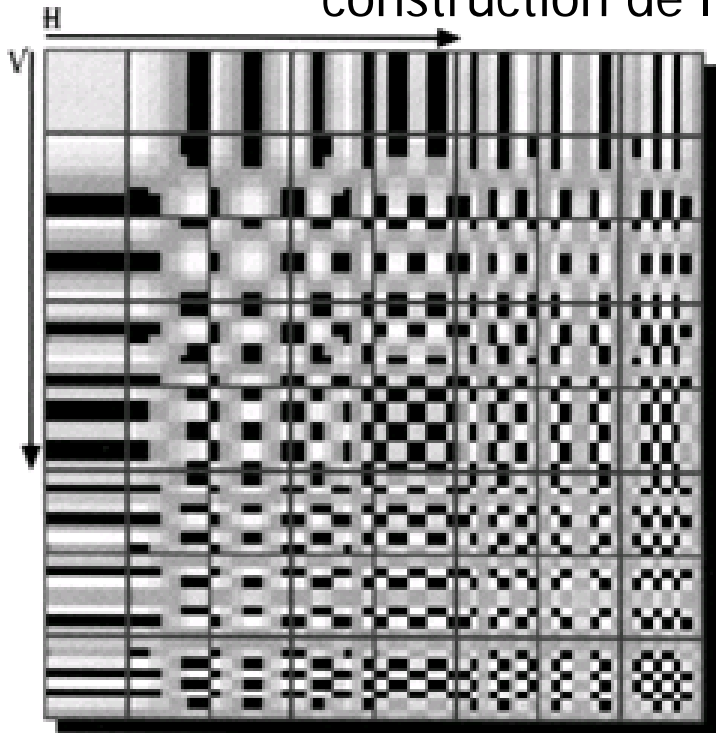
La Transformée en Cosinus Discrète dans la norme JPEG

- La DCT fournit une **représentation spectrale bidimensionnelle** (H et V) du bloc.
- Le bloc est décomposé en une combinaison linéaire de « fonctions » images de base, de la même manière que la transformée de Fourier décompose un signal temporel en une somme de fonctions sinus et cosinus.



La Transformée en Cosinus Discrète dans la norme JPEG

- Les « fonctions » de base de la DCT sont représentées dans le tableau ci-dessous (ici pour la composante de luminance si le format d'entrée est YCrCb).
- Chaque coefficient de la DCT représente la contribution de l'image située à la même place dans la matrice dans la construction de l'image.



Fonctions de base

Exemple : Un extrait de la matrice des coefficients DCT d'un bloc de pixel est le suivant (0 dans toutes les autres cases)

2	1			
3				

Cela signifie qu'il faut « superposer » 2 images $I_{0,0}$, 1 image $I_{1,0}$ et 3 images $I_{1,0}$ pour créer le bloc.

Calcul des coefficients DCT (à titre indicatif)

■ DCT

N = 8 dans la norme JPEG

Valeur de la composante du pixel situé ligne x et colonne y

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

Calcul du coefficient situé ligne u et colonne y dans la matrice DCT

■ DCT⁻¹

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

Où la fonction C() est définie par :

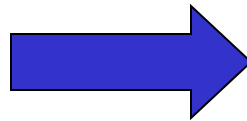
$$C(u) = \begin{cases} \frac{1}{\sqrt{2}} & \text{si } u = 0 \\ 1 & \text{si } u \neq 0 \end{cases}$$

Calcul des coefficients DCT : un exemple simple

- Pour l'exemple, bloc de taille 4×4.
- Calcul sur la luminance Y

Coefficients Y de
luminance du bloc

	0	1	2	3	y
0	10	5	0	0	
1	0	2	1	0	
2	4	0	0	0	
3	0	0	0	0	
x					



Coefficients DCT
correspondants

	0	1	2	3	v
0					
1		?			
2					
3					
u					

$$DCT(u, v) = \frac{2}{N} \cdot C(u) \cdot C(v) \cdot \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} \left[Y(x, y) \cdot \cos \frac{(2x+1)u\pi}{2N} \cdot \cos \frac{(2y+1)v\pi}{2N} \right]$$

Calcul des coefficients DCT : un exemple simple

■ Calcul du premier coefficient DCT

Coefficients de
luminance du bloc

	0	1	2	3
0	10	5	0	0
1	0	2	1	0
2	4	0	0	0
3	0	0	0	0

Coefficients DCT
correspondants

	0	1	2	3
0	5,5			
1				
2				
3				

$$DCT(0,0) = \frac{2}{4} \cdot c(0) \cdot c(0) \cdot \sum_{x=0}^{3-1} \sum_{y=0}^{3-1} Y(x,y) \cdot \underbrace{\cos \frac{(2x+1) \times 0 \times \pi}{2 \times 4}}_{=1} \cdot \underbrace{\cos \frac{(2y+1) \times 0 \times \pi}{2 \times 4}}_{=1}$$

$$= \frac{2}{4} \times \frac{1}{\sqrt{2}} \times \frac{1}{\sqrt{2}} \cdot \sum_{x=0}^{3-1} [Y(x,0) + Y(x,1) + Y(x,2) + Y(x,3)]$$

$$= \frac{2}{4} \times \frac{1}{\sqrt{2}} \times \frac{1}{\sqrt{2}} \cdot \left[\begin{aligned} & (Y(0,0) + Y(0,1) + Y(0,2) + Y(0,3)) \\ & + (Y(1,0) + Y(1,1) + Y(1,2) + Y(1,3)) \\ & + (Y(2,0) + Y(2,1) + Y(2,2) + Y(2,3)) \\ & + (Y(3,0) + Y(3,1) + Y(3,2) + Y(3,3)) \end{aligned} \right]$$

$$= \frac{1}{4} \cdot [(10 + 0 + 4 + 0) + (5 + 2 + 0 + 0) + (0 + 1 + 0 + 0) + (0 + 0 + 0 + 0)]$$

$$= 5,5$$

Calcul des coefficients DCT : un exemple simple

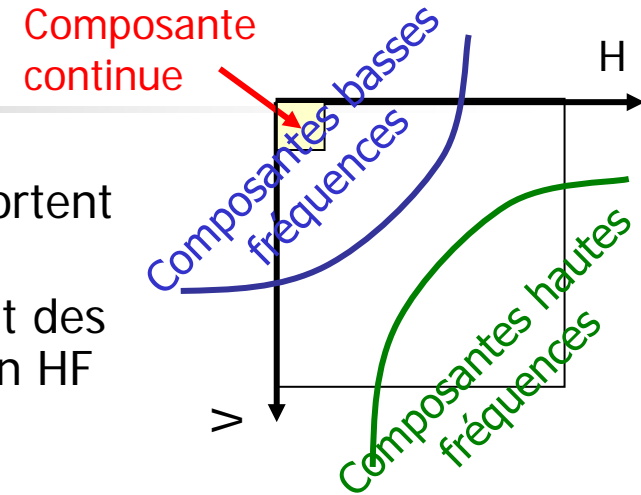
- En répétant la formule pour tous les coefficients DCT, on obtient :

	0	1	2	3	v
0	5,5	5,38	1,5	-0,07	
1	4,76	4,52	0,69	-0,84	
2	2	2,50	1	-0,49	
3	2,35	3,66	2,96	0,98	
u					

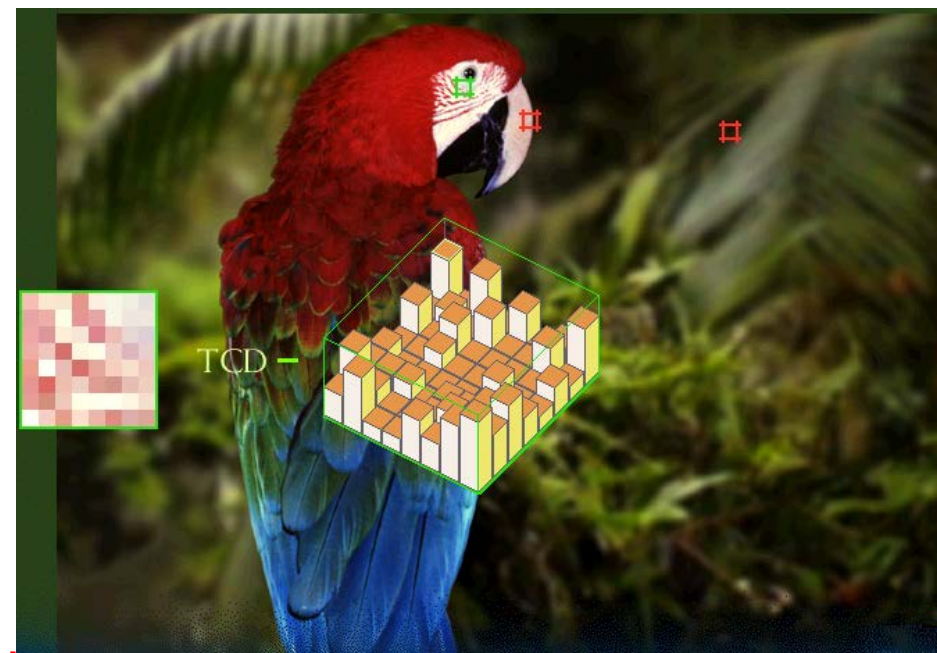
- En pratique, le calcul est toujours fait par ordinateur ou par un circuit électronique dédié !!!

Propriétés de la DCT

- On constate que les coefficients **hautes fréquences** portent les informations **de détails** de l'image.
- Ainsi, les blocs des zones homogènes d'une image ont des coefficients non négligeables en BF, et négligeables en HF (exemple 1)
- Seuls les blocs des zones très hétérogènes ont des coefficients importants en HF (exemple 2)



Exemple 1

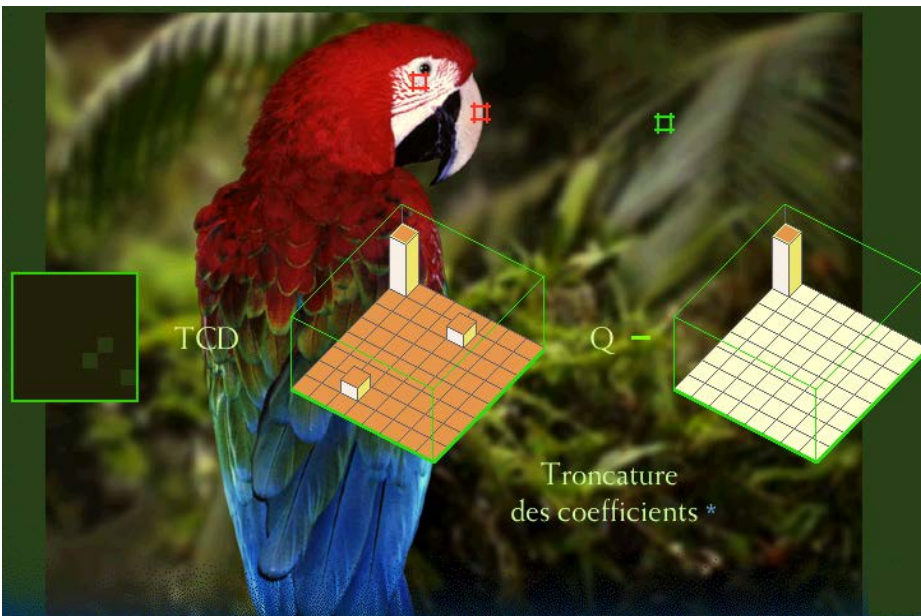


MPEG

Exemple 2

Application à la compression

- Puisque les hautes fréquences portent l'information de détail de l'image, on va diminuer leur valeur, voire l'annuler. C'est l'opération de **quantification**.
- L'image va être dégradée, mais compte-tenu de l'imperfection de l'œil humain, cette dégradation sera peu ou pas perceptible (tout dépendra de la sévérité de la quantification).



Exemple 1

Images MPEG



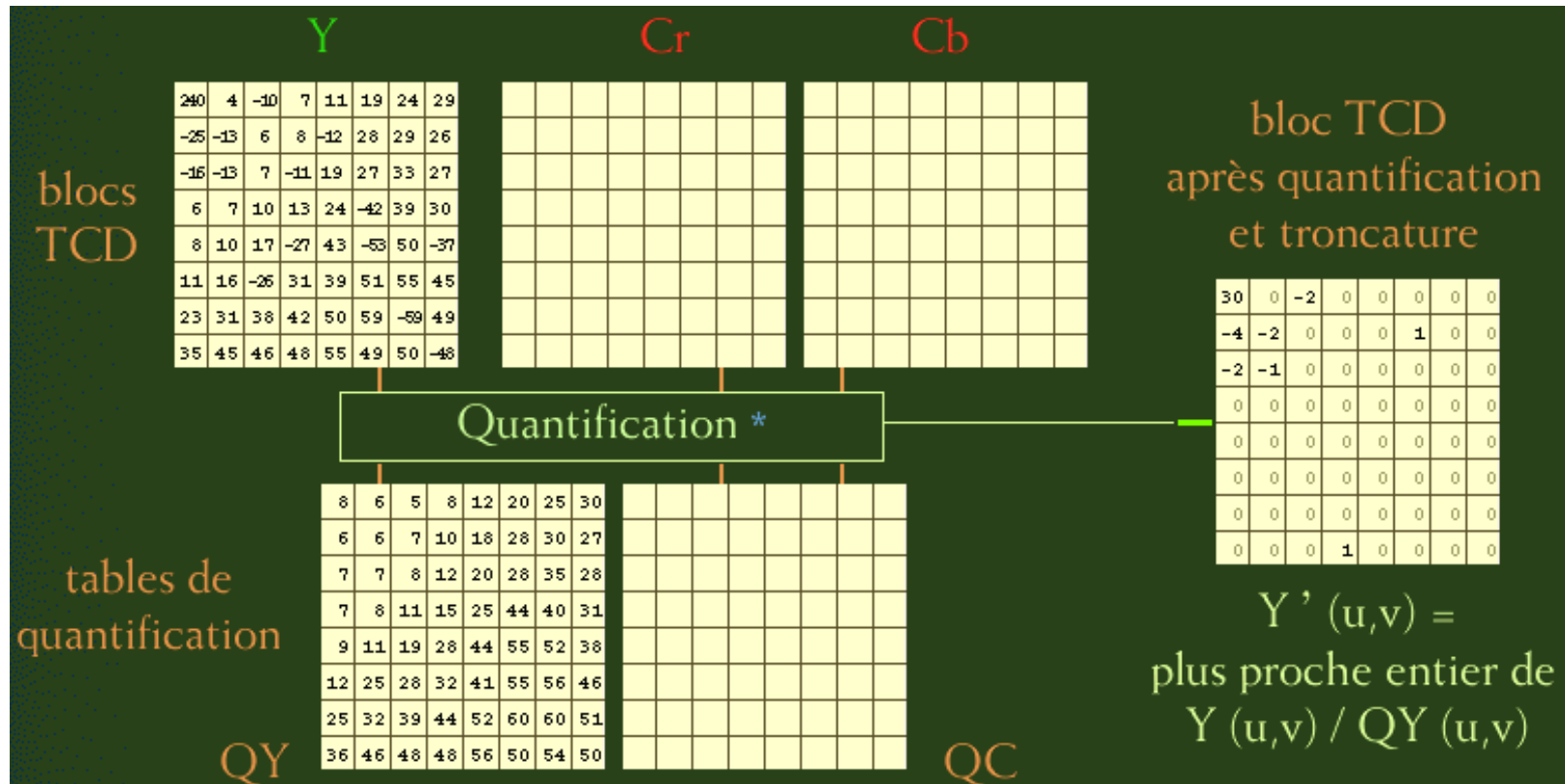
Exemple 2

La quantification

$$Q(x, y) = 1 + (1 + x + y)F_q$$

- La norme JPEG fournit des tables, qui permettent de retoucher la valeur des coefficients DCT de manière à diminuer la valeur de ceux qui contribuent le moins au sens de l'image.
- Les coefficients de haute fréquence sont les plus réduits. Le rapport de compression dépend directement de la table utilisée.

Exemple :
Table Q1



Quantification

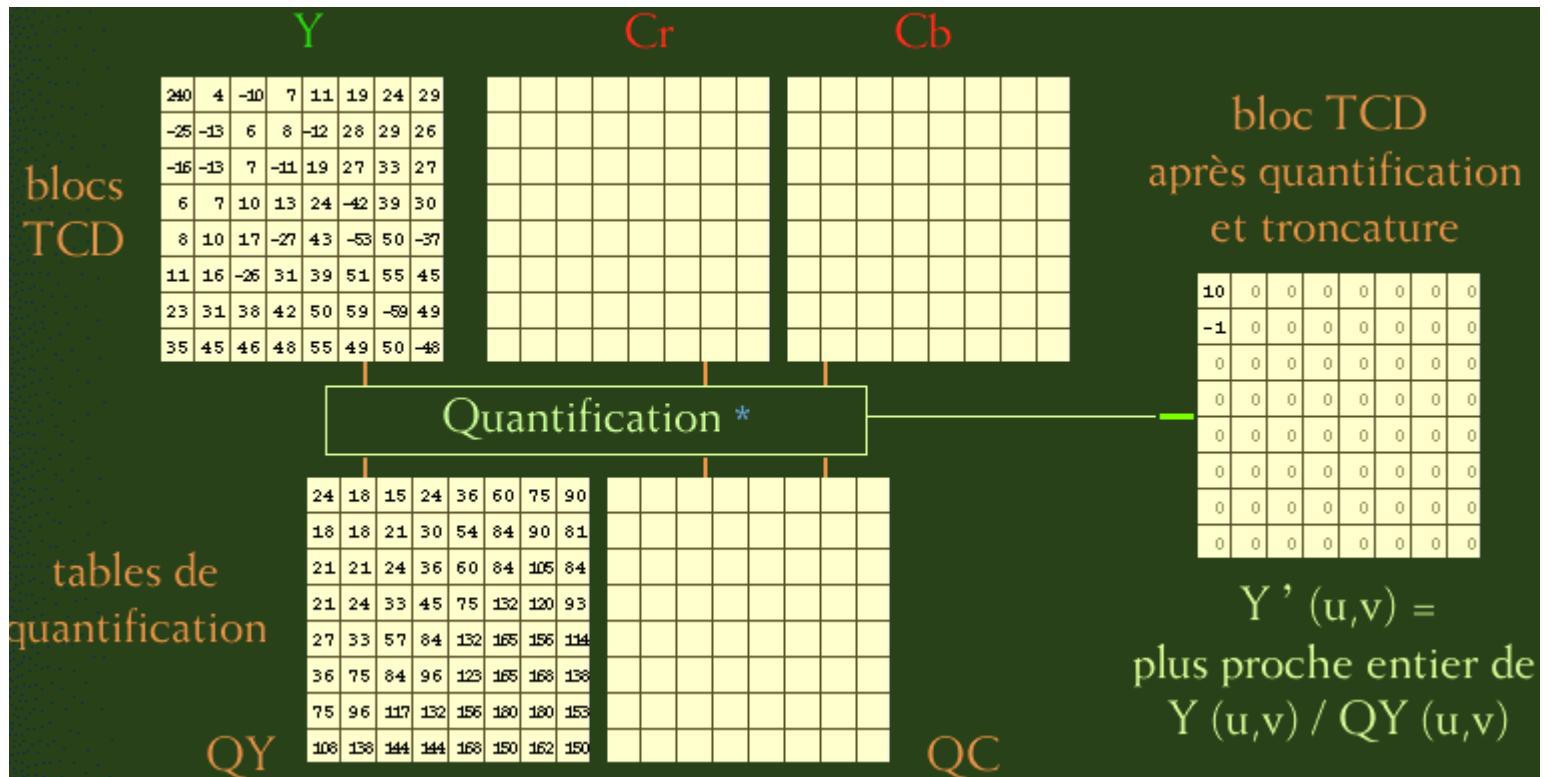


Image INA Formation

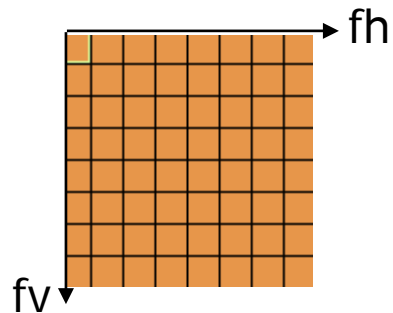
- Des tables de quantification sont définies pour chaque composante (Y, Cr et Cb).
- La quantification est moins sévère pour la luminance car l'œil y est plus sensible.

Illustration de l'effet de la quantification

Images INA Formation



Conservation
des 64
coefficients =
quantification
sans perte



Conservation
de 32
coefficients

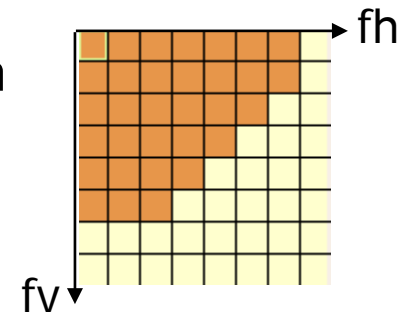
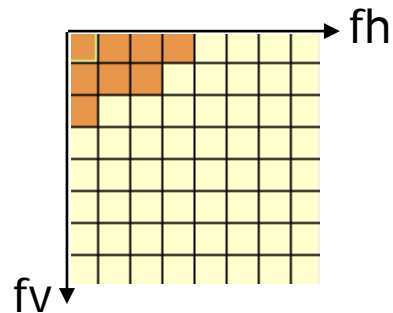


Illustration de l'effet de la quantification

Images INA Formation



Conservation
de 8
coefficients



Conservation
de 4
coefficients

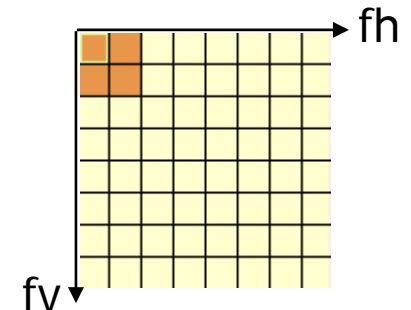
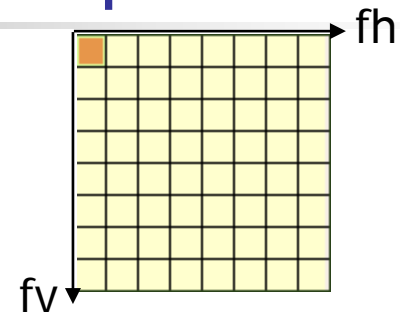


Illustration de l'effet de la quantification

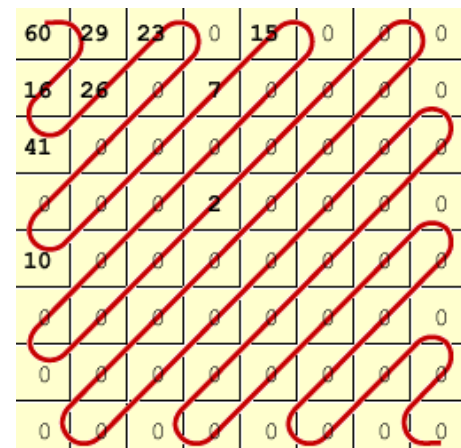
*Image INA
Formation*



Conservation
d'un unique
coefficient = la
composante
continue (DC)

- A retenir : Un codeur peut travailler
 - À taux de réduction constant (au détriment de la qualité)
 - A qualité constante : la performance du codeur en matière de taux de réduction risque d'être médiocre.
- Ne pas oublier que la nature de l'image intervient sur l'efficacité de la compression ! Une image simple peut être de bonne qualité avec un fort taux de réduction. Pour ce même taux, une image complexe sera certainement de piètre qualité...

Le codage



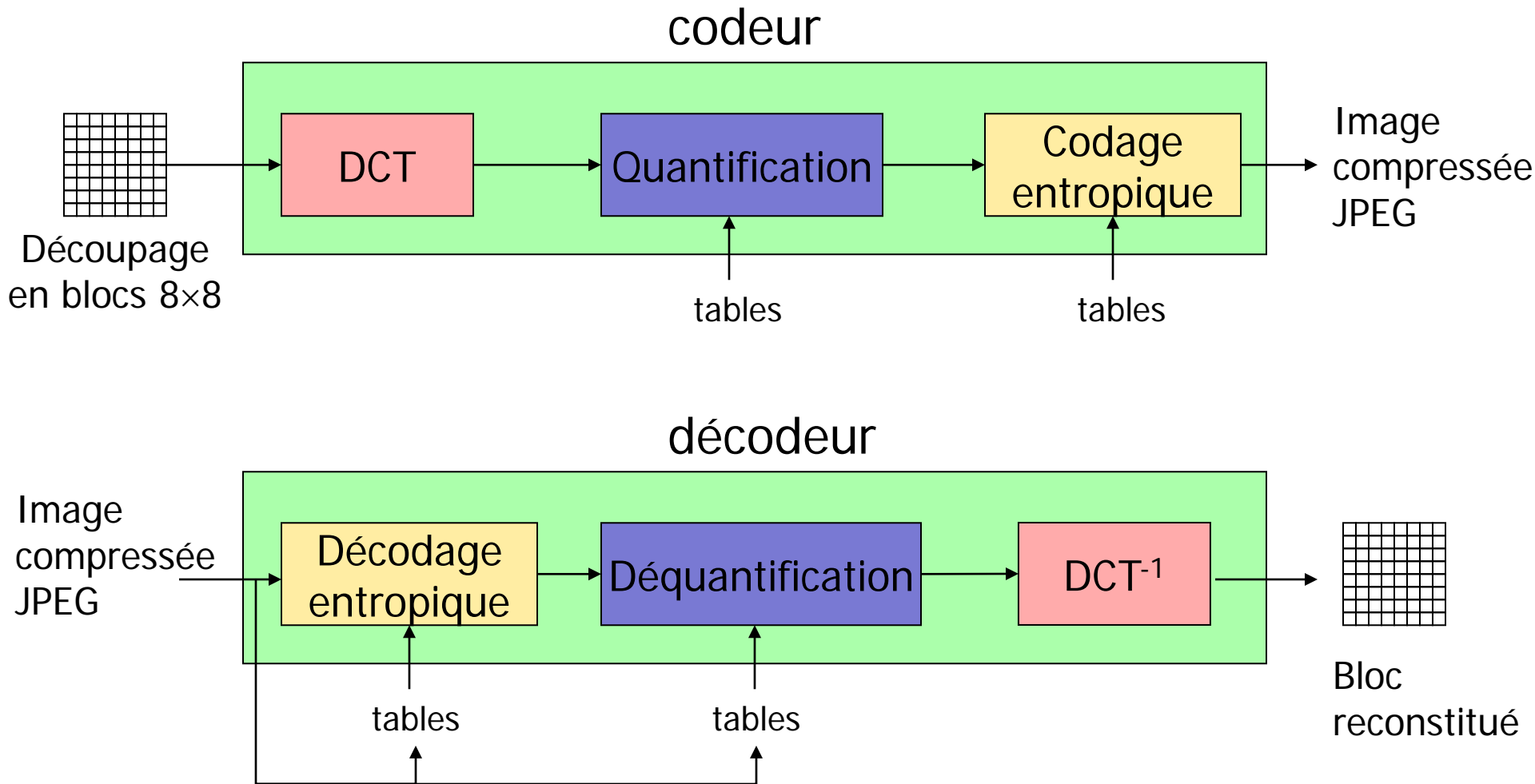
- Les coefficients de la DCT sont balayés en zig-zag, des basses fréquences vers les hautes fréquences.
- Les suites de 0 sont transmises sous forme condensée : couple (nb de « 0 » précédant le coeff non nul, coeff).
- Dans l'exemple : 0 60 0 29 0 16 0 41 0 26 0 23 4 10 2 7 0 15...
- Puis codage à longueur variable : les valeurs les plus courantes sont codées par un mot court, les valeurs les plus rares sont codées par un code long. (Principe du code Morse, ex. « e » = . et « q » = --.-).
- C'est le code de Huffman qui est employé



Pour résumer : algorithme de compression JPEG

1. Décomposition en blocs de 8×8 pixels par composante.
2. Passage du domaine spatial au domaine fréquentiel par DCT.
3. Quantification et seuillage des coefficients DCT.
4. Balayage en zig-zag.
5. Codage entropique

Synoptiques fonctionnels d'un codeur et d'un décodeur JPEG





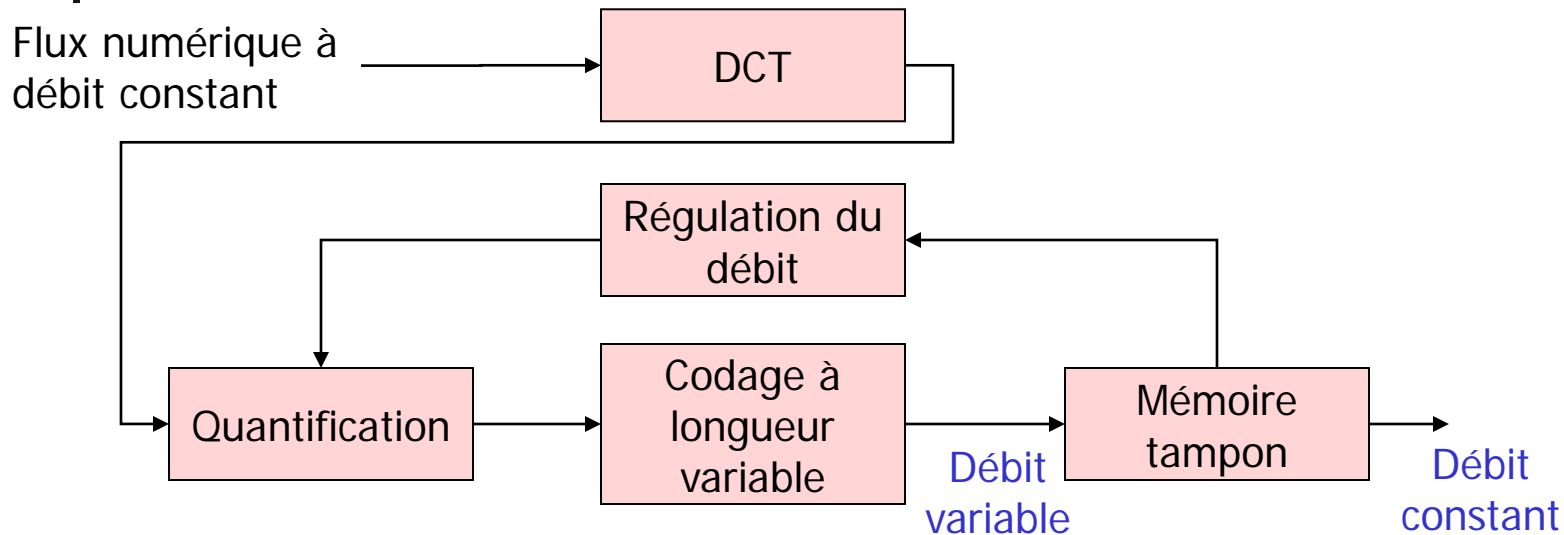
La compression MJPEG



MJPEG

- *Moving* JPEG (MJPEG) **n'est pas une norme !!** C'est un algorithme propriétaire, donc inadapté d'un appareil à l'autre.
- MJPEG permet de compresser les images successives d'une vidéo comme si celles-ci étaient chacune une photo.
- Origine : Le montage d'images compressées dépendantes les unes des autres est difficile car il y a rupture de la séquence. Les monteurs ont appliqué la compression JPEG individuellement aux photogrammes constituant une vidéo
⇒ « MJPEG ».

Fonctionnalités d'un codeur MJPEG



- Afin de garder un débit constant, le signal est stocké dans une mémoire tampon.
- Lorsque la mémoire déborde (par exemple, car images complexes donc DCT large) on sous-quantifie les prochaines images en changeant de table de quantification, voire on supprime certaines images.
- Les tables utilisées sont celles normalisées dans MPEG.



La compression vidéo MPEG 1-2

- Etapes de la compression
- Spécificités des normes MPEG1 et MPEG2

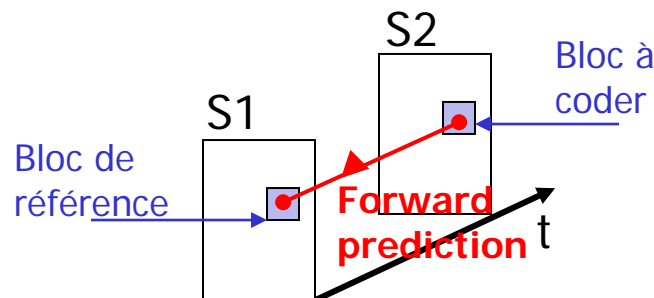


Principe de la compression MPEG

- La compression repose sur la transformation DCT de l'image, comme dans JPEG.
- L'algorithme part du principe que deux images successives dans une vidéo se ressemblent \Rightarrow on peut alors transmettre uniquement les informations qui ont changé entre les deux images.
- Ainsi, les images d'une vidéo peuvent être codées de deux manières :
 - En intra : c'est-à-dire, comme une photographie indépendante des autres images de la vidéo, selon un algorithme identique à celui de JPEG.
 - En prédiction : on ne transmet que les différences (compressées) entre l'image courante et l'image intra la précédant (ou la suivant).

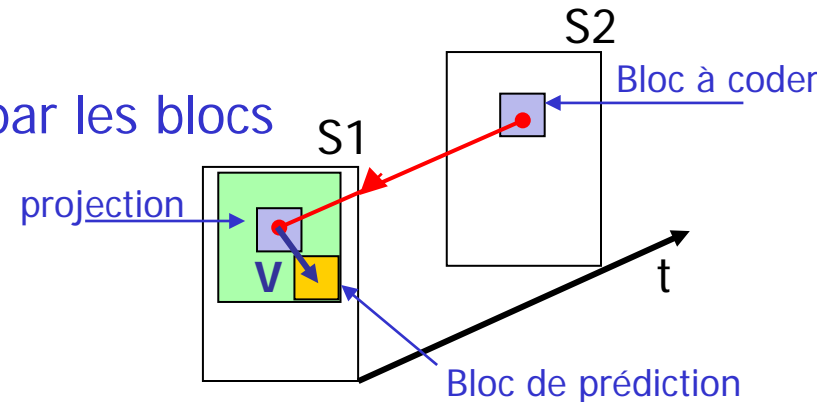
La compression MPEG

- Mesure de la différence entre deux images successives
- Image S1 codée en intra.
- Calcul pour chaque bloc de 8×8 pixels l'erreur $\varepsilon = S2 - S1$. ε est une matrice de 8×8 valeurs.
 - Si les deux images se ressemblent, les coefficients de la matrice sont proches.
- Calcul de la DCT de la matrice ε .
 - Si les deux images se ressemblent, les coefficients basse fréquence de la matrice ε sont plus importants que les coefficients haute fréquence.
 - Donc on peut les quantifier efficacement (i.e. sans perdre trop d'information).
- Au lieu de transmettre l'image S2, on **transmet la matrice des coefficients DCT de ε quantifiés**, qui est bien moins volumineuse.



La compression MPEG

- Estimation du mouvement subi par les blocs
Compensation de mouvement



- Si beaucoup de mouvement,
 - S2 est très différente de S1
 - Les coefficients de la matrice ε sont très différentes
 - Les coefficients DCT haute fréquence de ε sont non négligeables : on ne peut pas les quantifier sans dégradation importante.
- \Rightarrow Le codeur va **estimer le déplacement du bloc dans l'image**.
- Le codeur ouvre une fenêtre de recherche dans S1 autour du bloc de référence.
- Il balaie la fenêtre et calcule ε pour chaque bloc rencontré dans la fenêtre l'erreur (opération de block matching).
- Le bloc de S1 qui minimise l'erreur est retenu.
- Le codeur transmet
 - Les **coefficients DCT quantifiés** de la matrice d'erreur
 - Le **vecteur de mouvement du bloc**.

Illustration de la Forward Prediction

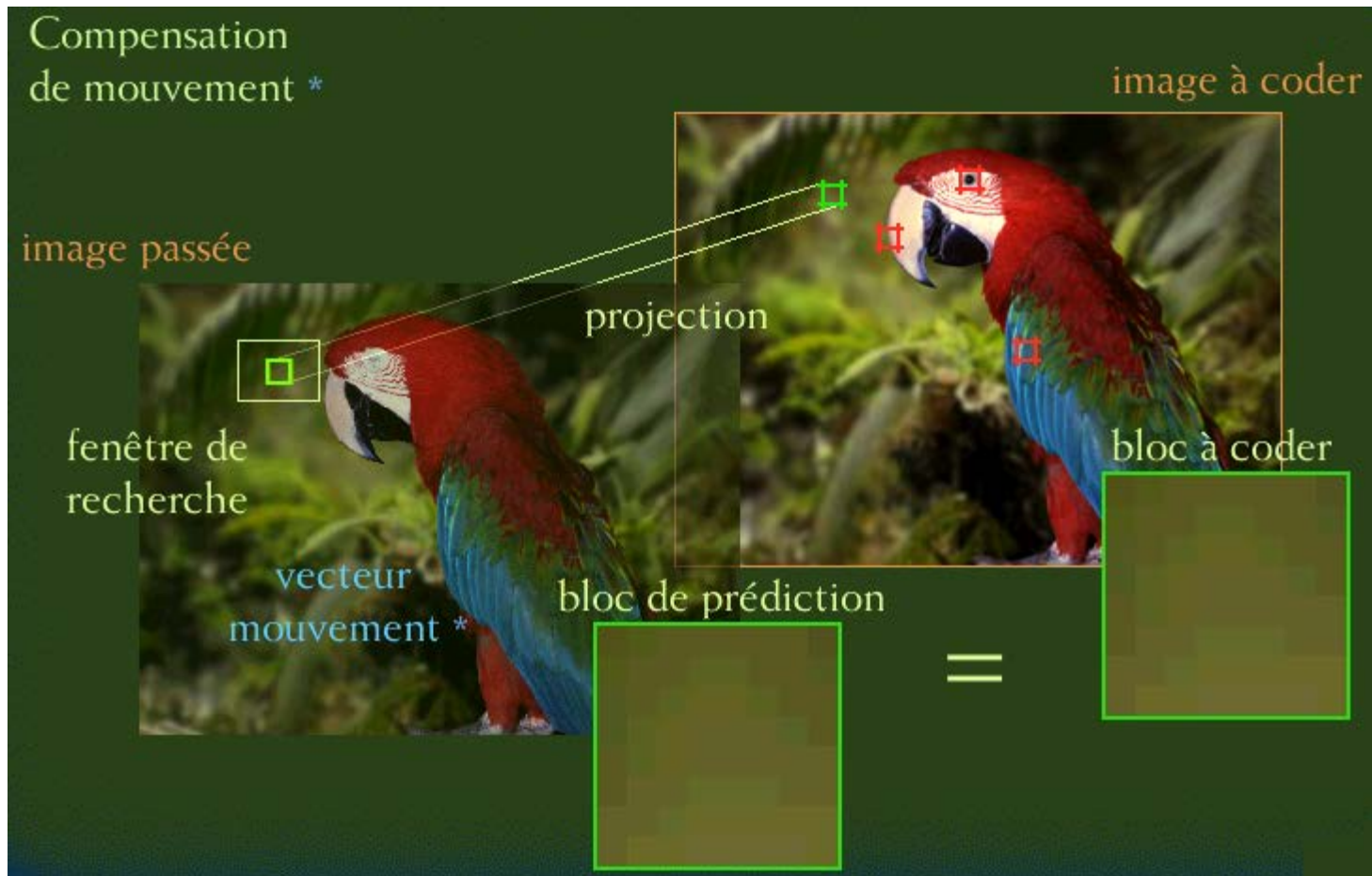
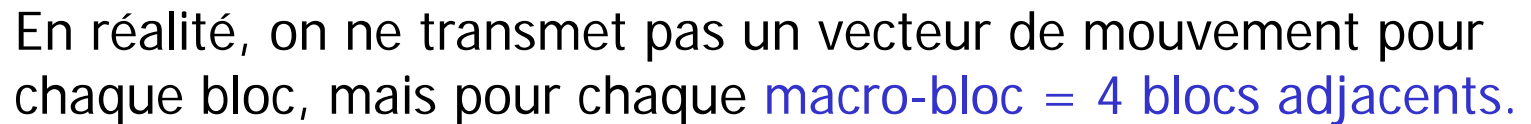
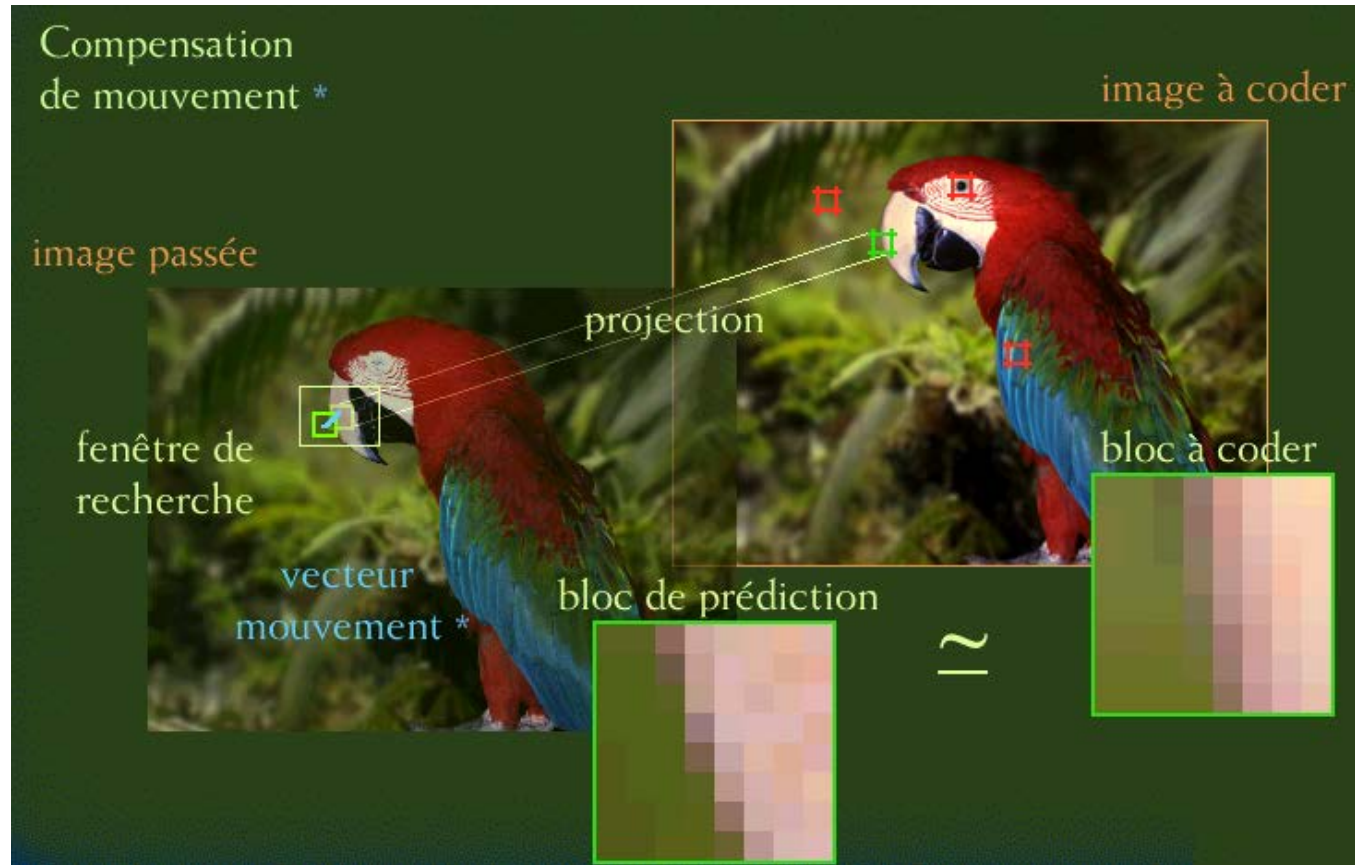


Image INA Formation



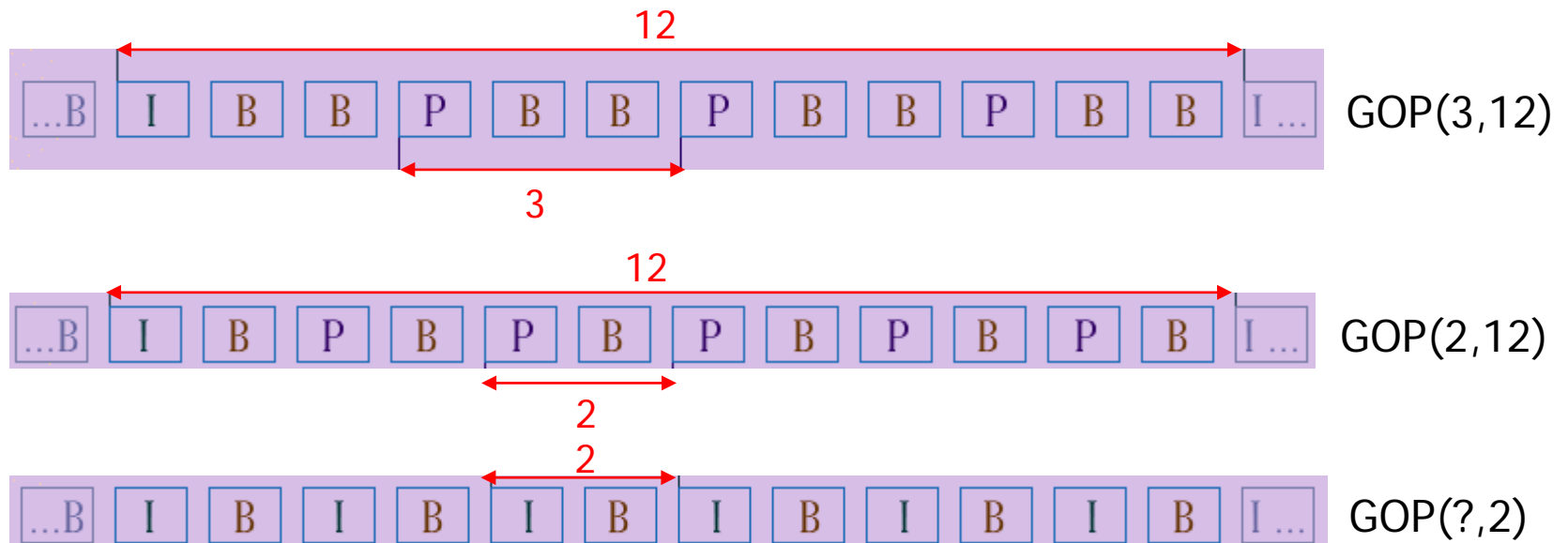


Les types d'images d'une séquence codée en MPEG

- Dans une même séquence vidéo, trois types de codage des images :
 - Image I : en Intra
 - Image P (prédite, *forward prediction*) : par rapport à une image passée
 - Image B (bidirectionnelle, *forward prediction* et *backward prediction*) : par rapport à une image passée et une image future. Le codeur transmet la moyenne des erreurs et transmet les 2 vecteurs.
- Le codeur **choisit** le mode de prédiction le plus rentable.
- A l'intérieur d'une image P : macroblochs I ou P.
A l'intérieur d'une image B : macroblochs I, P ou B.
- **En moyenne**, pour un même contenu,
 - $\text{volume}(P) = \text{volume}(I) / 2$
 - $\text{volume}(B) = \text{volume}(I) / 4$
 - Mais dépend de l'algorithme utilisé par le constructeur.

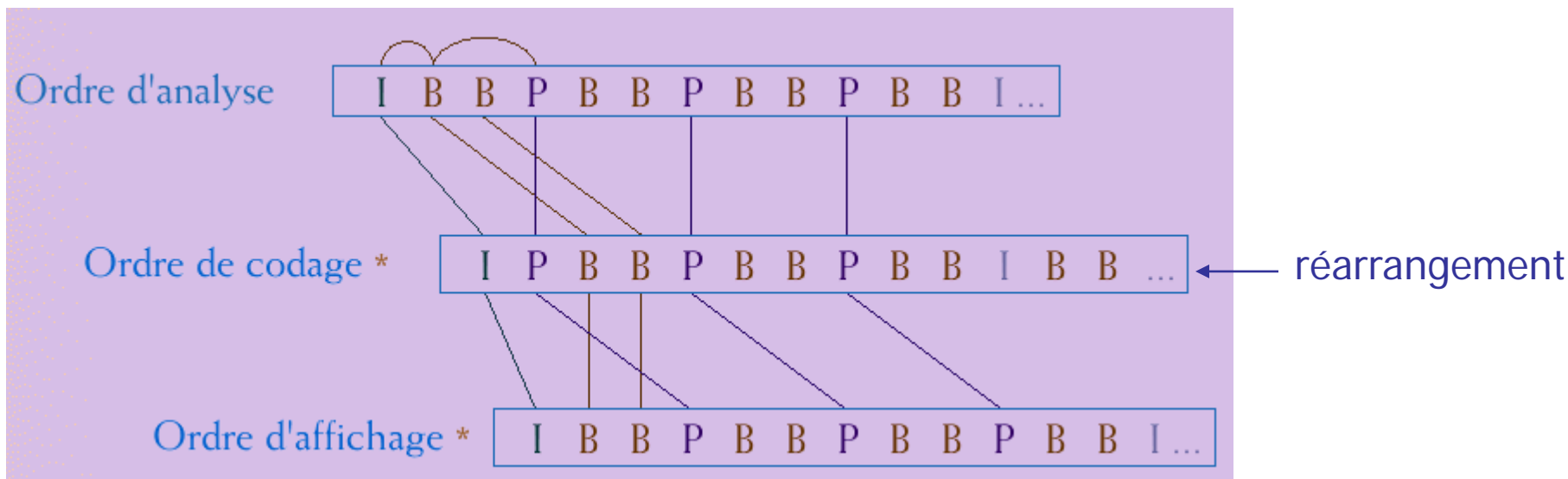
Le GOP

- Le GOP (*Group Of Picture*) définit l'enchaînement des images I, P et B dans le flux vidéo.
- Le GOP commence toujours par une image I.
- On note $\text{GOP}(M,N)$, où M est la distance entre deux images P et N la distance entre deux images I.

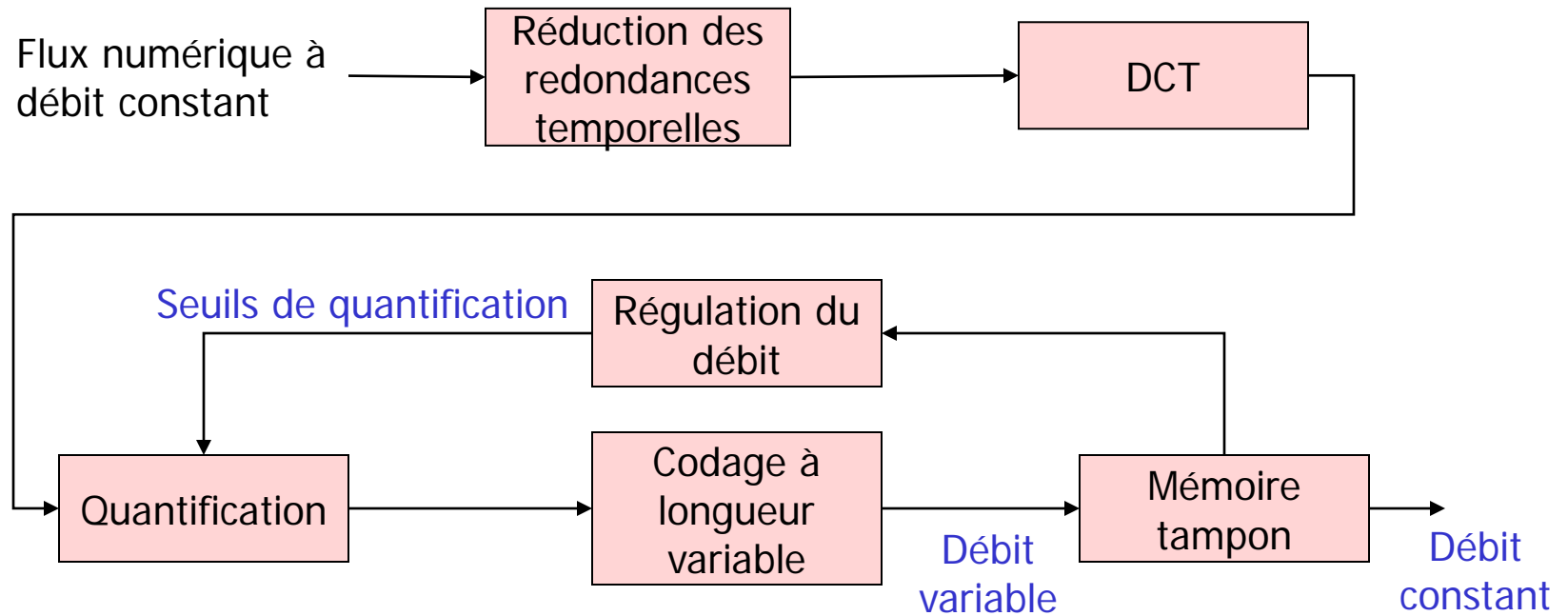


Réarrangement du GOP

- Si l'enchaînement des images contient des images B, le codage et décodage ne peuvent être effectués dans l'ordre d'arrivée de la séquence : **réarrangement** préalable des images.



Synoptique du codeur MPEG





La compression vidéo MPEG 1-2

- Etapes de la compression
- Spécificités des normes MPEG1 et MPEG2



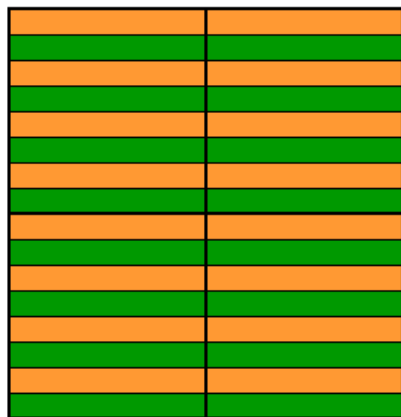
Paramètres des normes MPEG

- Structure d'échantillonnage
 - En théorie, un algorithme de compression MPEG accepte n'importe quel format en entrée (4:2:2, 4:2:0, SIF, SDTV, HDTV, ...).
 - La restriction se fait au niveau des profils (recommandations) décrits dans les normes.

Paramètres des normes MPEG

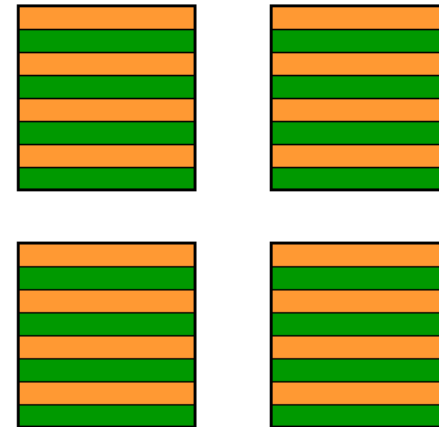
Gestion de l'entrelacement

- Dans un signal vidéo entrelacé, on peut créer les macroblocs en mode image ou en mode trame.
- \Rightarrow Image avec mouvement : Mode trame préférable car les trames sont peu cohérentes entre elles
- \Rightarrow Image sans mouvement : Mode image plus efficace car les trames sont fortement corrélées

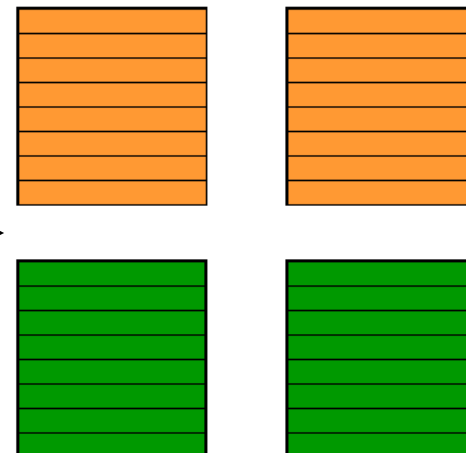


1 macrobloc = 4 blocs

Codage en mode image



Codage en mode trame



MPEG



La norme MPEG1

- Application : l'enregistrement
 - CD-ROM, disques optiques
 - VCD (magnétoscope numérique en Chine)
 - Utilisée par la TV numérique aux USA en attendant la mise au point du MPEG2
- Débit
 - Comprime des images animées + son stéréo avec un débit de 1.5Mbit/s
- Format d'entrée : en général, **format SIF**
 - Performances réduites pour un signal entrelacé (ex. : signal TV...)
- MPEG1 définit
 - 3 types d'images : I, P, B
 - 2 paramètres : GOP(M, N)



La norme MPEG2

- Le norme MPEG2 a été mise au point pour le **stockage et la diffusion de la vidéo numérique**.
- C'est la norme de compression utilisée dans la **télévision numérique terrestre** (TNT ou DVB-T).
- A amélioré MPEG1
 - En tenant compte de l'entrelacement
 - En rajoutant des outils de gestion des erreurs de transmission
- Formats d'entrée acceptés :
 - 4:2:2, 4:2:0, 4:4:4
 - EDTV, HDTV, futurs formats super HDTV
 - RVB et composantes
 - Formats informatiques
 - Balayage entrelacé ou progressif : à l'intérieur d'une même image, le codeur peut choisir de traiter des blocs trame à trame ou en désentrelacé.



La norme MPEG2

- S'adapte à diverses qualités d'images en télévision
 - 2Mbit/s : qualité comparable au VHS
 - 3 à 5 Mbit/s : qualité comparable au PAL
 - 8 à 10 Mbit/s : qualité comparable au 4:2:2
- Définit
 - 4 niveaux (définition des paramètres utilisés)
 - 6 profils (applications dédiées)
 - 3 types d'images : I, P, B
 - 2 paramètres : M et N
- La donnée d'un couple profil/niveau garantit l'interopérabilité des équipements.

Profils et niveaux de la norme MPEG2

Difficulté technique
croissante

CODEC bas
coût

TV numérique

Studio

4 niveaux *	6 profils *					
	Simple pas d'images B	Main 4:2:0	SNR TV améliorée	Spatial TVHD	High 4:2:2	4:2:2 post-prod.
High TVHD		80	structures »			100
High 1440 TVHD		60	60	60	80	
Main TV	15	15	15	15	20	50
Low 1/4 TV		4				

(débits en Mb/s)

MPEG

Qualité croissante 45



Scalabilité de la norme MPEG2

- La qualité et la résolution sont adaptables au mode de transmission vidéo. Entre autres :
- Scalabilité SNR (*Signal to Noise Ratio*)
 - Si le support de transmission est peu robuste, on peut dégrader la qualité de la vidéo et augmenter sa résistance au bruit (en gardant le même débit).
- Scalabilité spatiale
 - En attendant la migration des postes de TV de la SDTV à la HDTV, transmission simultanée des 2 qualités
 - \Rightarrow MPEG2 exploite la redondance entre le signal basse définition et le signal haute définition pour réduire la quantité d'information transmise.



La compression audio MPEG1 - 2

- Propriétés psycho-acoustiques exploitables pour la compression
- Les compressions audio MPEG 1 et 2

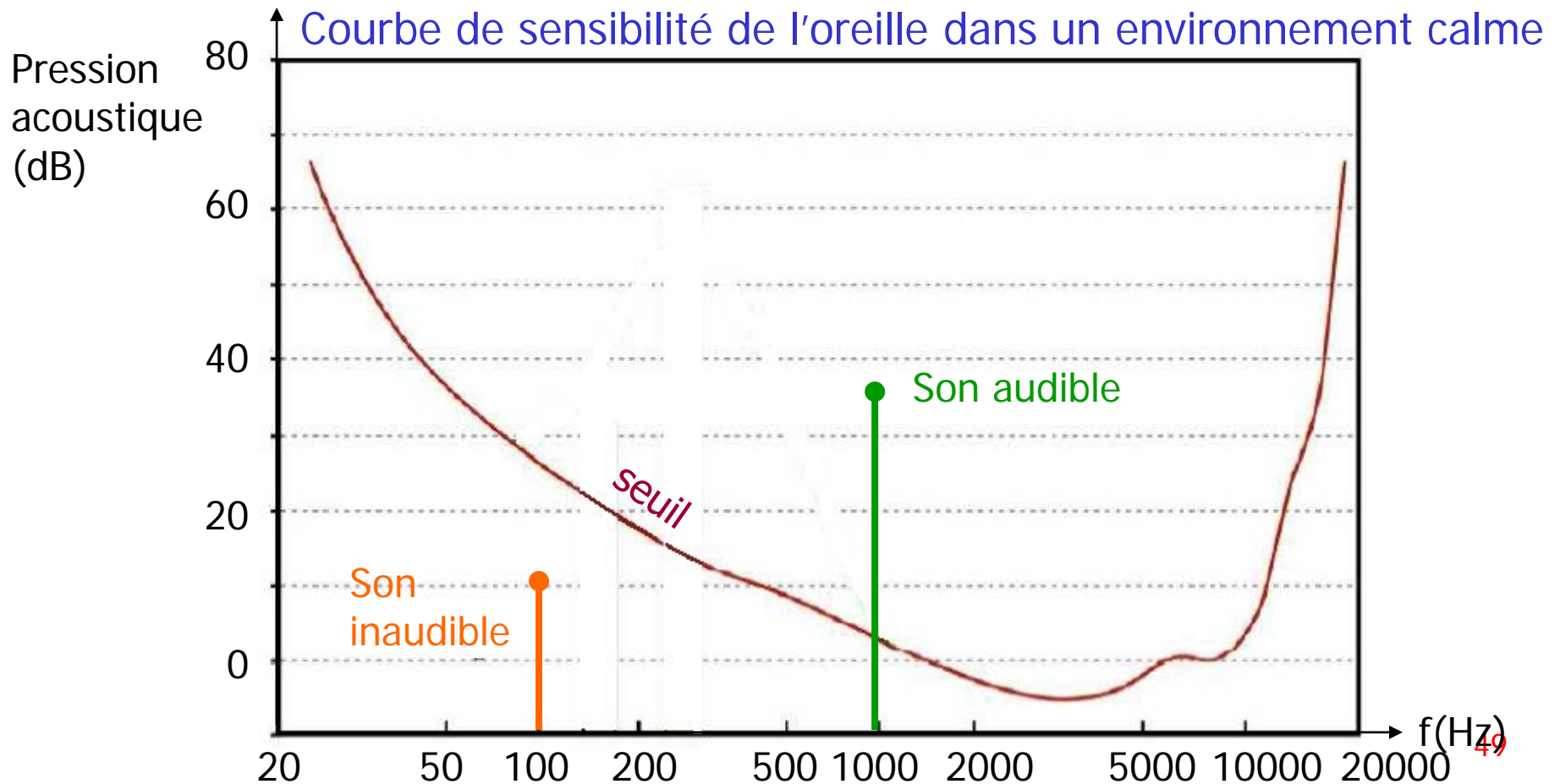


Le codage perceptuel

- Comme la compression vidéo MPEG utilise les failles de la vision humaine, la compression audio MPEG exploite les propriétés de l'audition humaine.
 - Compression de l'image : ne coder que l'information indispensable, dans la mesure où l'œil est un capteur imparfait.
- ⇒ Compression du son : ne coder que l'information indispensable, dans la mesure où l'oreille est un capteur imparfait.

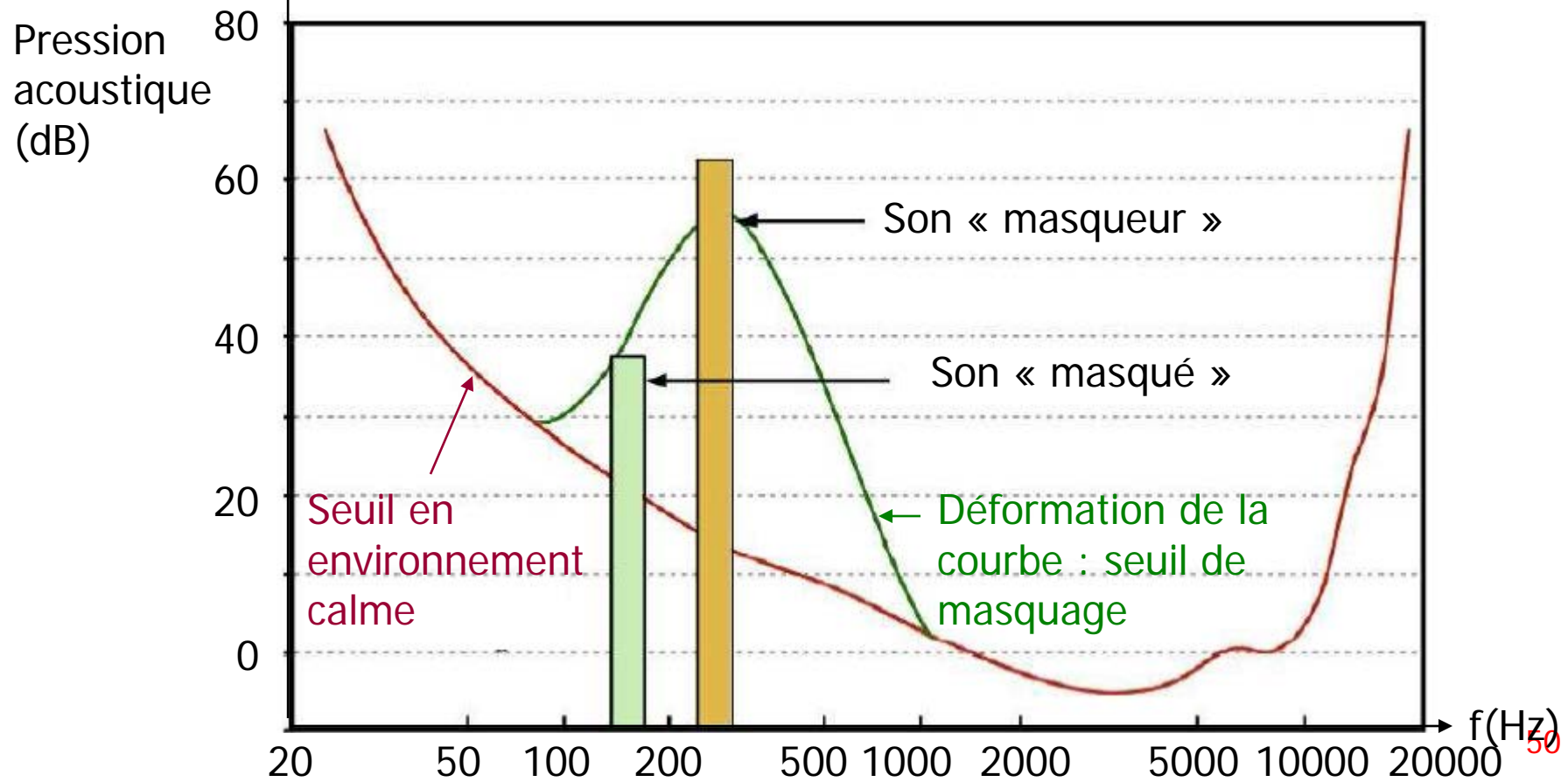
Courbe de sensibilité de l'oreille humaine

- A une fréquence donnée, notre oreille ne peut percevoir un son qu'à condition que son niveau de pression acoustique soit supérieur à un seuil. Le seuil est différent pour chaque fréquence de la bande audible.



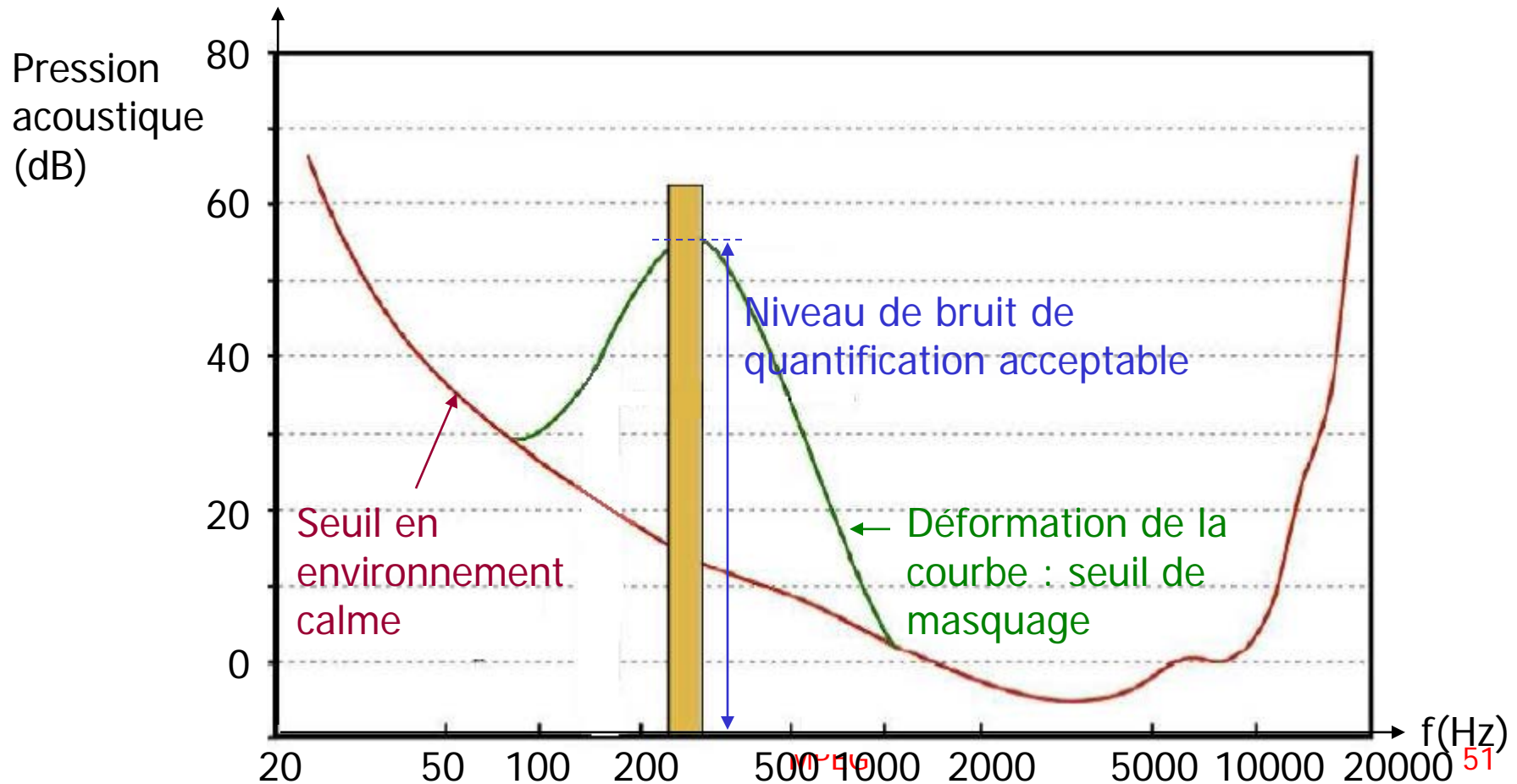
Le masquage fréquentiel

- Sous l'effet d'une excitation, le seuil d'audibilité minimum est amplifié au voisinage de la fréquence du son.
- Par conséquent, un autre son de fréquence proche, émis en même temps, peut être inaudible.



Exploitation du masquage en fréquence pour la compression

- On ne code pas les sons masqués (inaudibles)
- On adapte la profondeur de codage : nombre de bits minimal pour que $N < \text{seuil d'audibilité}$.





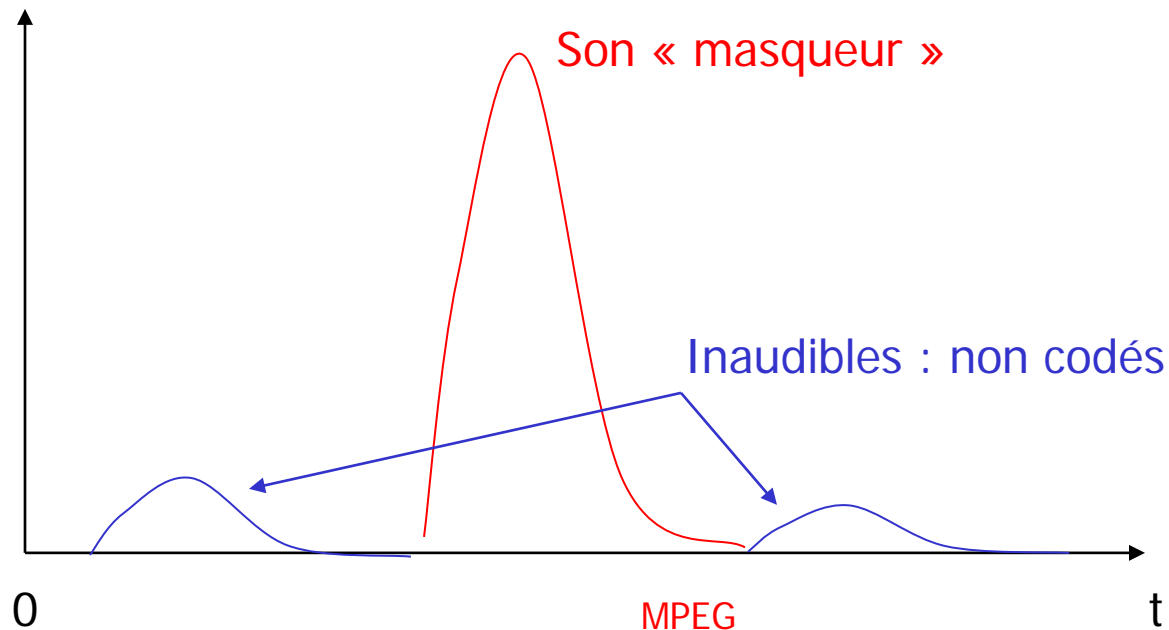
Exploitation du masquage fréquentiel par les normes MPEG1 et 2

- Le spectre audible est divisé en 32 sous-bandes.
- Pour chaque sous-bande, on détermine la fréquence dominante et on en déduit la nouvelle courbe de seuil d'audibilité (à l'aide de modèles psycho-acoustiques).
Ces modèles sont sauvegardés dans la ROM des codeurs.
- Dans chaque sous-bande, le codeur choisit le pas de quantification adapté au niveau du seuil d'audibilité.
- Le flux numérique contient les valeurs du pas de quantification de chaque sous-bande pour que le décodage soit possible.

Le masquage temporel

- L'oreille n'entend pas un son s'il est produit dans un court délai avant ou après un son **de même fréquence** de plus grande amplitude.
- ⇒ On ne code pas les sons masqués temporellement.

Pression
acoustique





Le masquage de canal

- L'oreille humaine est incapable de détecter la provenance d'un son s'il a une fréquence trop basse.
- \Rightarrow Coder en mono les sons basse fréquence.



La compression audio MPEG

- Propriétés psycho-acoustiques exploitables pour la compression
- Les compressions audio MPEG 1 et 2



La compression audio MPEG1

- Fréquences d'échantillonnage : 32 à 48 kHz
- Débits : 32 à 384 kbit/s
- Trois niveaux (*layers*) sont définis

Niveau	Technique	Utilisation - Performances
Niveau 1	Masquage fréquentiel	Qualité CD
Niveau 2	Masquage fréquentiel Masquage temporel	Qualité CD Utilisé dans le DVB-S
Niveau 3	Masquage fréquentiel avec largeur des sous-bandes adaptée Masquage temporel Codage entropique Exploitation de la cohérence entre les voies stéréo	« MP3 »



La compression audio MPEG2

- Fréquences d'échantillonnage : 16 à 48 kHz
- Débits : 8 à 384 kbit/s
- Elle reprend l'essentiel de la norme MPEG1 audio.
- Gère les signaux multicanaux jusqu'à 5 voies

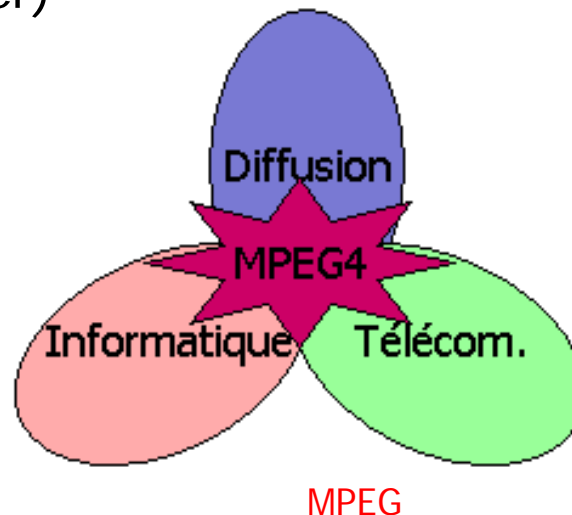


La compression MPEG4

- L'approche objet
- Le codage de la vidéo
- L'intégration de la 3D
- Le codage des images de synthèse
- La norme MPEG4 audio
- La description de scène
- Profils et niveaux

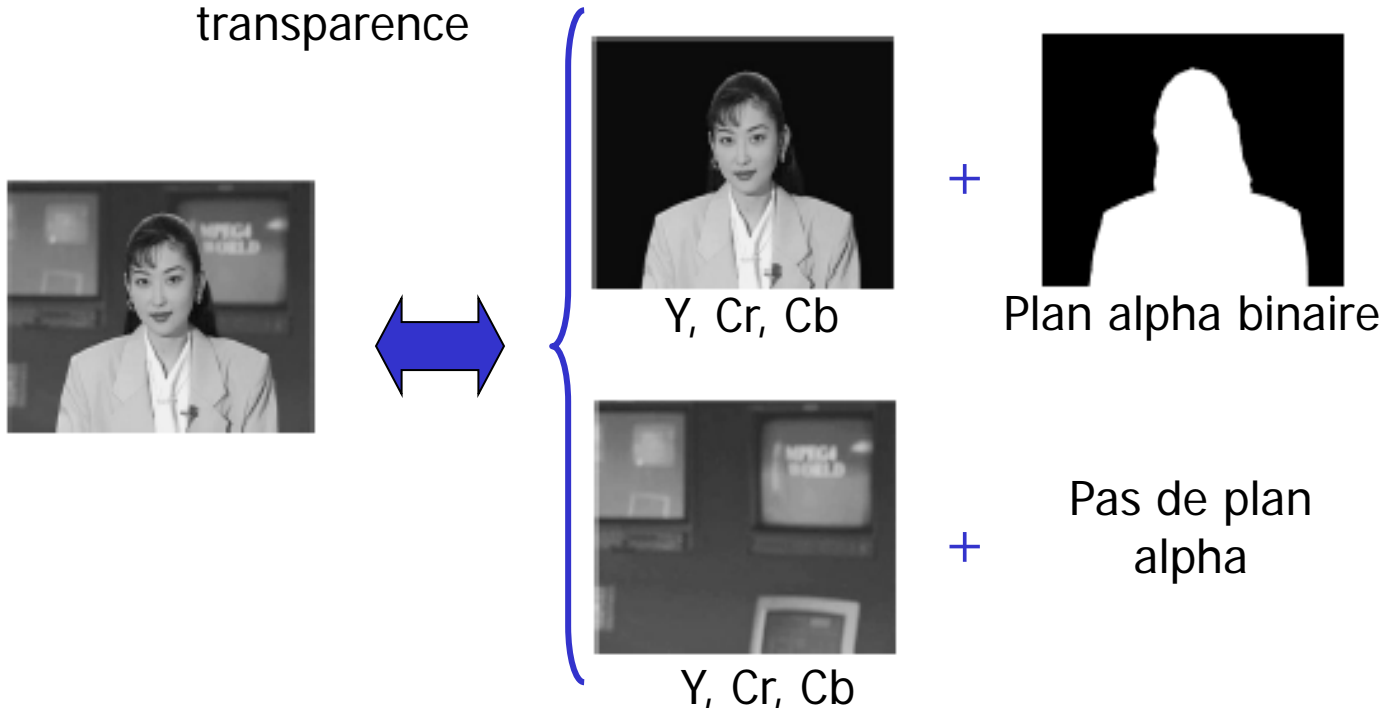
L'approche objet

- Nouveauté majeure : découpe des scènes en « objets »
- Intérêt : MPEG4 se veut une norme de compression
 - pour les applications **multimédia**,
 - supportée sur **n'importe quel type de support** (⇒ émergence récente de la mobilité)
 - et gérant **l'interactivité** (possibilité de modifier la composition d'une scène et les objets dans le flux numérique sans avoir à les décoder)



Codage de la vidéo : le plan alpha

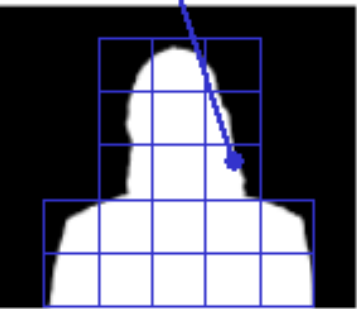
- Un objet est décrit par
 - Ses composantes Y, Cr, Cb : codées comme en MPEG 1 et 2 (sauf images fixes, cf. supra)
 - Un plan alpha
 - Codé en binaire : 1 à l'intérieur de l'objet, 0 ailleurs
 - Ou codé sur 8 bits : définit en plus des niveaux de transparence



Codage de la vidéo : le plan alpha

- Codage d'un plan alpha binaire : technique du codage de contour

16×16 pixels



- L'objet est découpé en blocs 16×16 pixels
 - Plusieurs (7) types de BAB (Bloc Alpha Binaire) sont définis par la norme, en fonction du mouvement et de la « déformation » subis dans la scène
 - Le codage de chaque type de BAB est défini par la norme : Intra, ou Inter avec compensation de mouvement
 - Chaque bloc de l'objet est donc codé suivant son BAB
- Codage d'un plan alpha sur 8 bits
 - Composante binaire : par technique de codage de contour
 - Les niveau de gris des blocs intérieurs (= niveaux de transparence) sont ensuite codés par DCT.

Codage d'une vidéo : les sprites

- **Sprite** : « une grande image composée de tous les pixels d'un objet donné visibles à un moment ou un autre de la séquence »
 - Ex. : un sprite d'objet « fond » = une vue panoramique d'un paysage sans les personnages passant au premier plan
- C'est une image fixe : elle est codée par DCT en Intra et émise au début du flux MPEG.



- **Intérêt**
 - Economie de bits
 - Manipulations intéressantes sur la vidéo

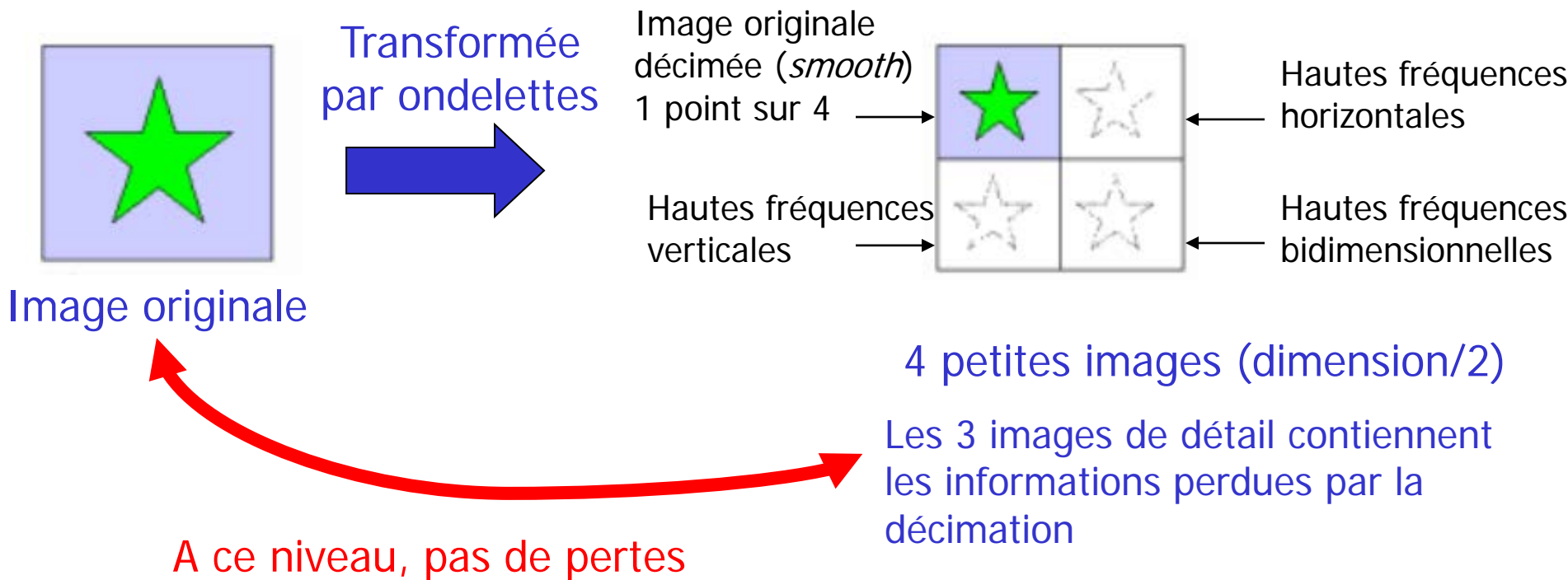


L'intégration de la 3D

- Pour habillage des structures 3D par des images naturelles (fixes), MPEG4 utilise la technique de compression par **ondelettes**, plus efficace que la DCT.
- La compression par ondelettes consiste en une analyse multi-résolution de l'image.

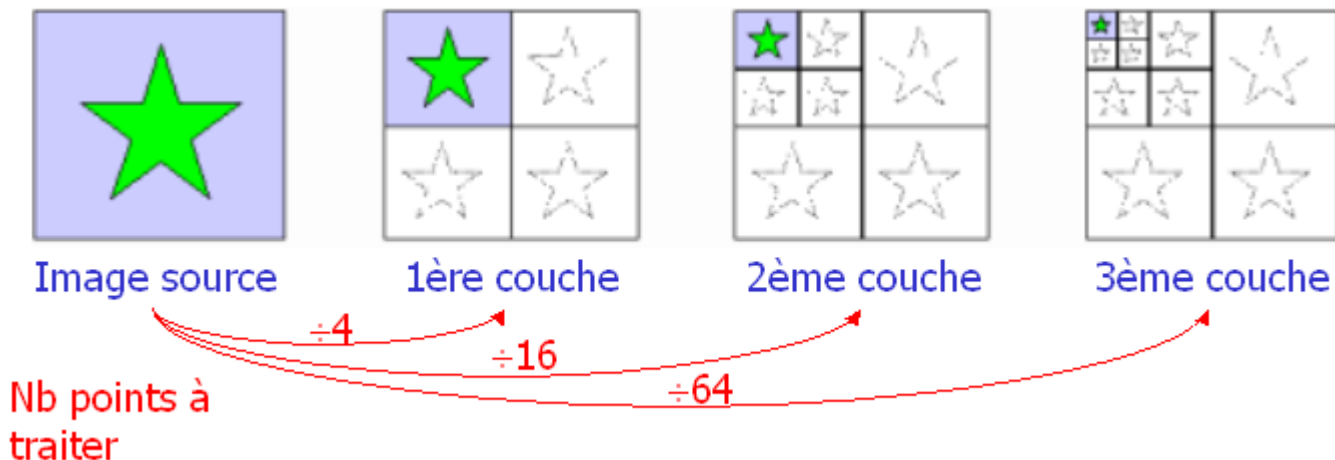
La compression par ondelettes

- Réduction de la résolution de l'image par décimation
- Séparation des détails horizontaux, verticaux, et bidimensionnels par filtrage



La compression par ondelettes

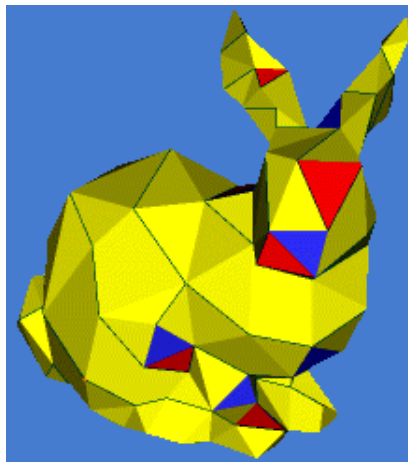
- On peut réitérer l'opération sur l'image décimée (smooth) plusieurs fois : analyse « multicouches ».
 - L'image smooth est plus petite à chaque couche : moins d'informations à coder
 - On peut **quantifier** les images de détails et leur appliquer un codage entropique.



- Intérêt pour la 3D : le multicouche permet d'adapter le rendu des détails à la distance de visualisation (position de l'utilisateur dans l'environnement 3D).

Le codage des images de synthèse

- La norme inclut des algorithmes permettant la description des visages et leur animation.
- Elle permet aussi la construction d'objets par treillis actifs
 - Les objets sont décrits par un treillis triangulaire
 - Le décodeur calcule l'évolution de la texture d'une scène à l'autre pour chaque triangle par interpolation



Un treillis



La norme MPEG4 audio

- Codage des sons naturels
 - Débits de 2 à 64 kbit/s
 - La norme n'a pas pu proposer 1 seul algorithme pour tous les types de sons (musique, parole, etc. ...)
 - ⇒ Elle inclut plusieurs standards de codage, adapté à chaque type de signal sonore
 - Les algorithmes proposés dans la norme MPEG2 audio sont notamment repris.

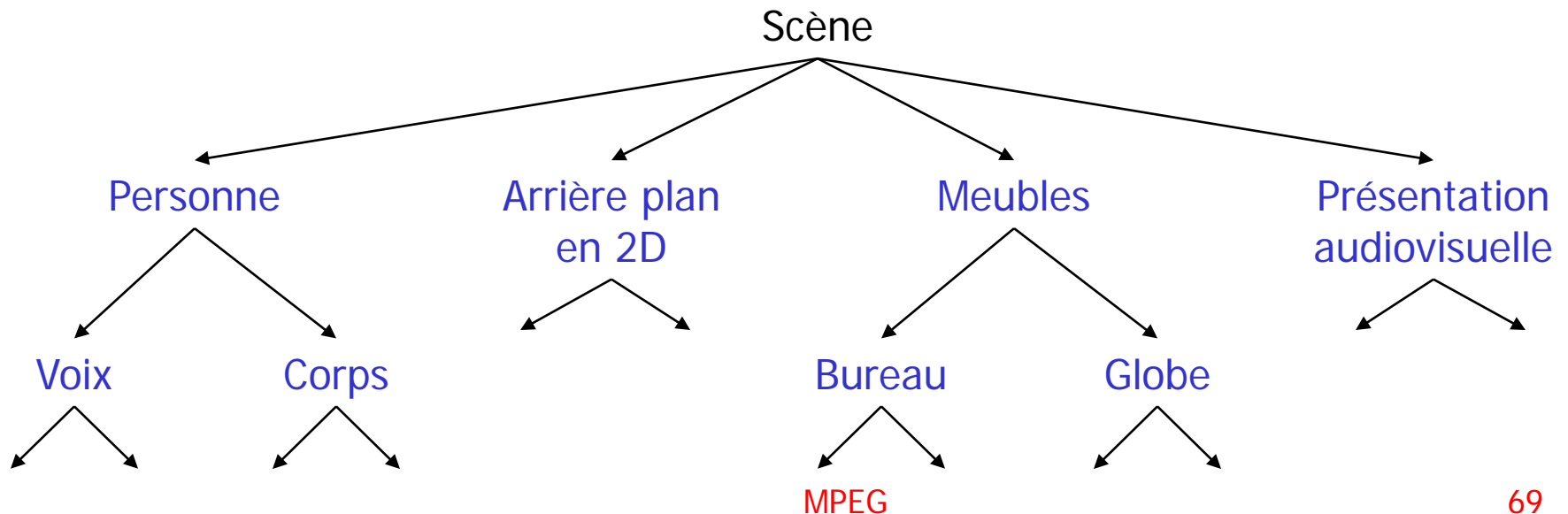
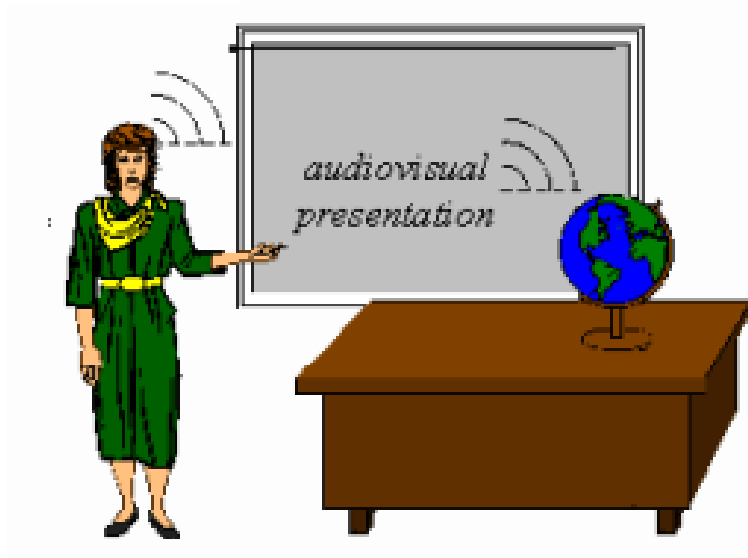
- Synthèse de sons : la norme inclut plusieurs algorithmes de génération de sons (pluie, bruits de pas, instruments de musique...).



Description de scènes

- Les scènes sont organisées sous forme d'une structure hiérarchique d'objets.
- La norme inclut un standard de description de la scène en flux binaire : BIFS (*Binary Format for Scene Description*).
- Elle permet l'utilisation du VRML (langage de modélisation de la réalité virtuelle) pour permettre l'interface de la 3D et de la vidéo.

Description de scène : un exemple





Profils et niveaux MPEG4

- 5 profils
 - Profils visuels : codage visuel des données naturelles, synthétiques ou hybrides
 - Profils audio : nombre d'outils implémentés et le débit visé
 - Profils graphiques : éléments graphiques et textuels utilisables dans une scène
 - Profils de descripteurs de scène : types d'informations pouvant constituer une scène MPEG4 (audio, 2D, 3D...)
 - Profils de descripteurs d'objets : outils disponibles pour décrire un objet (descripteur d'objet, synchronisation, information sur le contenu des objets)

Profils et niveaux MPEG4

	NIVEAU 1	NIVEAU 2	NIVEAU 3	NIVEAU 4	NIVEAU 5
VISUEL (naturels)	SIMPLE VISUAL	SIMPLE SCALABLE	CORE VISUAL	MAIN VISUAL	N-BIT VISUAL
VISUEL (synthétiques Hybrides)	SIMPLE FACIAL ANIMATION VISUAL	SCALABLE TEXTURE VISUAL	BASIC ANIMATED 2D TEXTURE VISUAL	HYBRID VISUAL	
AUDIO	SPEECH	SYNTHESIS	SCALABLE	MAIN	
GRAPHIQUE	SIMPLE 2D GRAPHICS	COMPLETE 2D GRAPHICS	COMPLETE GRAPHICS		
DESCRIPTEUR DE SCENE	AUDIO	SIMPLE 2D	COMPLETE 2D	COMPLETE	
DESCRIPTEUR D'OBJET	MAIN				



Bibliographie

- « Compression Numérique : JPEG, DV, MPEG », support de cours, INA Formation, mars 2003
- « Compression MPEG-1 à MPEG-4 », E. Fert, S. Jeannin, Techniques de l'Ingénieur - Traité Télécommunications
- « Compression des images : nouveaux besoins, nouvelles technologies », support de cours, J. Weiss, Supelec, 2000