

# Compression video

F. Denis  
F. Dupont



# Objectifs du cours

---

- ✖ Présenter les normes actuelles en compression vidéo
- ✖ Etudier les principes fondamentaux de la compression et la spécificité de la vidéo
- ✖ Mieux comprendre les enjeux pour les télécommunications

# SIGNAL VIDÉO

- 
- ~1950 ✕ **NTSC :**  
standard TV utilisé en Amérique du nord et au Japon  
**525 lignes**  
fréquence des trames : 60Hz → fréquence d'affichage : **30 images/s**
  - 1956 ✕ **SECAM (brevet Henri de France)**  
standard français  
**625 lignes (dont 576 lignes visibles)**  
fréquence des trames : 50Hz → fréquence d'affichage : **25 images/s**
  - 1962 ✕ **PAL (brevet Walter Bruch)**  
standard TV couleur, mis point en Allemagne vers 1960 et utilisé en Europe (sauf en France), Asie, Afrique, Australie, Amérique du sud, etc.  
**625 lignes (576 lignes affichées)**  
fréquence des trames : 50Hz → fréquence d'affichage : **25 images/s**
  - 1982 ✕ **Standard CCIR 601** ( CCIR = *Comité Consultatif International des radiocommunications*)  
**standard vidéo numérique** référencé désormais par ITU-R BT.601

# SIGNAL VIDÉO

## Signal vidéo numérique

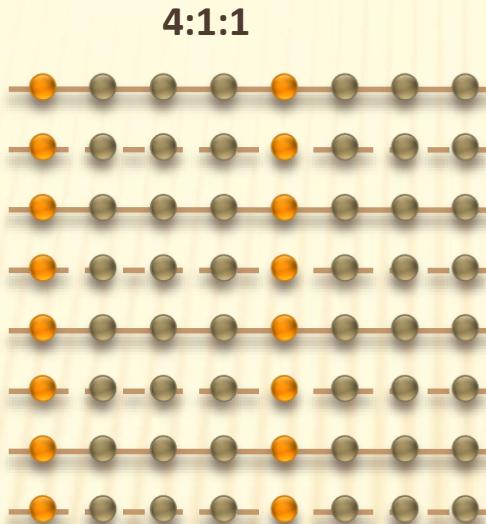
→ différents formats d'image possibles



Y : 8 bits  
Cb : 8 bits 1 pixel sur 2  
Cr : 8 bits 1 pixel sur 2  
→ En moyenne 16 bits par pixel



TV numérique professionnelle

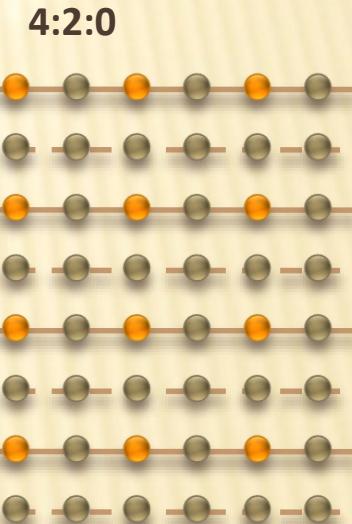


Y : 8 bits  
Cb : 8 bits 1 pixel sur 4  
Cr : 8 bits 1 pixel sur 4  
→ En moyenne 12 bits par pixel



Magnétoscopes et caméscopes, DVD, télévision numérique (TNT, satellite,...)

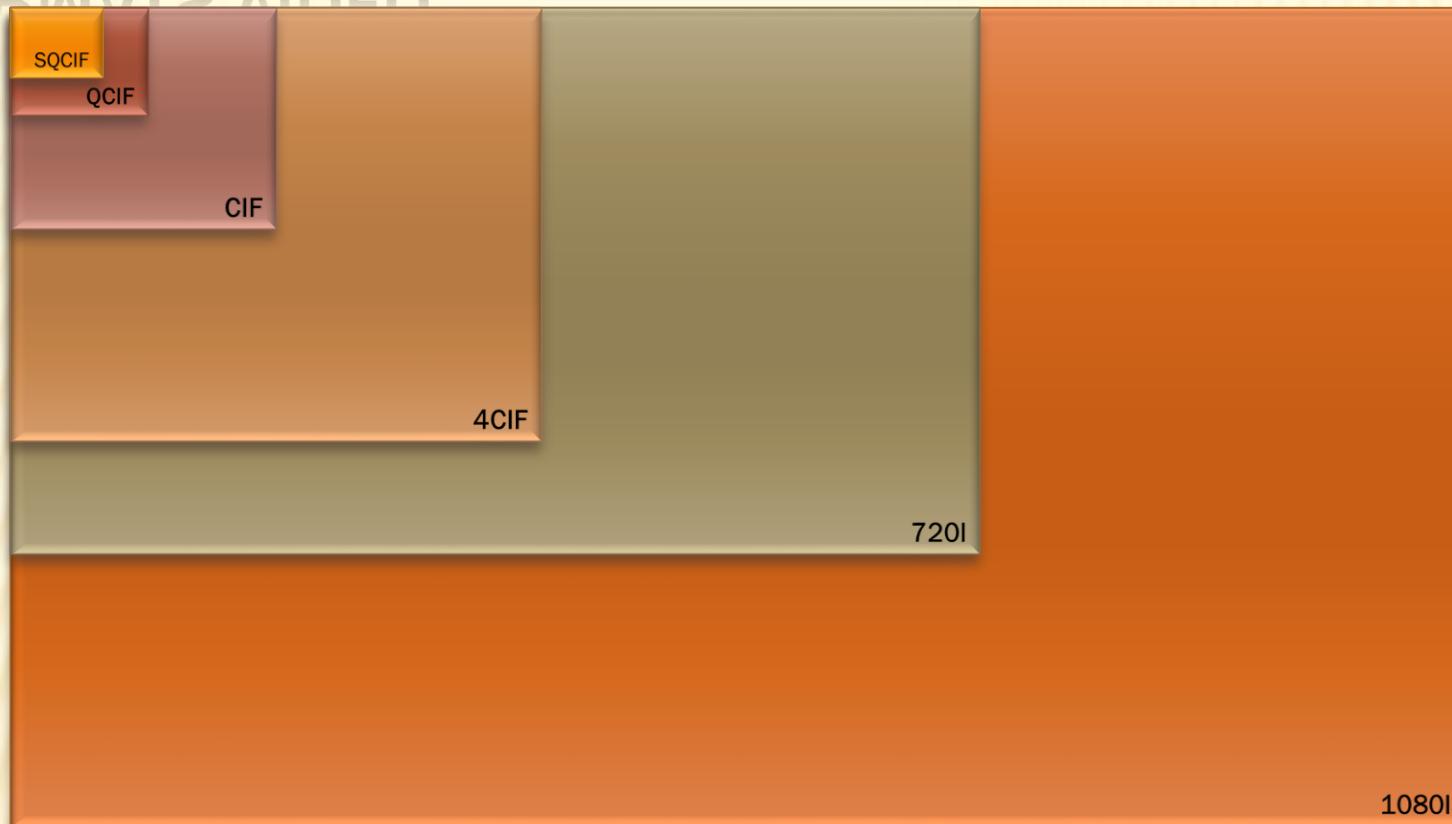
- Luminance
- Luminance + chrominance



Y : 8 bits  
Cb : 8 bits 1 pixel sur 4  
Cr : 8 bits 1 pixel sur 4  
→ En moyenne 12 bits par pixel



# FORMATS VIDEO



Format	Résolution Luminance (Horizontal × Vertical)	Bits par image (YV12, 8bit)	Application
SQCIF	128 × 96	147456	Vidéo Mobile
QCIF	176 × 144	304128	Vidéo Conférence
CIF	352 × 288	1216512	Monitoring Vidéo
4CIF	704 × 576	4866048	TV, DVD
720i50	1280 × 720	11059200	HD TV
1080i50	1920 × 1080	24883200	HD DVD

**4K**

**2K**

**1080p**

**720p**

**DVD**  
**VCD**

# COMPRESSION VIDÉO

- ✖ **Signal vidéo numérique non compressé**  
en format 4/3 : 720 pixels sur 576 lignes actives  
→  $720 * 576 * 25 * 16 \approx 166$  Mbits/s en 4:2:2  
→  $720 * 576 * 25 * 12 \approx 124$  Mbits/s en 4:1:1 ou 4:2:0
  
- ✖ **Débits actuels en vidéo compressée**  
→ moins de 2 Mbits/s

# COMPRESSION VIDÉO

## ✖ Réduction de la redondance

- + Spatiale
  - + Statistique
  - + Temporelle
- } cf. compression d'images fixes

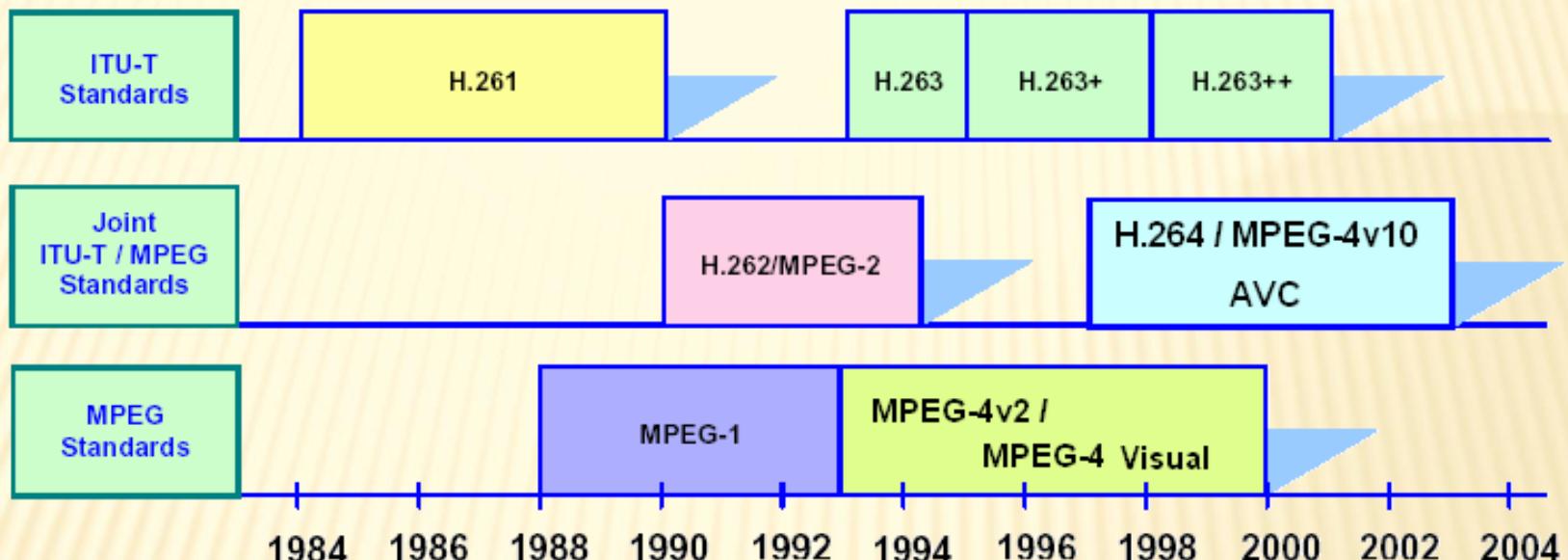
## ✖ Compression JPEG image par image

- M-JPEG ou Motion JPEG
- + accès aléatoire aux images → adapté au montage numérique

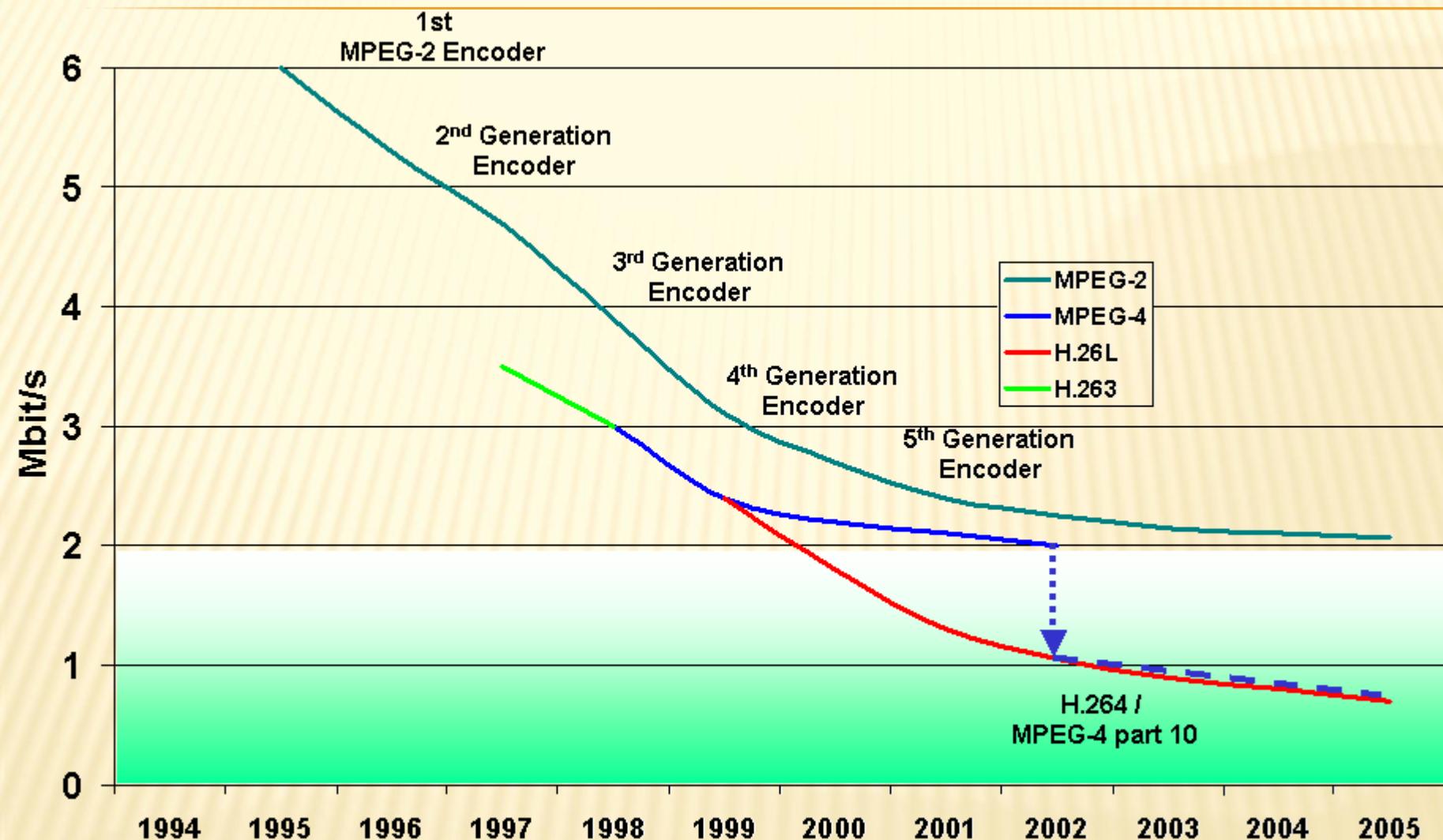
## ✖ Compression par compensation de mouvement

- MPEG (Moving Pictures Expert Group)

# STANDARDS ET NORMES VIDÉOS



- ✖ **ITU-T propose les standards H.26x**
  - + meilleur que MPEG4 en 2001
- ✖ **MPEG et ITU-T s'allient pour former le Joint Video Team**
  - + continuent de développer H.26x
  - + proposent H.264 et MPEG-4 Part 10 connu sous le nom standard JVT et Advanced Video Coding (AVC)



# STANDARDS ET NORMES VIDÉOS

- ✖ **MPEG-1, développé en 1988**
  - + compression des données vidéos et des canaux audio associés (jusqu'à 2 canaux pour une écoute stéréo)
  - + stockage de vidéos à un débit de 1.5Mbps dans une qualité proche des cassettes VHS sur un support CD appelé VCD (Vidéo CD).
- ✖ **MPEG-2**
  - + standard dédié à l'origine à la télévision numérique (HDTV)
  - + qualité élevée à un débit pouvant aller jusqu'à 40 Mbps, et 5 canaux audio
  - + Permet une identification et une protection contre le piratage.
  - + format utilisé par les DVD vidéos.
- ✖ **MPEG-4**
  - + standard destiné aux applications multimedias (codage de données multimédia sous formes d'objets numériques)
  - + plus grande interactivité → adapté au Web et aux périphériques mobiles

# **STANDARDS ET NORMES VIDÉOS**

## **✖ MPEG-7**

- + standard visant une représentation standard des données audio et visuelles afin de rendre possible la recherche d'information**
- + également intitulé Multimedia Content Description Interface.**

## **✖ MPEG-21, en cours d'élaboration**

- + destiné à standardiser la gestion de ces contenus, les droits d'accès, les droits d'auteurs, ... et à fournir un cadre de travail pour l'ensemble des acteurs du numériques (producteurs, consommateurs, ...)**

# MPEG1

- **Objectif :**
  - compression des plans (redondance spatiale)
  - compression inter-plan (redondance temporelle)
- **Approche simple : généralisation du savoir-faire 2D**
  - traiter les plans successifs de manière indépendante
  - traiter la différence des plans successifs
- **Approche plus adaptée : tenir compte du mouvement**
  - estimer le mouvement entre 2 images successives
  - compenser le mouvement
  - traiter la différence des image  $i$  et image  $(i+1)$  compensée

# ESTIMATION DU MOUVEMENT PAR BLOC

- ✖ Block Matching : recherche de blocs similaires
  - + l'image instant i : image de référence
  - + pour chaque bloc (fils) de l'image instant (i+1), recherche d'un bloc similaire (père) dans l'image i, autour de la position centrale du bloc fils
  - + ressemblance entre père et fils mesurée par une distance D ( $L_2$  ou  $L_1$ ) à minimiser
  - + V : vecteur de déplacement du bloc père

# ESTIMATION DU MOUVEMENT PAR BLOC

Image i

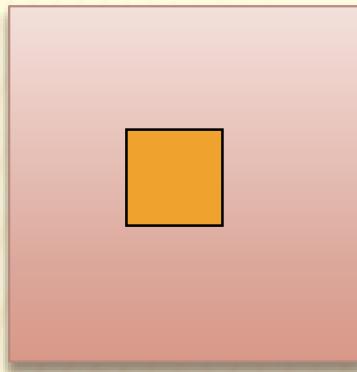
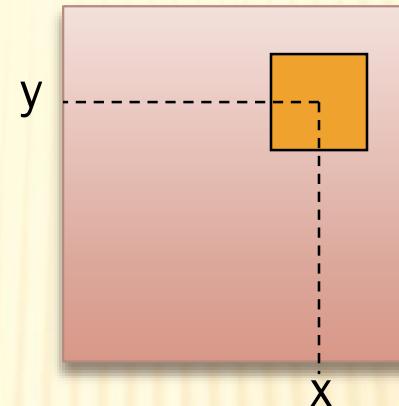
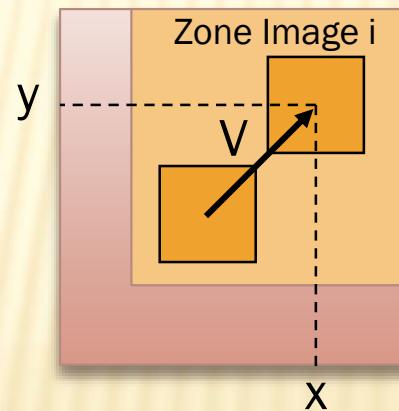


Image  $i+1$



Zone de recherche  
dans l'image i  
autour de  $(x,y)$



Distance minimale D  
Vecteur de déplacement  
du bloc père V

# ESTIMATION DU MOUVEMENT PAR BLOC

## ✖ exemple

trame n



trame n+1



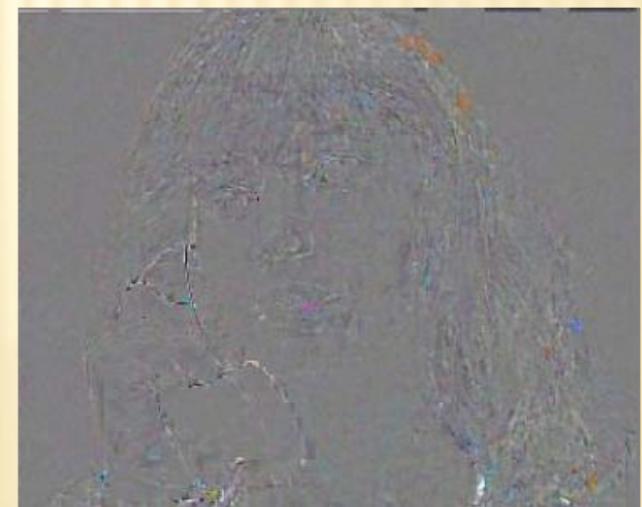
résidu



Vecteurs mouvement



Résidu avec compensation de mouvement



# BLOCK MATCHING : TYPES DE BLOC

- Bloc compensable
  - D est inférieure à un seuil S, alors il y a ressemblance
  - On ne transmet que le vecteur V
- Bloc non compensable
  - D est supérieure au seuil S, alors ressemblance très faible (changement de luminosité, partie cachée) ou c'est un nouveau bloc
  - soit on transmet V et le bloc différence codé JPEG
  - soit on transmet comme un nouveau bloc codé JPEG

# MPEG1 : INTERPOLATION BIDIRECTIONNELLE

## + 3 types de trames

### ✗ I : codage Interne

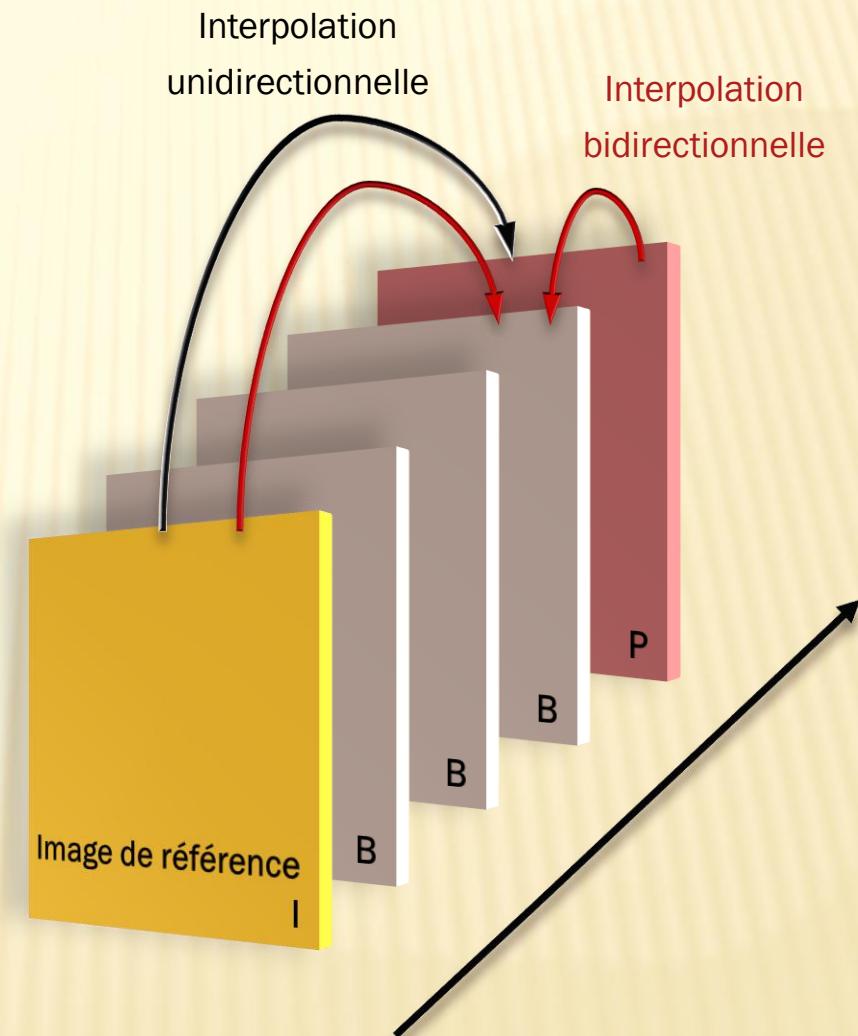
→ codées séparément sans faire référence aux images précédentes

### ✗ P : codage Prédicatif

→ décrites par différence avec les images précédentes

### ✗ B : codage prédictif Bidirectionnel

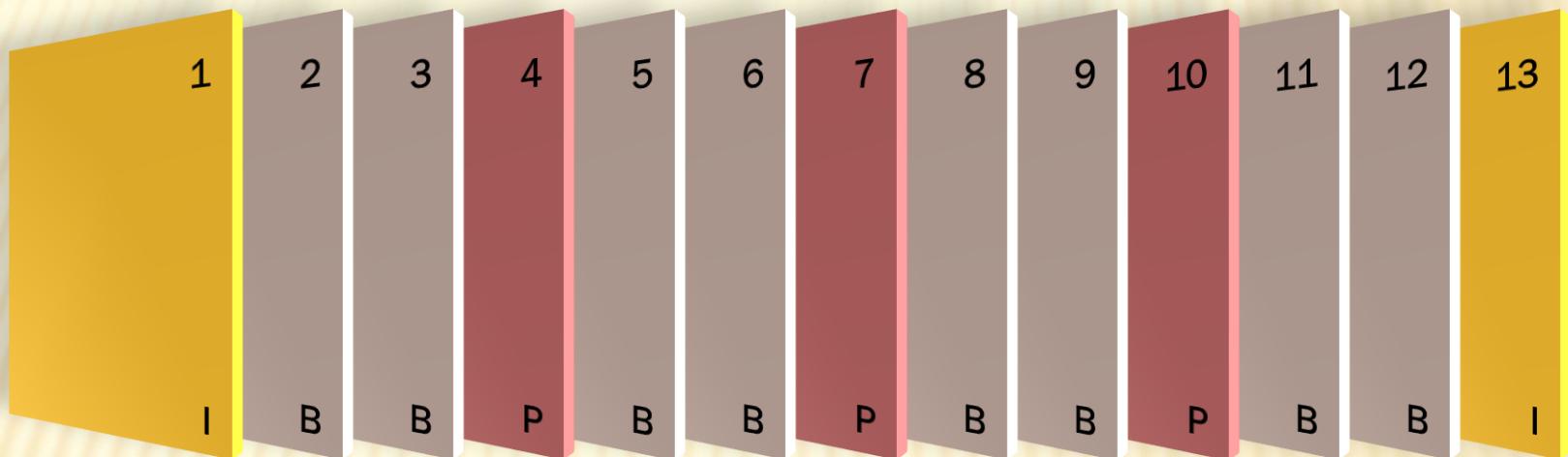
→ décrites par différence avec l'image précédente et l'image suivante



# GOP (Group Of Pictures) : EXEMPLE

Distance inter P : m=3

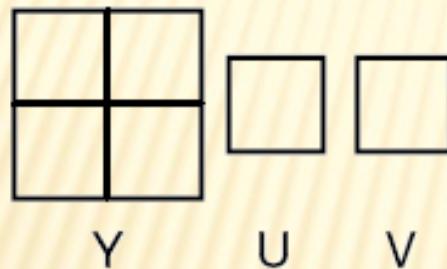
Distance inter I : n=12



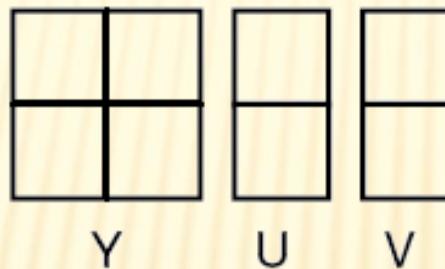
- ✗ **1 et 13 (I) sont complètes**
- ✗ **4 est prédite à partir de 1 par différence, de même 7 est prédite à partir de 4**
- ✗ **2 est interpolée à partir de 1 et 4, de même que 3**
- ✗ **Ordre de codage et décodage : 1 4 2 3 7 5 6 10 8 9 13 11 12 16 14 ...**

# FORMATAGE DES DONNÉES : MACROBLOCS

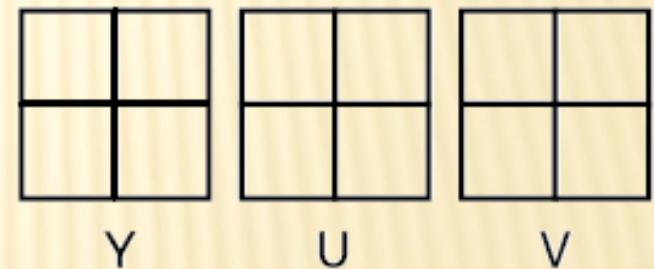
- + Image → rangées de 16 lignes de pixels
- + Rangée → macroblocs 16x16
- + Macroblock → 4 blocs 8x8 pour Y  
1, 2 ou 4 blocs 8x8 pour U et V



Macroblock 4 :2 :0

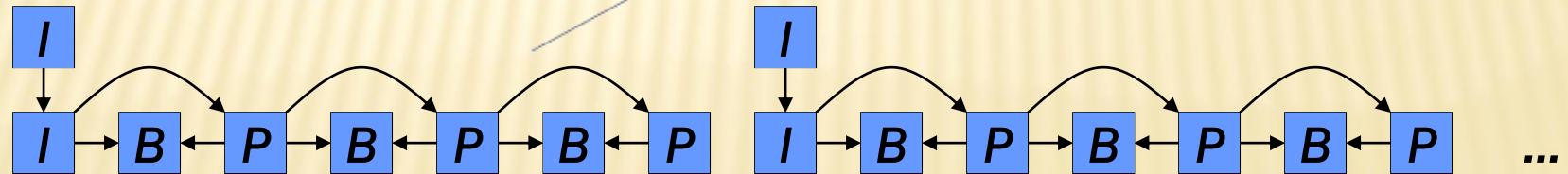
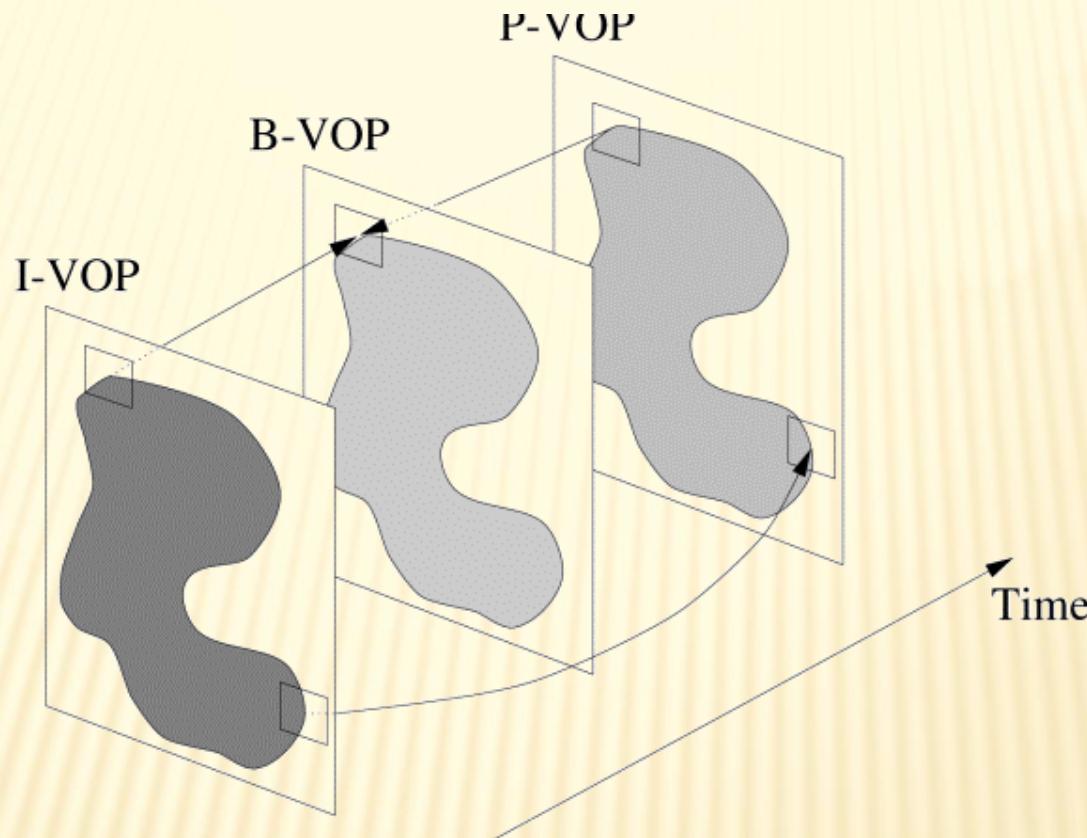


Macroblock 4 :2 :2



Macroblock 4 :4 :4

# INTERPOLATION BIDIRECTIONNELLE PAR BLOC



# MPEG1 : IMAGES I, P ET B

## ✖ Images I

- + Images de référence
- + Elles servent de points de synchronisation
- + Elles sont codées en JPEG, codage **INTRA**
- + Il ne doit pas y avoir plus de 400 ms entre deux images I

## ✖ Images P

- + Images prédites à partir d'une image I ou P → codage **INTER**
- + L'image erreur est codée par JPEG dérivé avec un fort taux de compression
- + Les vecteurs mouvement sont codés par DPCM
- + Le terme d'erreur est transformé par DCT
- + La quantification utilise un parcours zig-zag, RLE, Huffman
- + La table de quantification différente de celle des images I

## ✖ Images B

- + Images prédites à partir de 2 images I et P, avec prédiction bidirectionnelle
- + Codage identique à l'image P

# MPEG1 : APPLICATIONS / DÉBITS

- Applications
  - enregistrement/lecture sur support numérique (DAT, CD-ROM)
  - VCR
- Qualité souhaitée : cassette vidéo VHS
- Format d'image : SIF (quart d'image TV)
- Débits
  - débit total : 1.5 Mbits/s
  - 1.25 Mbits/s vidéo
  - 32 à 384 kbits/s audio

# MPEG2 : APPLICATIONS / DÉBITS

- ✖ Applications
  - + TV numérique définition standard (satelliten cable, ...)
  - + TVHD
  - + DVD
- ✖ Débits
  - + TV numérique entre 3 (qualité TV analogique) et 10 Mbits/s
    - ✖ luminance :  $720 * (486 \text{ ou } 576 \text{ lignes})$
    - ✖ chrominance :  $360 * (486 \text{ ou } 576 \text{ lignes})$
    - ✖ codage : 4:2:2
  - + TVHD entre 10 et 15 Mbits/s (débit max. pour MPEG2)
- ✖ Soutenu par le consortium DVB (Digital Video Broadcasting)

# DVB

---

- ✖ Les normes DVB sont définies par l'ETSI, the European Telecommunications Standards Institute.
- ✖ Le DVB a spécifié les normes pour les décodeurs :
  - + câble (DVB-C)
  - + satellite (DVB-S et DVB-S2 pour HDTV avec MPEG4-AVC)
  - + hertzien/terrestre (DVB-T)
  - + application portable (DVB-H, H pour « handheld »)
- ✖ Spécifications
  - + modulation de la porteuse (QPSK, 64-QAM, 256-QAM)
  - + protection contre les pertes (FEC, pour Forward Error Correction)
  - + codes correcteurs d'erreurs bit (e.g. algorithme de Viterbi, Reed-Solomon).

# MPEG2 : MPEG1 amélioré

- ✖ Les plus de MPEG2 :
  - + Prise en compte de signaux d'entrée entrelacés
  - + scalabilité → codage hiérarchique : contrôle de flux pour permettre au client d'ajuster le flux vidéo à sa configuration (en débit et en qualité)
    - ✖ temporel : en jouant sur le nombre des images
    - ✖ spatial : en jouant sur la taille des images
    - ✖ qualité : en jouant sur le pas de quantification
    - ✖ hiérarchie des coefficients DCT : on envoie d'abord les BF, puis les HF
  - + adaptation aux différents types d'applications
    - ✖ niveau : fréquence image, taille image, nombre de pixels ligne,..

# NORMES ITU-T

- ✖ Norme H.261 (1990) : vidéoconférence
- ✖ Common Intermediate Format (CIF) :
  - $Y = 352*288$ ,  $Cb/Cr = 176*144$
  - codage : 4:2:2
  - qualité vidéoconférence
- Quarter Common Intermediate Format (QCIF) :
  - $Y = 176*144$ ,  $Cb/Cr = 88*72$
  - codage : 4:2:2
  - qualité vidéophone

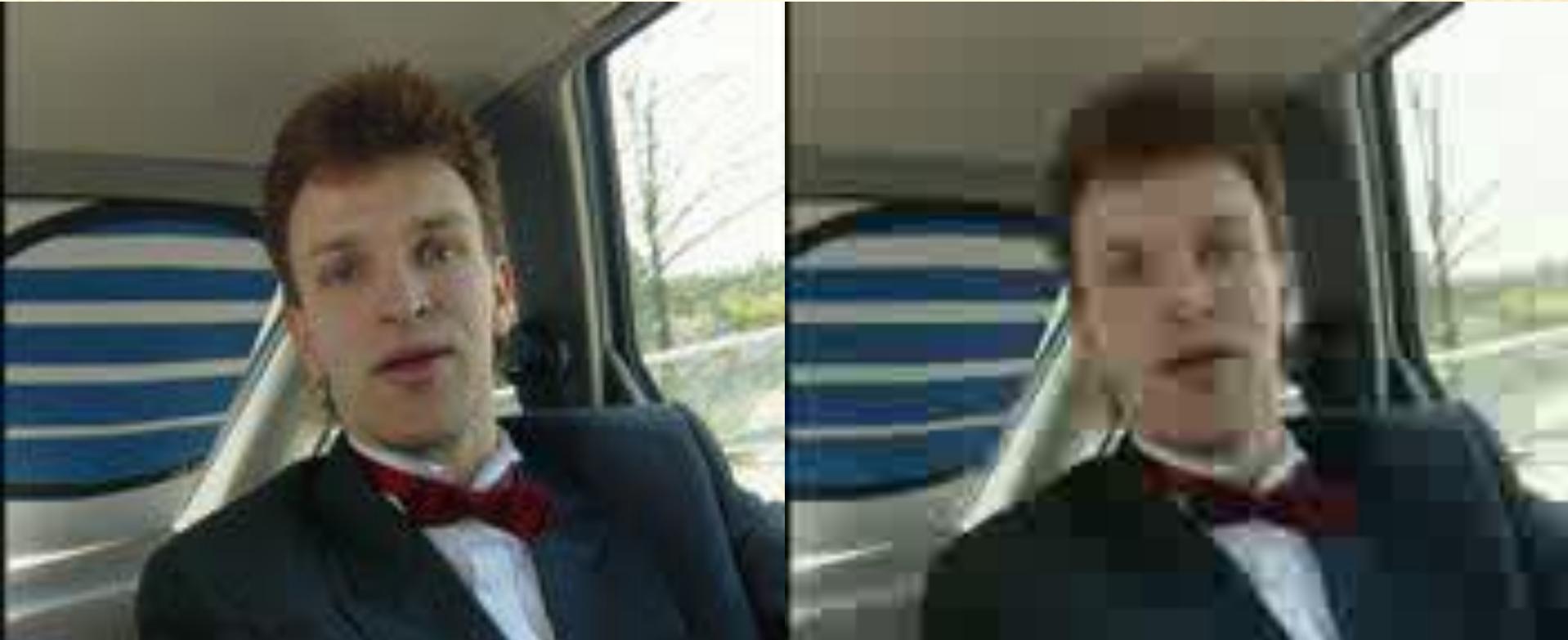
# NORMES ITU-T

- ✖ Norme H.261 : vidéoconférence
- ✖  $p * 64 \text{ kbits par sec. (} p = 1 \text{ à } 30 \text{)}$
- Méthode de compression :
  - Image de référence : codée par JPEG
  - Images suivantes : par prédition : l'erreur est codée par JPEG dérivée
  - pas de recommandation pour le rafraîchissement
- Norme H.263 (1995) :
  - basée sur H.261; améliore certains points comme la prédition
  - vise les applications visiophone
- Norme H.264 (2002) :
  - basée sur H.261 ; objectif : très bas débit, terminaux mobiles
  - couche codage vidéo
  - couche d'abstraction de réseau (NAL) avec des entêtes adaptées aux couches de transport)

# VIDÉO : BANDES PASSANTES

<i>Qualité</i>	<i>Technique</i>	<i>Bande passante en Mbps</i>
HDTV (1920x1080/60)	non compressé	2000
HDTV	MPEG-2	25 à 34
Studio TV	non compressé	166
Studio TV	MPEG-2	3 à 6
TV Broadcast	MPEG-2	2 à 4
VCR	MPEG-1	1.2
Vidéo-conférence	H.261	0.1

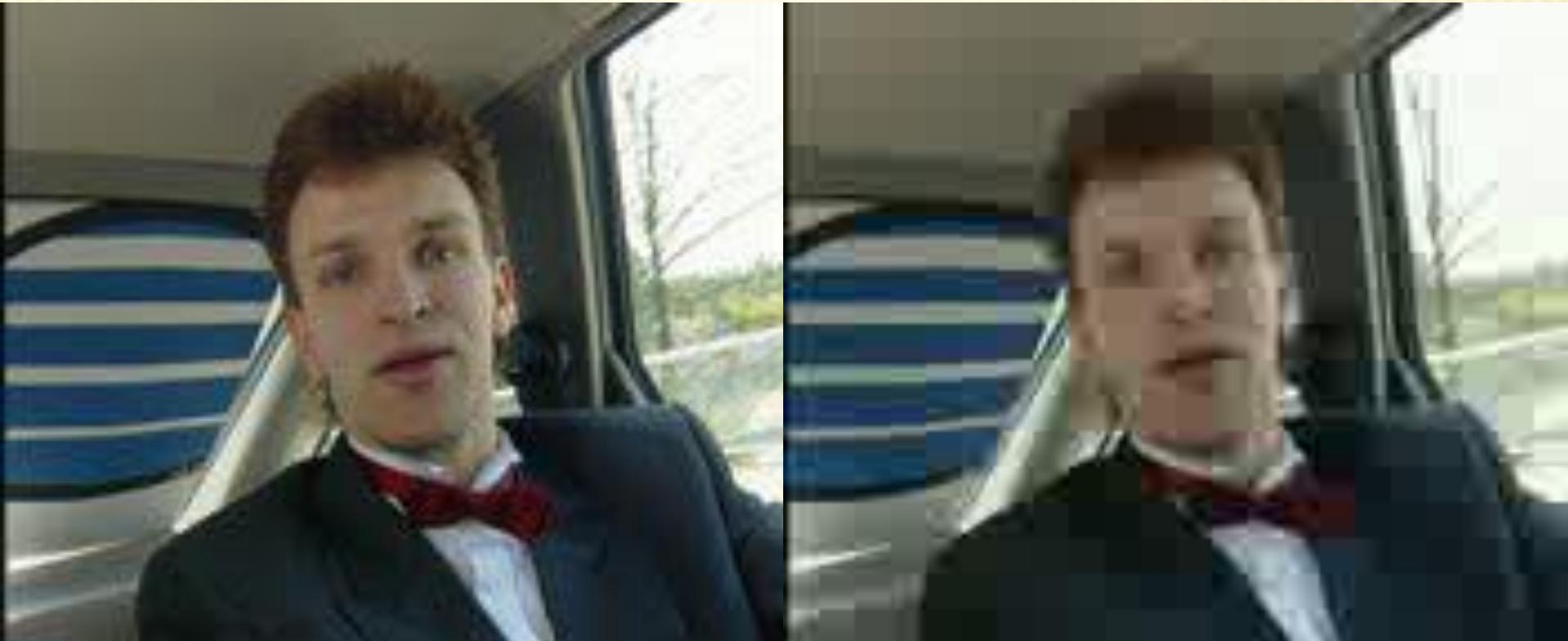
# EXEMPLES AVEC H.263



Séquence "car-phone"  
30 im/s ; 9124 kbps

10 im/s ; 105 kbps  
 $T=29$  ;  $PSNR=36$  dB

# EXEMPLES AVEC H.263



Séquence "car-phone"  
30 im/s ; 9124 kbps

10 im/s ; 22 kbps  
 $T=139$  ;  $PSNR=29.8$  dB

# COMPARAISON DE CODEURS VIDÉO

Original



H.263 baseline (33 dB)



H.263+ (33.5 dB)



MPEG-4 core (33.5 dB)



H.264 (42 dB)



Vidéo Foreman (QCIF, 30 fps, 100 kbit/s)

# MPEG-4

## ✗ Objectifs

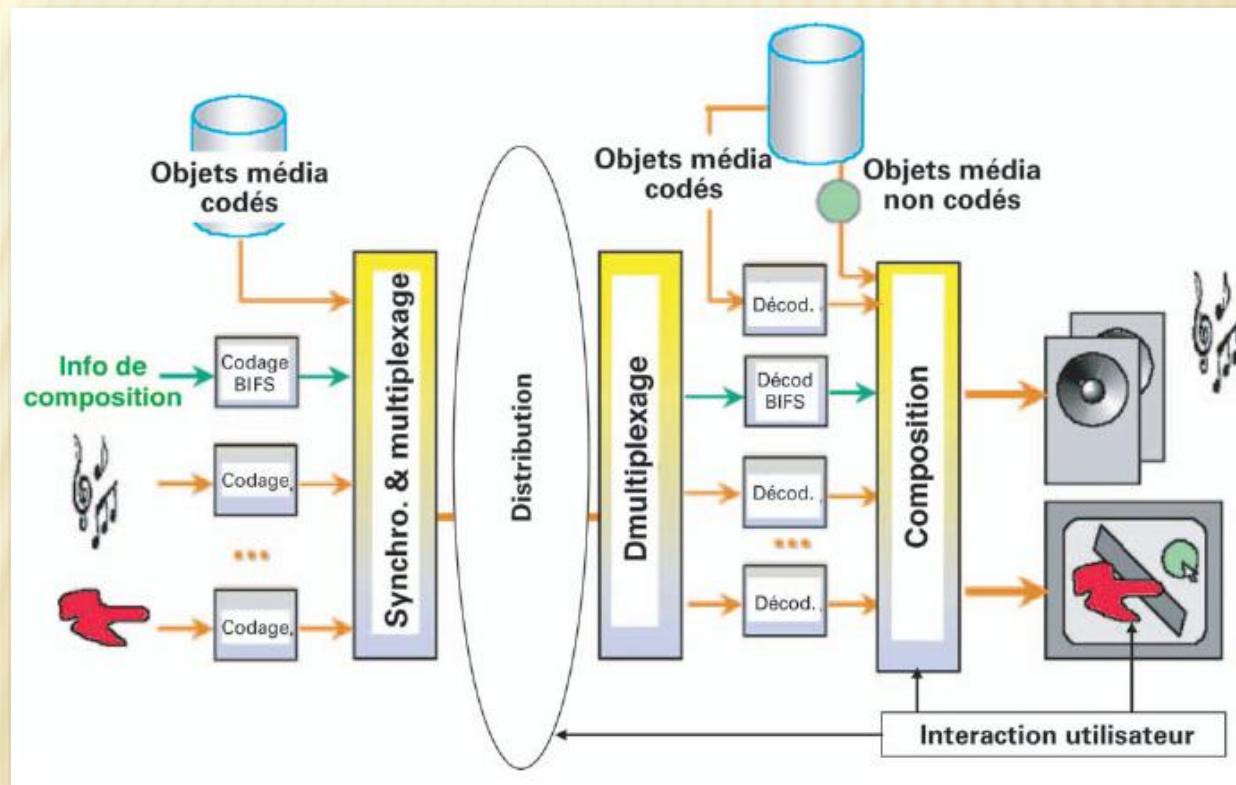
- + Amélioration du débit
  - bas débit : téléphonie mobile
  - Haut débit : production vidéo haute définition, ...
- + Indépendance vis-à vis des réseaux de transmission
- + Accès direct aux données
- + Interactivité (jeux interactifs en réseaux, ...)
- + Streaming audio et vidéo (diffusion à la demande)

## ✗ Applications

- + Diffusion TV et TVHD (TNT, cable, satellite, ADSL), programmes à la demande
- + videoconférence
- + Services vidéo vers les mobiles 3G (actualités, bourse, météo, ...)
- + Application graphiques interactives (jeux en réseau)
- + Distribution sur supports optiques (DVD, Blue ray Disc)

# MPEG-4

- ✖ Représentation des contenus par des « objets media »
- ✖ Description de l'assemblage de ces objets pour former des scènes
- ✖ Multiplexage et synchronisation des objets
- ✖ Interactivité côté terminal



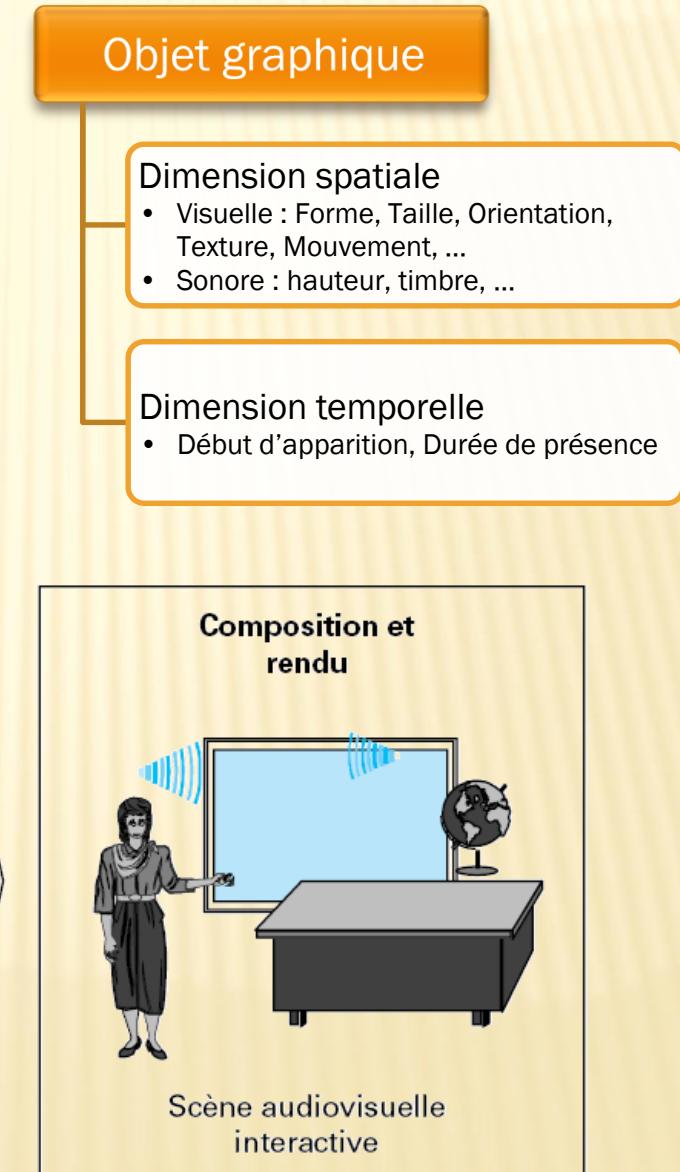
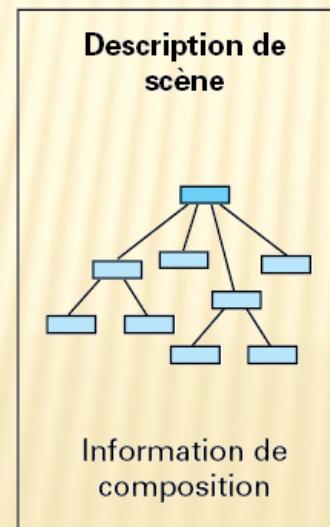
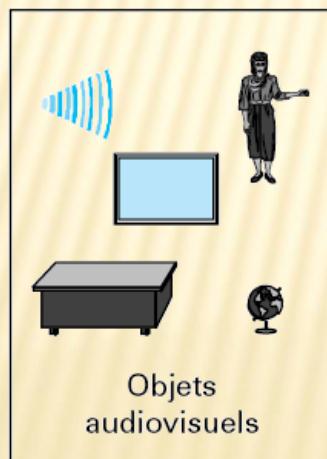
# OBJETS MEDIA

## Scène → organisation hiérarchique

- + Image fixes (fonds,...)
- + Objets vidéos animés
- + Objets audio (voix, musique,...)

## Objets primitifs normalisés

- + Textes et graphiques
- + Têtes parlantes
- + Corps animés
- + Sons de synthèses



# OBJETS MEDIA

## Composition

→ Description normalisée de la scène :  
BIFS (Binary Format for Scenes)

- + Placement des objets
- + Transformation → changement d'apparence géométrique ou acoustique
- + Groupement d'objets media primitifs en d'objets media composites
- + Modification d'attributs
- + Paramètres d'animation
- + Changement interactif de point de vue

	Codage visuel	Codage audio
Naturel	<b>Vidéo rectangulaire</b> Outils Simple Profile Outils Advances Simple. P AVC / H.264 <b>Scalabilité Vidéo</b> <b>Objets vidéo</b> Mosaïque (sprite) Forme Texture / Images fixes	<b>Codecs audio</b> AAC HE-AAC BSAC TwinVQ HLIN <b>"Scalabilité" Audio</b> <b>Codecs parole</b> CELP HVXC
De synthèse	<b>Graphismes (SNHC)</b> Graphismes 2-D Graphismes 3-D VRLM Animation visage Animation corps	<b>Audio de synthèse</b> Texte à Parole (TTS) Structured Audio <b>Rendu Audio</b>
Description des objets (OD) Description de scène (BIFS / XMT)		
Gestion et protection des droits (IPMP/ DRM)		
Format de fichier (MP4)		
Transport des données (Sync - DMIF - Multiplex)		
Interactivité (MPEG-J, JavaScript / Upstream)		

Boîte à outils MPEG-4

# PROFILS ET NIVEAUX

## ✖ Profils

spécification des outils MPEG-4 pour certaines classes d'applications

- + **Profils visuels** → types d'objets visuels pouvant être utilisés
- + **Profils audio** → types d'objets sonores
- + **Profils graphiques** → éléments graphiques autorisés
- + **Profils de graphes de scènes** → fonctions de composition autorisées (2D, 3D, translation, rotation, ...)
- + **Profils de descripteurs d'objets** → capacité de la couche de synchronisation et des outils du descripteur d'objets

## ✖ Niveaux

→ contraintes de débit, échantillonnage, mémoire, etc. pour chaque profil

# CODEUR VIDEO MPEG-4

## ★ 3 types de VOP ( Video Object Plane)

- + I-VOP : codage INTRA, sans compensation de mouvement
- + P-VOP : prédiction à partir des derniers I-VOP ou P-VOP
- + B-VOP : prédiction bidirectionnelle à partir des I-VOP ou P-VOP antérieurs ou postérieurs

## ★ Compensation de mouvement

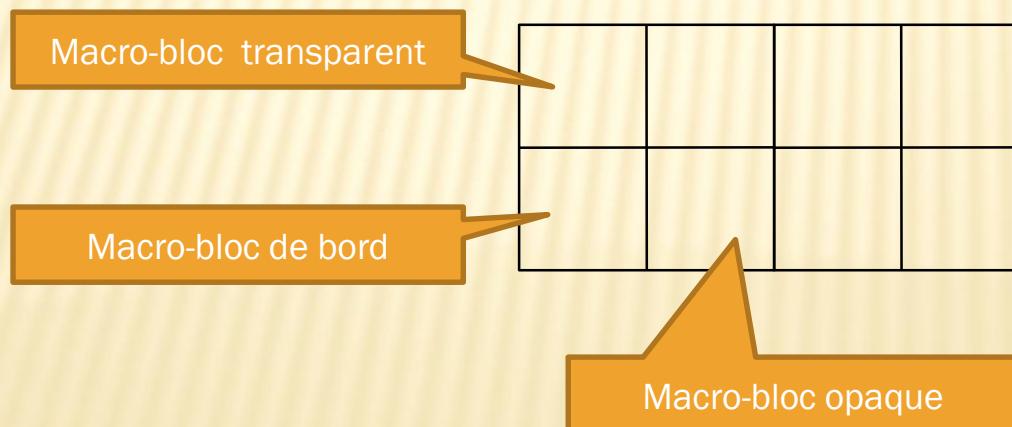
- + Locale (précision  $\frac{1}{2}$  et  $\frac{1}{4}$  de pixel)
  - ✗ 1 MV : 1 vecteur mouvement par macro-bloc
  - ✗ 4 MV : 4 vecteurs mouvement par macro-bloc

codage prédictif des vecteurs mouvement à partir de ceux des blocs adjacents. Codage de l'erreur de prédiction des vecteurs par VLC

- + Globale (GMC : Global Motion Compensation)  
les vecteurs dont le déplacement est cohérent avec le mouvement global ne sont pas spécifiés

# CODAGE DES OBJETS NON RECTANGULAIRES

- ✖ Codage de la forme par un masque (alpha mask)
  - + binaire → support de l'objet
  - + en niveaux de gris → effets de montage avec transparence
- ✖ Codage du contenu
  - + Macro-blocs opaques
  - + transparents
  - + de bord → remplissage des pixels transparents par extrapolation pour ne pas induire de coefficients DCT élevés



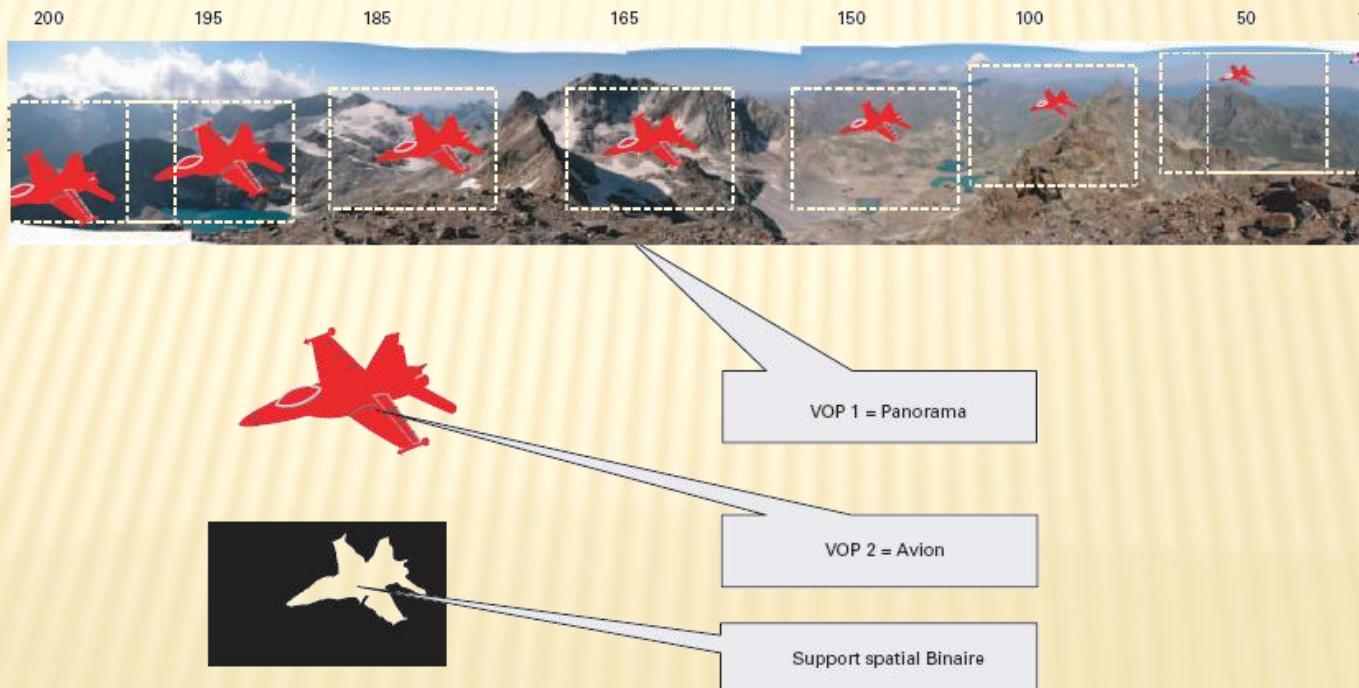
# CODAGE ADAPTATIF

## ✖ Couches VOL (Video Object Layers)

- + Couche de base : image en qualité réduite
- + Couche d'amélioration : détails
- scalabilité temporelle : sous-échantillonnage temporel
- scalabilité spatiale: sous-échantillonnage spatial
- scalabilité qualitative (SNR): codage des coefficients DCT par plans de bits en commençant par les poids forts

# CODAGE PAR MOSAIQUE

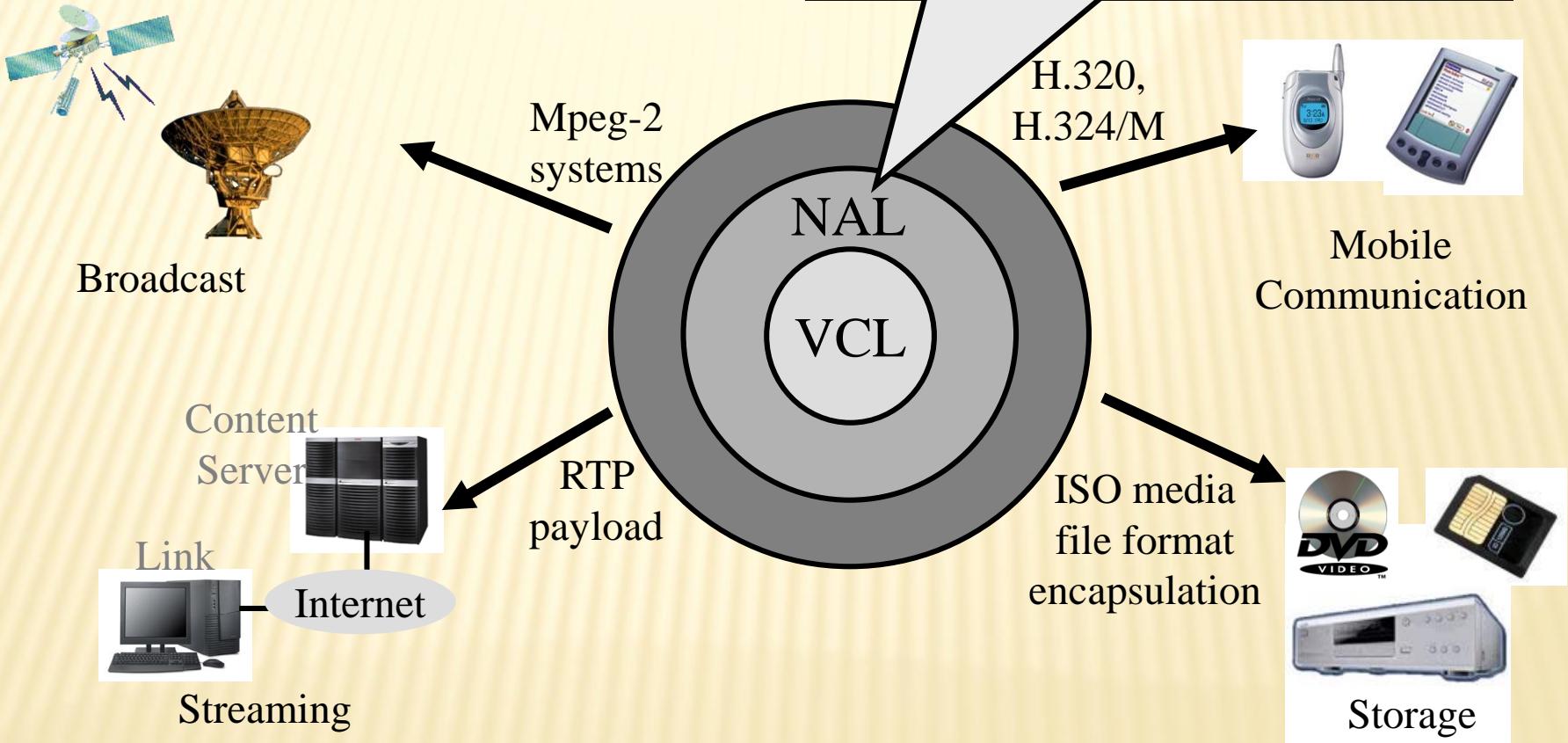
- ✖ **Image d'arrière plan**  
→ transmise une fois en début de séquence
  - ✖ **Objet , support binaire et paramètres de taille position, orientation**
  - ✖ **Paramètres de mouvement**
- La scène est recréée au décodage



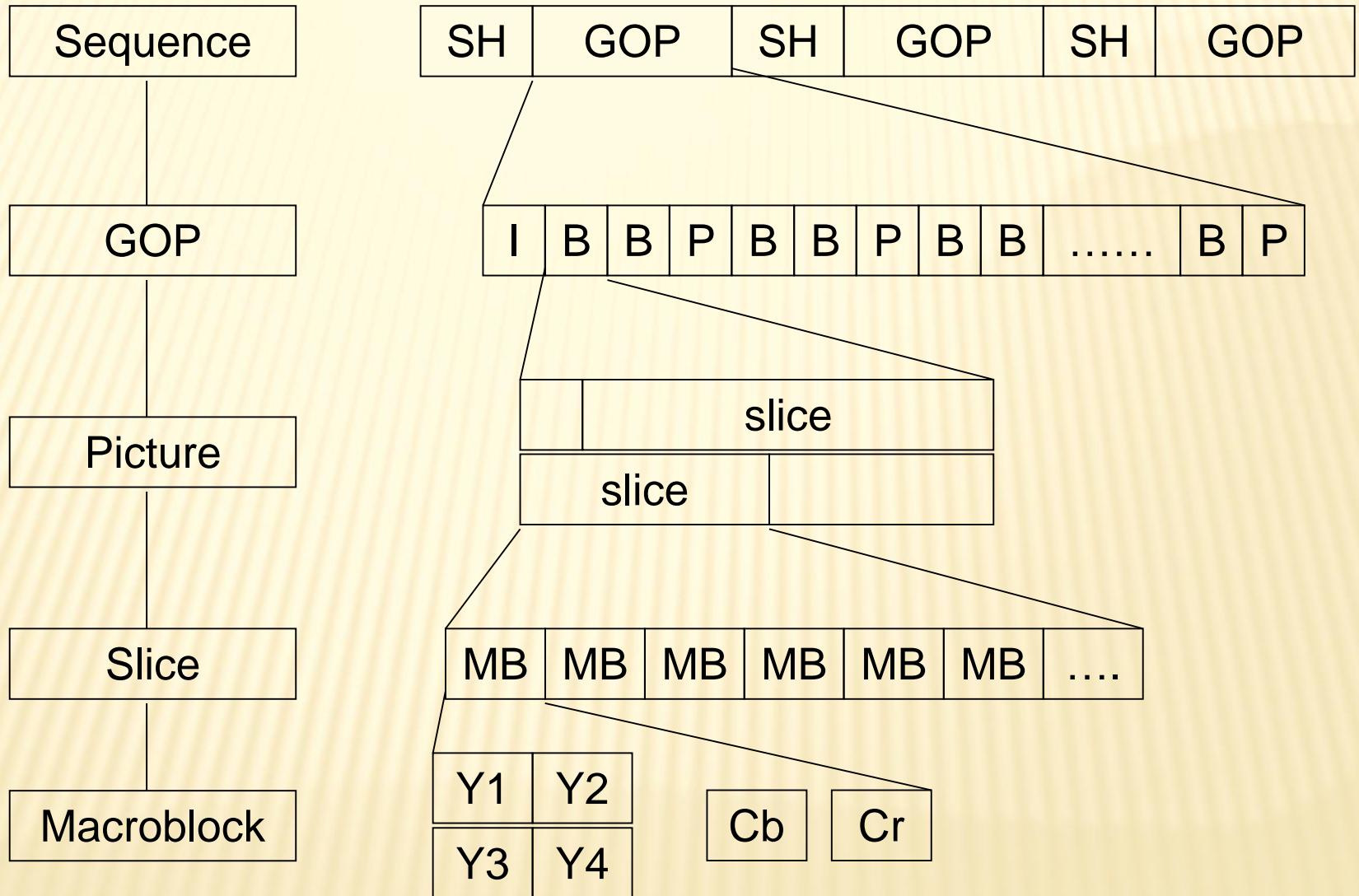
# Applications

## DMB Multimedia Service

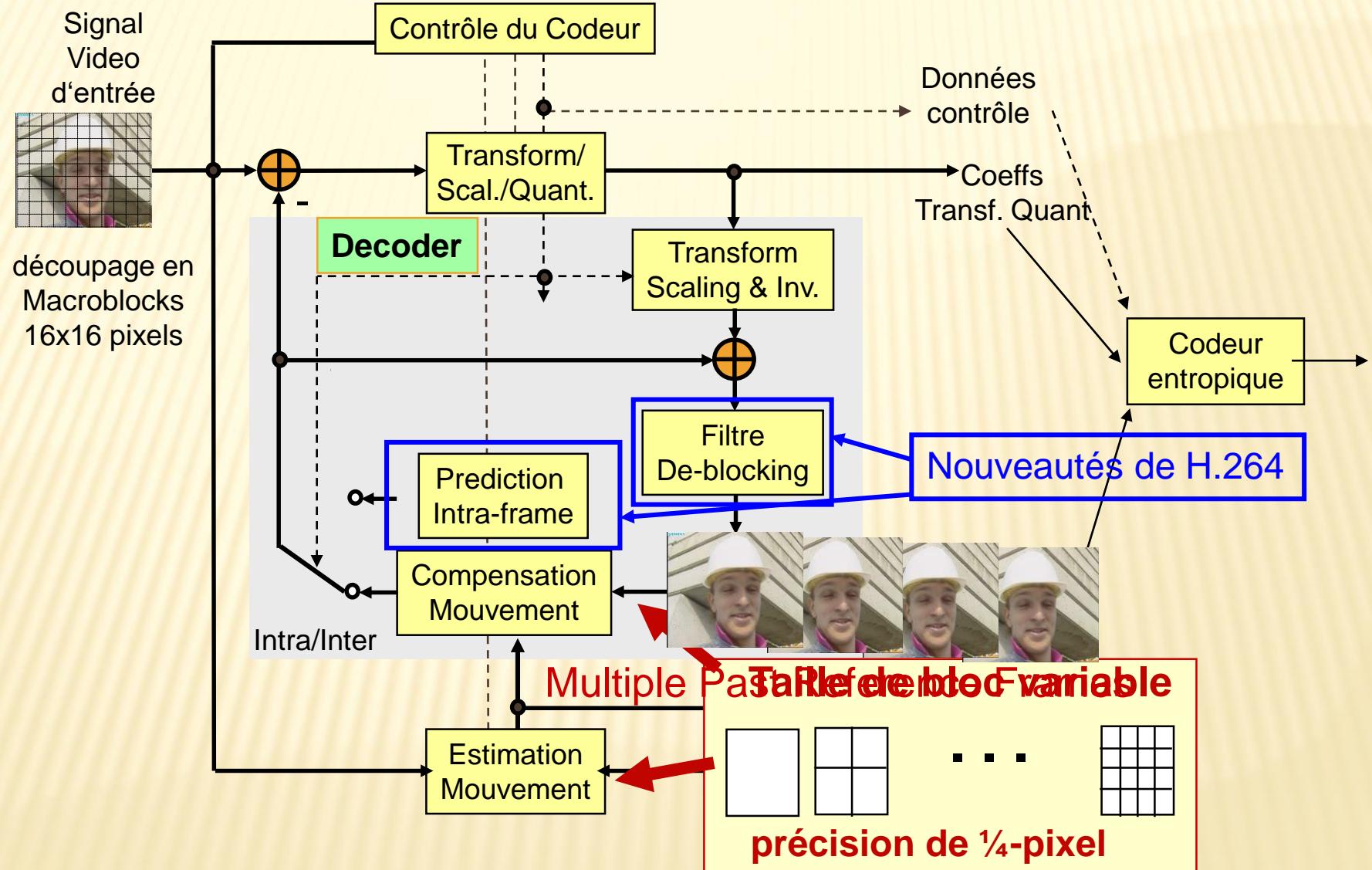
NAL gives VCL network independent interface



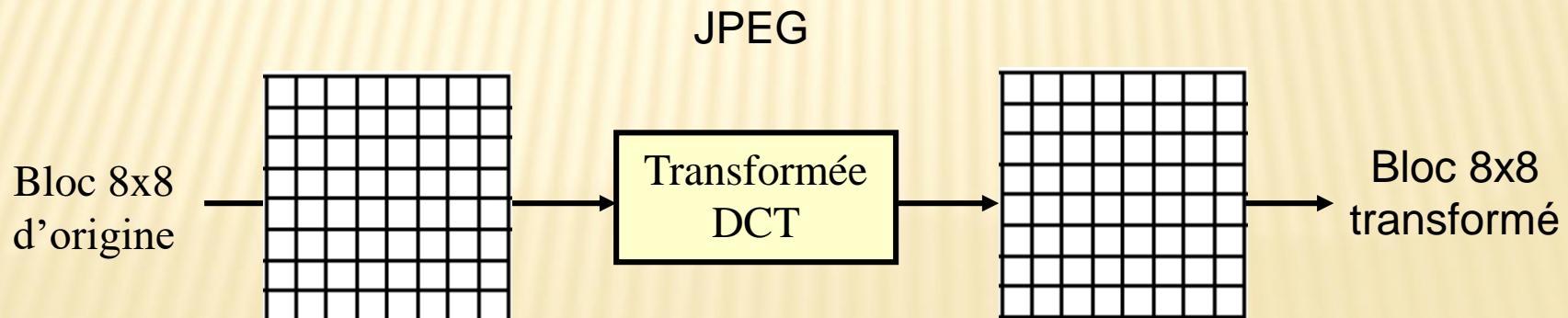
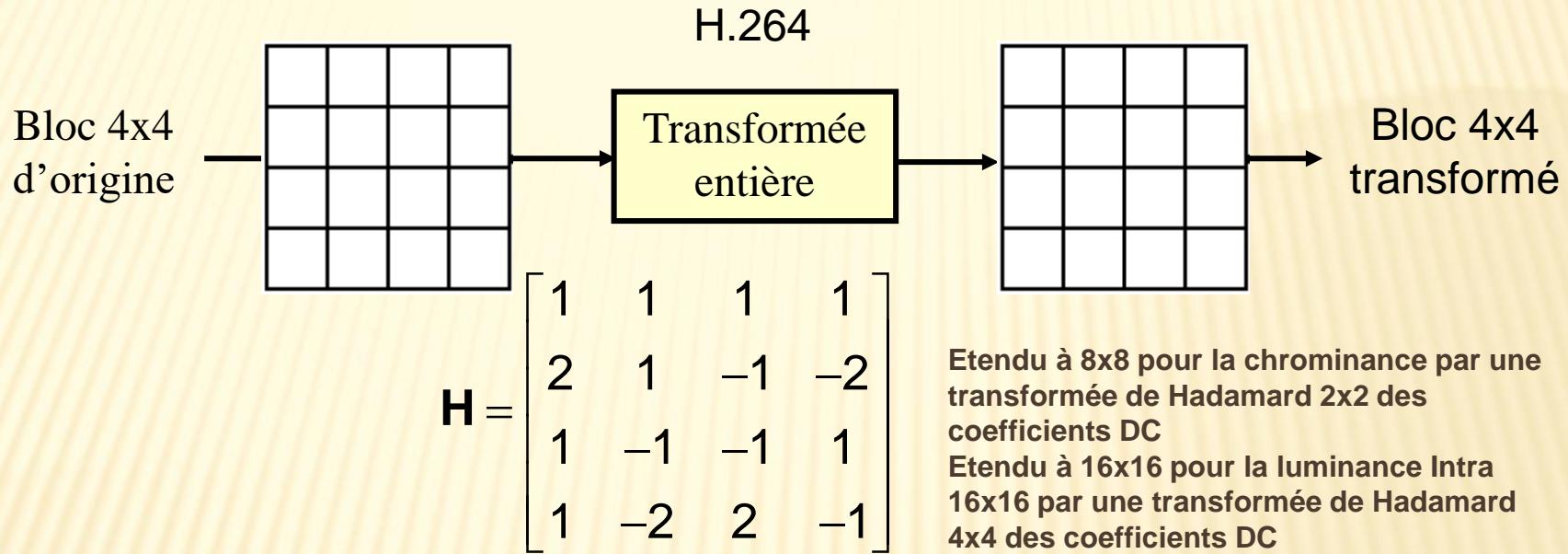
# Structure de données



# Principe de codage d'un *macroblock*



# Transformation



# Filtre de suppression d'effet de bloc

- ✖ “Deblocking Filter”
  - + Artefacts d'effet de bloc importants
    - ✖ transformation 4\*4 transforms et utilisation de compensation de mouvement
  - + Gain de l'ordre de 6~9% en débit
  - + Améliore la qualité subjective et le rapport signal sur bruit de l'image décodée



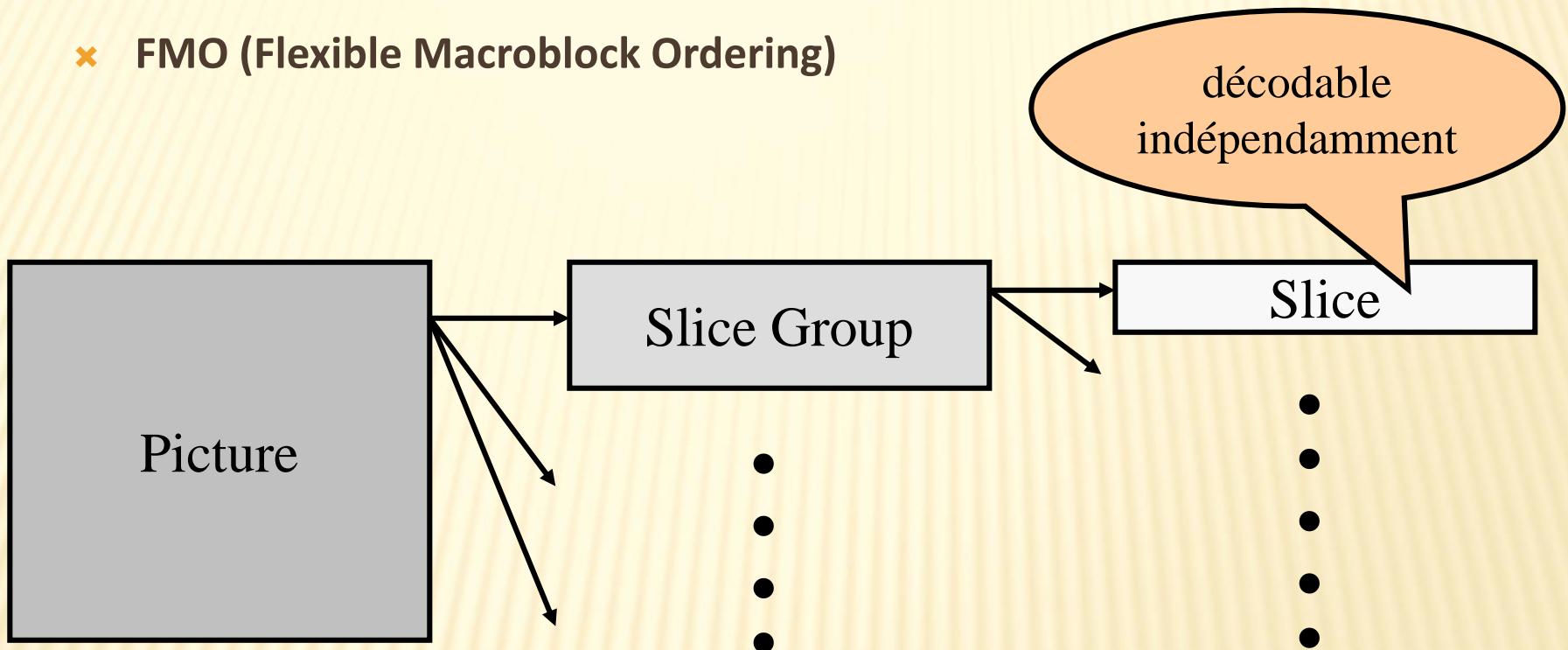
Sans filtrage



Avec “Deblocking Filter”

# Utilisation de « slice »

## ✗ FMO (Flexible Macroblock Ordering)



### Utilisation de Slice (FMO)

→ améliore la robustesse à la perte de données

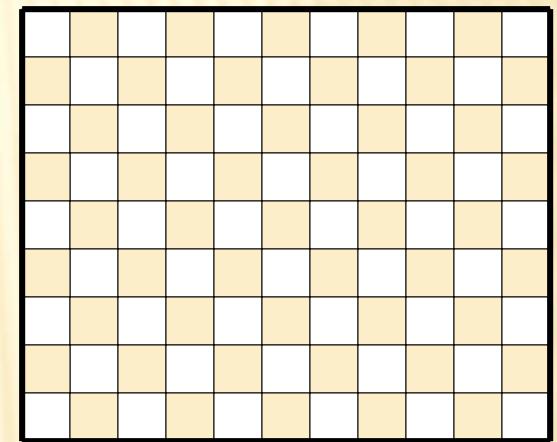
# Utilisation de « slice »



Découpage d'une image  
en “slices”  
sans utilisation de FMO



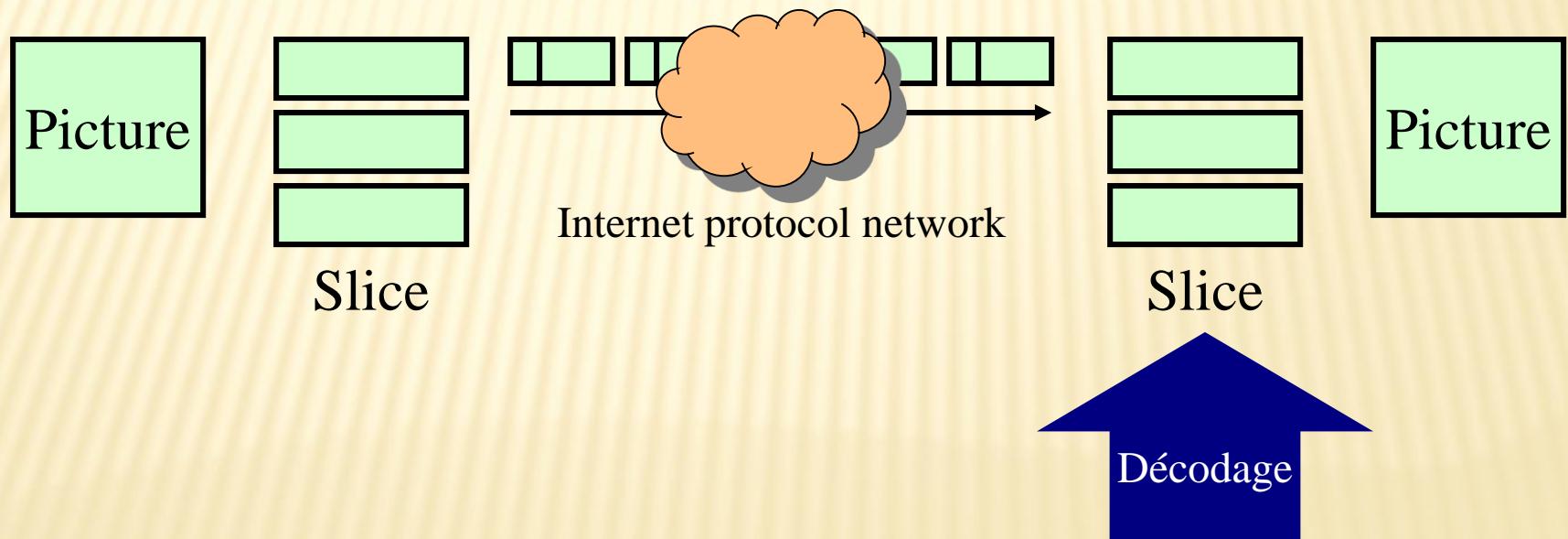
Découpage d'une image en “slices”  
avec utilisation de FMO



- Slice Group #0
- Slice Group #1

# Utilisation de « slice »

- ✖ ASO (Arbitrary Slice Ordering)
  - + décodable indépendamment (améliore la résistance aux pertes)
    - ✖ permet l'envoi et la réception d'une **slice dans un ordre quelconque**
    - ✖ Améliore **le délai de bout-en-bout** dans des applications temps-réel

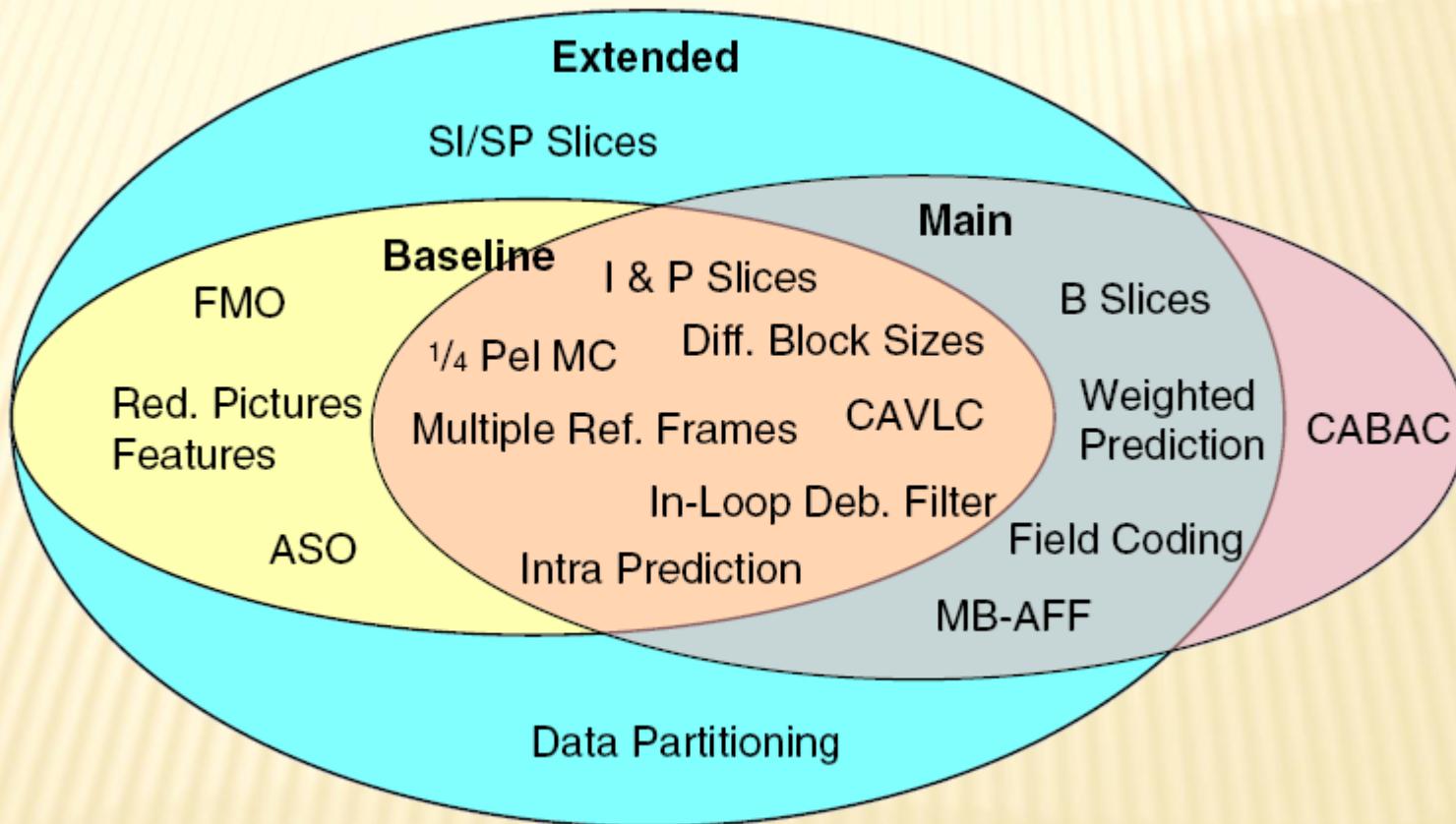


# Codeurs arithmétiques

---

- ✖ CAVLC (Context Adaptive Variable Length Coding)
  - + Codeur entropique dynamique pour le codage des résidus (profil de base)
- ✖ CABAC (Context Adaptive Binary Arithmetic Codes)
  - + Codeur entropique binaire à contexte variable : gain de 5 à 15% par rapport à CAVLC

# Profiles

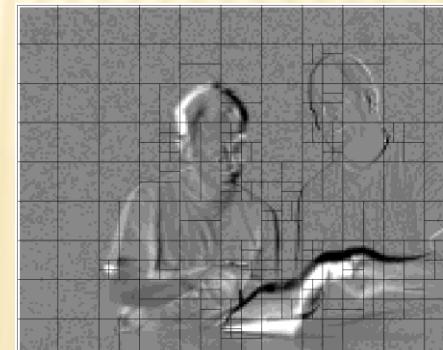
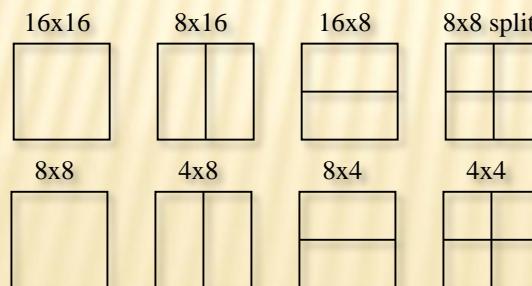


# MPEG4-AVC (MPEG4-PART10, H264, H264/AVC)

- ✖ Normalisé en 2003
- ✖ Gain en qualité et en débit
  - + Filtre anti-bloc dans la boucle de codage
    - ✖ Améliore la qualité subjective et le rapport signal sur bruit de l'image décodée
    - ✖ gain en débit de 6 à 9%

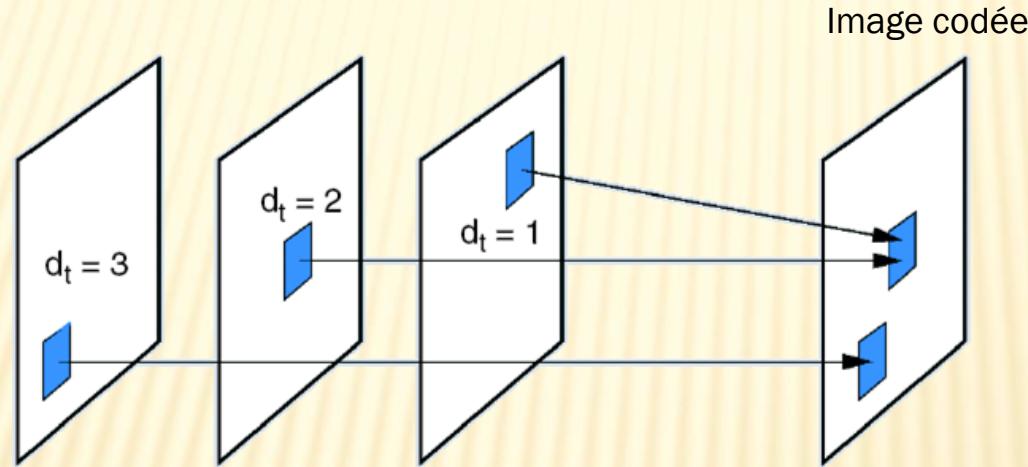


- + Taille de blocs variable pour la compensation de mouvement



# MPEG4-AVC (MPEG4-PART10, H264, H264/AVC)

- + Prédiction spatiale directionnelle pour le codage INTRA
- + Utilisation d'images référence multiples
  - ✗ Poids variables
  - ✗ Pas de direction temporelle
  - ✗ Peut utiliser une "slice" de référence



# MPEG4-AVC (MPEG4-PART10, H264, H264/AVC)

- ✖ Complexité accrue !
- ✖ Utilisation du processeur :

