



*La numérisation
De La Haute Définition*

Du MPEG-2 au MPEG-7

Et le D-CINEMA JPEG 2000

De la Production analogique à la production Numérique

<u>Support</u>	<u>Format</u>	<u>mm</u> <u>Hauteur</u>	<u>mm</u> <u>Largeur</u>	<u>mm²</u> <u>Superficie</u>
70 mm	2:28	23,01	52,48	1207,56
Super 35 mm	2:37	18,67	24,92	465,26
35 mm	1:66	16	21,95	351,20
	1:85	16	21,95	351,20
Super 16mm	1,66	07,42	12,35	91,64
16mm	1:37	07,49	10,26	76,85
Super 8mm	1:34	04,14	5,79	23,97

La Définition D'une Image Analogique.

- *Sensibilité de la pellicule.*
- *L'optique: la focale utilisée.*
- *Le niveau de luminosité: le diaphragme.*

Transfert Film vers le support HD

- ❑ *Majorité des tournages sont en 24 images/sec*
- ❑ *En Europe le Standard HD = 25 images/sec*
- ❑ *4/5 du monde fonctionne en 30 images/sec*

Les producteurs plaident pour une uniformisation :

Le 24 images/sec en Progressif 24 P

La Numérisation HD

- ▣ *1 . Le signal est numérisé par un convertisseur analogique / numérique :*

Échantillonnage = découpage temporel du signal

Quantification = convertir en valeur numérique binaire

Aux normes ISO (INTERNATIONAL STANDART ORGANISATION)

1. La Fréquence d'échantillonnage

- *Un signal est numérisé par un convertisseur analogique / numérique.*
- *Il est découpé à un rythme régulier ou période. On définit ainsi sa fréquence d'échantillonnage.*
- *La loi de Nyquist détermine que la fréquence d'échantillonnage doit être au moins le double de la plus haute fréquence reproductible.*
- *La fréquence d'échantillonnage est au minimum de 13,5 Mhz, c'est-à-dire au moins deux fois supérieure à la fréquence maximale du signal qui est de 5,5 Mhz.*
- *A chaque période le signal est quantifié et converti en valeurs binaires 0 ou 1.*

2. La Quantification

- *Le signal est analysé en intensité électrique au bout d'une durée qui est fonction de la fréquence d'échantillonnage. Chaque intensité est quantifiée et mémorisée sous forme de mot binaire. A ce niveau arrive la notion de Résolution.*
- *En TVHD une quantification en **8 Bit** peut être suffisante.*

<i>SONY</i>	<i>HDCAM SR 10 Bit</i>
	<i>HDCAM 8 Bit</i>
<i>PANASONIC</i>	<i>D5 –HD 8 à 10 Bit</i>
	<i>VARICAM HD100 8 Bit</i>
- *En D-Cinéma on fait appel à **10 ou 16 Bit** au moins.*

<i>THOMSON</i>	<i>LA VIPER 10 Bit</i>
<i>DALSA</i>	<i>L'Origin 14 bit</i>
<i>ARRI</i>	<i>D20 12 Bit</i>
<i>PANAVISION</i>	<i>GENESIS 12 Bit</i>

La quantification : transformer un signal en codage binaire

La quantification sur 8 bits : de 0 à 256 valeurs

Un Bits peut connaître deux états: 0 ou 1. Les Bits en informatique sont généralement placés en série de 8 BITS appelés BYTES ou OCTET. Avec 8 BITS dans un BYTE ou OCTET, on peut différencier 256 valeurs numériques différentes comprises entre 0 et 255. (2^8)

0 = 00000000
254 = 11111110

1 = 00000001
255 = 11111111

2 = 00000010

0	0	0	0	0	0	0	0	= 0
0	0	0	0	0	0	0	1	= 1
0	0	0	0	0	0	1	0	= 2
1	1	1	1	1	1	1	0	= 254
1	1	1	1	1	1	1	1	= 255

La quantification sur 9,10,11,12,14 bits :

A chaque fois que l'on ajoute un bit supplémentaire, le nombre de combinaison est doublée:

9 Bits possède	512 états
10 Bits possède	1.024 états
11 Bits possède	2.048 états
12 Bits possède	4.096 états
13 Bits possède	8.192 états
14 Bits possède	16.384 états
15 Bits possède	32.768 états
16 Bits possède	65.536 états

- Le canal numérique à une bande passante énorme mais demande un stockage proportionnel.
- Jusqu'à ce jour on a fait appel à la compression. Les nouvelles caméras ouvrent la perspective de la non compression sur toute la chaîne du tournage à la diffusion en salle.

Le système couleur en HD non compressé : RGB – RVB

Soit en codage sur 16 bits (en 2006):

RGB - RVB

Red - Rouge

1	0	1	0	0	0	0	1
1	1	1	1				

= 65.536 niveaux

Green - Vert

1	0	1	0	0	0	0	1
1	1	1	1				

= 65.536 niveaux

Blue - Bleu

1	0	1	0	0	0	0	1
1	1	1	1				

= 65.536 niveaux

Le double canal HD - SDI permet sur une résolution de 16 Bits d'obtenir une bande passante de couleur RGB intégrale.

** La caméra Genesis Panavision, l'ARRI D 20,*

Le système couleur en télévision :

$$Y = 0,3 R + 0,59 V + 0,11 B$$

System RGB



Composant Couleur



$$(V = \frac{Y - 0,3 R - 0,11 B}{0,59})$$

Les différents modes d'échantillonnage en Haute Définition en Y Cb Cr

- En Cinéma Haute Définition l'échantillonnage idéal doit s'effectuer en 4 : 4 : 4 **non compressé**. (16 bits vers 2007)

* la Viper, la HDCAM S R.

4 : 4 : 4	Pixel 1	Pixel 2	Pixel 3	Pixel 4
Y	16 bits	16 bits	16 bits	16 bits
Cr	16 bits	16 bits	16 bits	16 bits
Cb	16 bits	16 bits	16 bits	16 bits

- 4 : 2 : 2

Pixel 1	Pixel 2	Pixel 3	Pixel 4
Y	16 bits	16 bits	16 bits
Cr	16 bits	Partagé	16 bits
Cb	16 bits	Partagé	16 bits

Compression
Réduction Couleur
de 33 %

- La Caméra HDW-F750 Sony, la Varicam Panasonic.
Bande passante 27Mhz et un débit de 216 Mb/sec en 8 bits

Les formats de Compression

1. La DCT (Discrete Cosinus Transform)

La DCT est utilisée pour le Jpeg en image fixe et pour le Mpeg-1, Mpeg-2, ou le Mpeg-4.

- 1. Division de l'image en Macro Blocs par bloc de 8 x 8 pixels.*
- 2. La DCT opère une transformation du domaine spatial de pixels en domaine d'espace de fréquence. Elle utilise une fonction mathématique basée sur la transformée de Fourier.*

On aboutit à une matrice dont les différentes valeurs vont directement représenter la quantité de détails dans la portion d'image concernée.

Les formats de Compression

1. La DCT (Discrete Cosinus Transform)

135	136	152	132	130	147	155	138
138	137	151	134	132	145	139	141
142	143	132	135	153	154	138	140
153	144	132	132	142	132	132	132
144	151	156	140	136	141	144	140
150	152	141	132	132	152	132	132
136	132	142	138	122	124	132	132
148	132	148	142	132	128	124	132

DCT



172	-18	15	-8	23	-9	-14	19
38	-34	25	-9	-10	11	14	12
21	9	-6	3	-5	8	5	-1
-10	6	-5	4	3	-4	-2	5
-8	3	4	5	-3	3	-7	1
4	-2	-2	-4	6	6	2	-4
4	-3	-4	5	6	5	3	-2
0	-8	-4	3	2	1	4	0

La quantification est non conservatrice et contrôle le débit. Elle applique des coefficients afin de réduire les écarts de niveaux dans les hautes fréquences où l'œil est moins sensible. Cette réorganisation s'effectue par ordre croissant, les détails les plus fins étant situés en bas et à droite de la matrice, la première valeur en haut à gauche, représentant la valeur moyenne de la matrice.

Les formats de Compression

2. La DWT (Discrete Wavelet Transform).

- ▣ *Les Ondelettes utilisées pour le JPEG2000 et le Digital Cinema*
 - *Elle utilise une autre logique mathématique, la DWT (Discrete Wavelet Transform). L'image d'origine est décomposée en blocs de taille identique.*
 - *Décomposition fréquentielle établissant une séparation entre les fréquences les plus basses, les plus hautes, les fonds continus, les détails fins, et les contours qui sont contenues dans l'image.*
 - *La DWT exploite la redondance spatiale et fréquentielle.*

JPEG 2000

Image Size:

4096 x 3112 pixels (fin 2005)

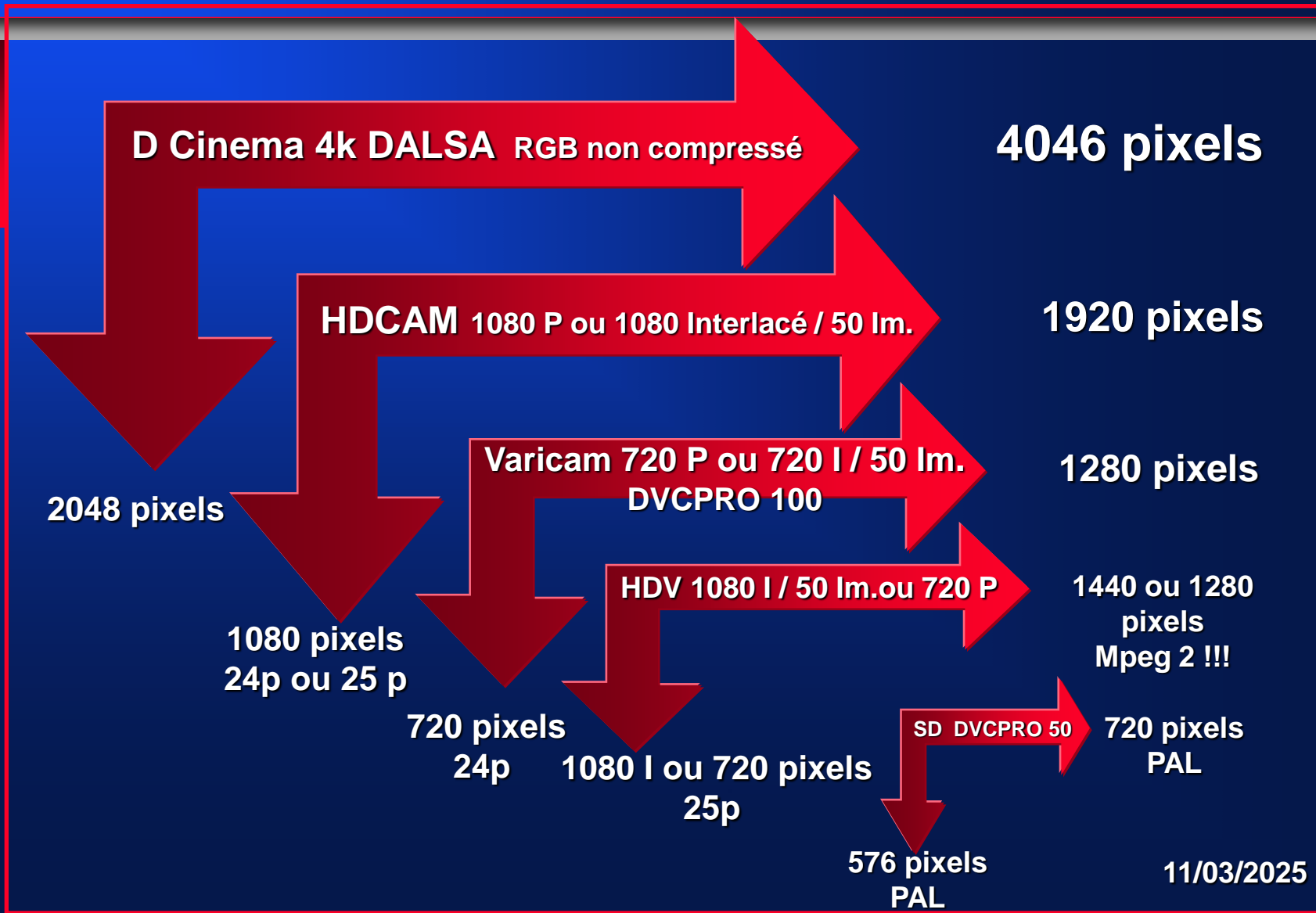
Image quality:

RGB 4:4:4 / 16 Bit

Avantage d'une compression « ondelettes »

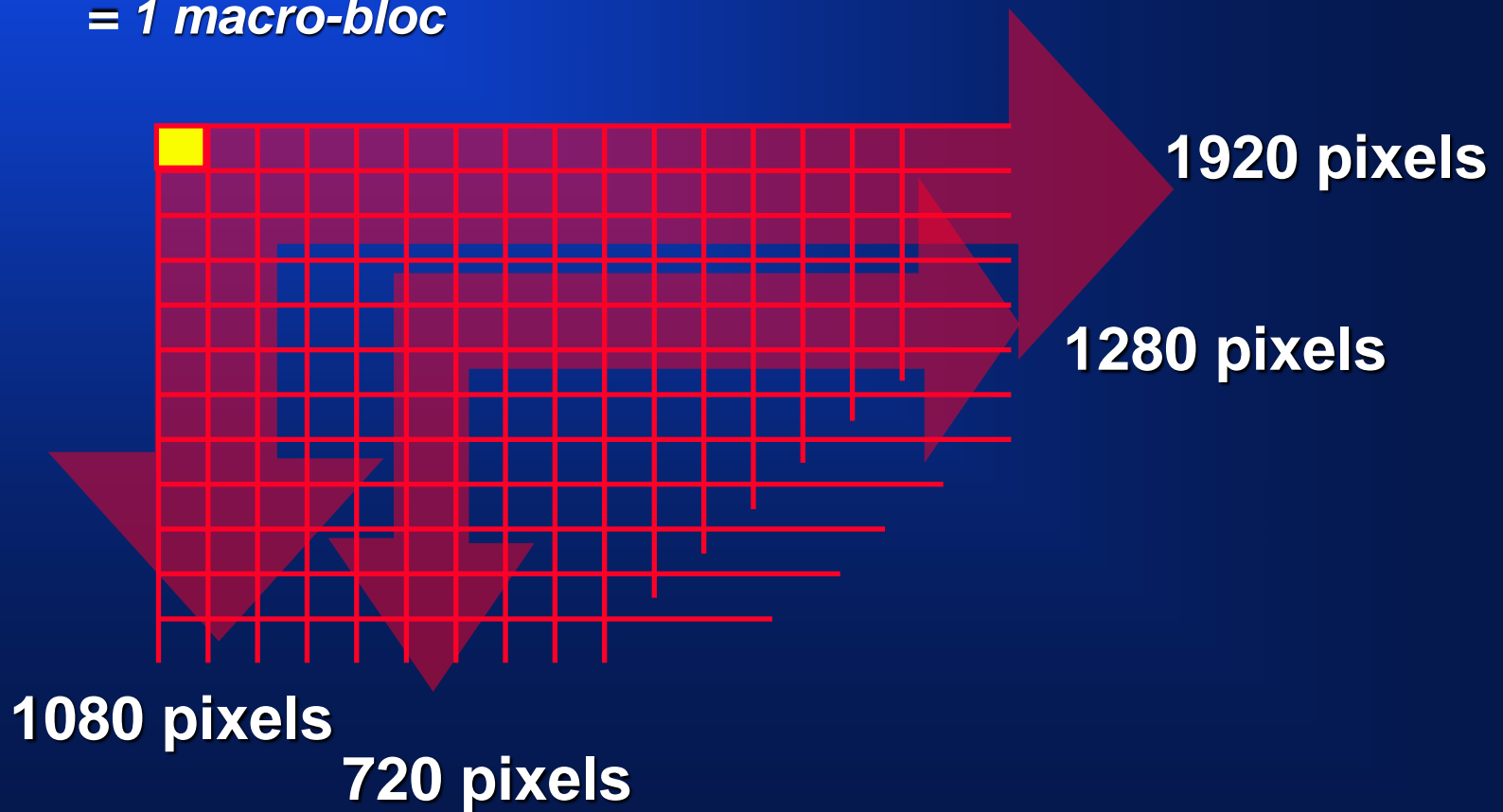
- *A un taux de compression fixe, la qualité image est supérieure aux autres codages*
- *A qualité égale, le résultat présente un débit inférieur de 50% par rapport à la DCT.*
- *la durée de la phase de compression et de décompression est identique*
- *la (dé)compression progressive offre à l'utilisateur un choix adaptés aux différentes bandes passantes d'un réseau. Exemple: le Streaming pour le VDSL.*
- *le codage est plus robuste que les autres aux erreurs de transmission*

Résolution d'une image HD en format 16/9 :



1080 P : 32.400 Macro Blocs
720 P : 14.400 Macro Blocs

8x8 pixels
= 1 macro-bloc

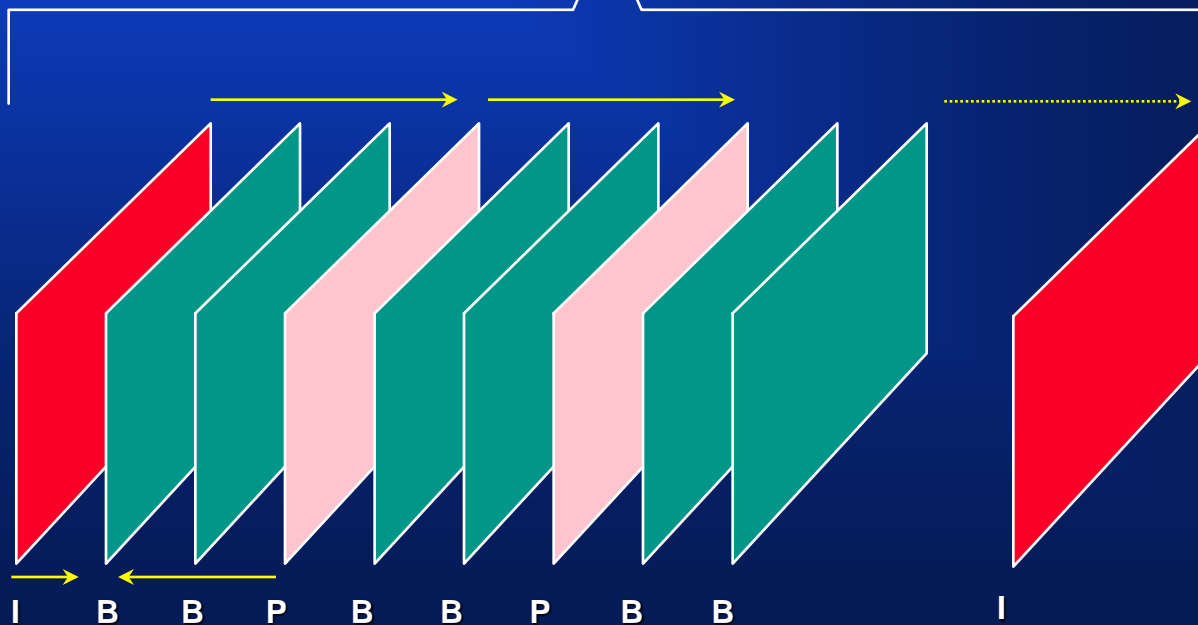


La Numérisation HD

Le Codage des sources

- *Représenter de manière compacte les grandeurs quantifiées (compression numérique)*
- *En 1990 le Betacam Numérique l'histoire de la compression numérique commençait !*
- *Objectif de la compression : réduire le débit en comprimant fortement le poids des images .*
- *Exploitation de :*
 - *Les redondances spatiales: intra image.*
 - *Les redondances temporelles: inter images*
 - *Les redondances subjectives: vision détail*
 - *Les redondances statistiques: codages entropiques.*
- *But: réduire le débit et conserver autant que possible la qualité des images originales.*

Le système de codage MPEG-2



**Group of pictures
GOP :**
12 - 16

■ I Intracoded frame

■ P Forward/predicted frame

■ B Bi-directional/interpolated frame

MPEG 2 choisit par la France pour la T.N.T

MPEG 2 pour le HDTV:

38Mbts par voie hertzienne = 5 x la Télévision Pal

Fréquence échantillonnage Vidéo de 27 Mhz limité à 8 Bits

SMPTE 274 M 1080 x 1920 ou 720 x 1280 p ou I (50 im)

Audio: dolby AC.3: 48 Khz à 16 bit/s= 384 Kbps

Le MPEG 4 choisit par la France pour les autres types de diffusion: Satellite, etc.

- ❑ *Le MPEG-4 est né pour permettre l'intégration du multimédia, la sémantique et l'interaction.*
- ❑ *En plus de son développement pour le Digital Cinema le MPEG-4 propose une réponse à des besoins pour des champs d'applications variés comme des services audiovisuels interactifs à la télésurveillance en bas débit, mais aussi pour répondre au souhait de l'industrie audiovisuelle de disposer de chaînes de production virtuelle et cela avec des débits importants plus de 300 Mb/sec.*
- ❑ *Le MPEG 4 couvre un grand champ d'action allant de l'application Internet très haut débit avec WINDOWS MEDIA PLAYER 10 de MICROSOFT aux caméras comme la Varicam de chez Panasonic ou la XDCam SR de chez Sony mais aussi dans le segment DVCAM HD.*
- ❑ *Elle est, avec le MPEG 7 et MPEG 21, la clé de voûte de la convergence entre l'informatique, le cinéma, la télévision et les télécommunications.*

LES FORMATS HD EN FONCTION DE LA TAILLE DE L'ECRAN DE DIFFUSION

<i>DALSA 4K</i>	<i>4:4:4</i>	<i>4046 x 2048</i>	<i>1:1</i>	<i>14.000 Ko/i</i>	<i>3 x 1.208Mb/s</i>
<i>ARRI D20</i>	<i>4:4:4</i>	<i>3018 x 2200</i>	<i>1:1</i>	<i>8.800 Ko/i</i>	<i>1.575 Mb/s</i>

Destination: Projection en salle 4K

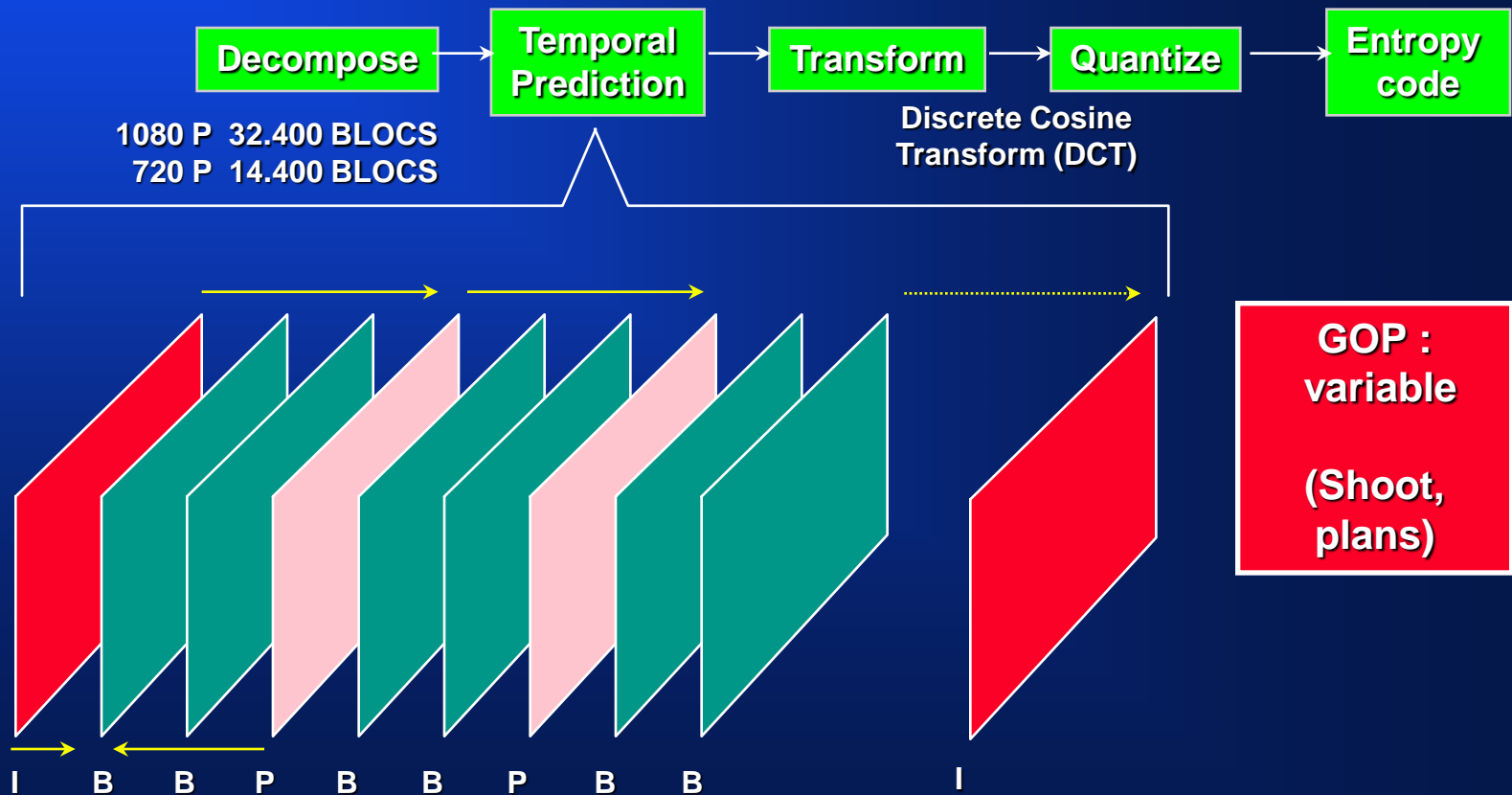
<i>HDCAM SR</i>	<i>4:4:4</i>	<i>1080 x 1920</i>	<i>2,7:1</i>	<i>4.400 Ko/i</i>	<i>880 Mb/s</i>
<i>VIPER</i>	<i>4:4:4</i>	<i>1080 x 1920</i>	<i>1:1</i>	<i>8.800 Ko/i</i>	<i>1.500 Mb/s</i>

***Destination: Projection en salle 2K
(Kinopolis Bruxelles Salle13)***

<i>HDCAM</i>	<i>4:2:2</i>	<i>1080 x 1920</i>	<i>4,3:1</i>	<i>700 Ko/i</i>	<i>140 Mb/s</i>
<i>VARICAM</i>	<i>4:2:2</i>	<i>720 x 1280</i>	<i>6,7:1</i>	<i>500 Ko/i</i>	<i>100 Mb/s</i>
<i>HDV</i>	<i>4:2:0</i>	<i>1080 x 1920</i>	<i>60:1</i>	<i>125 Ko/i</i>	<i>25 Mb/s</i>

Destination: Ecran Plasma Home Cinéma

Le système de codage MPEG-4



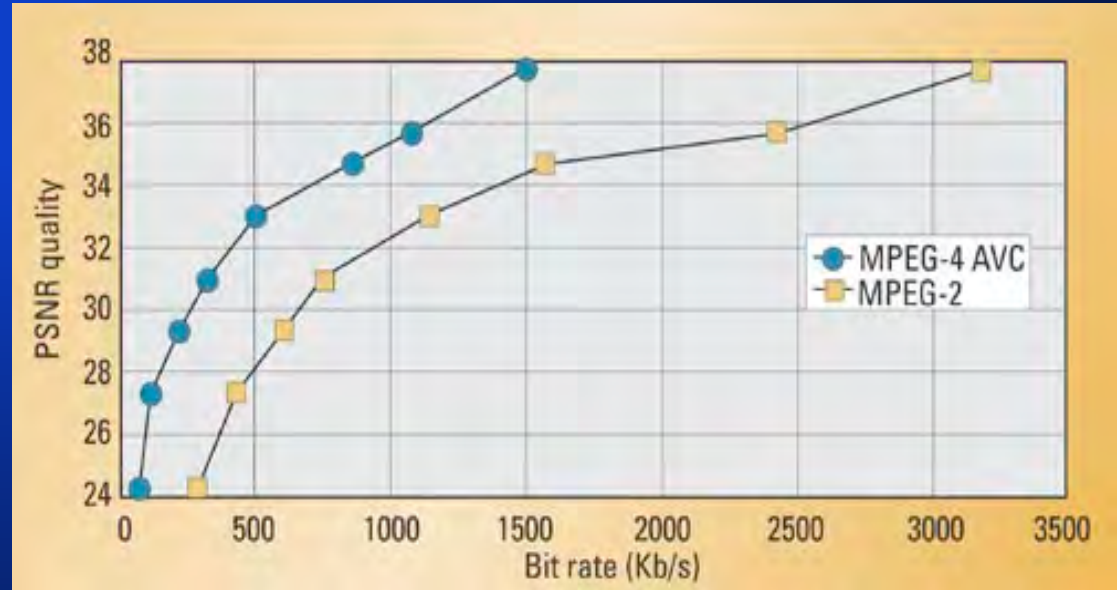
I Intracoded frame

P Forward/predicted frame

B Bi-directional/interpolated frame

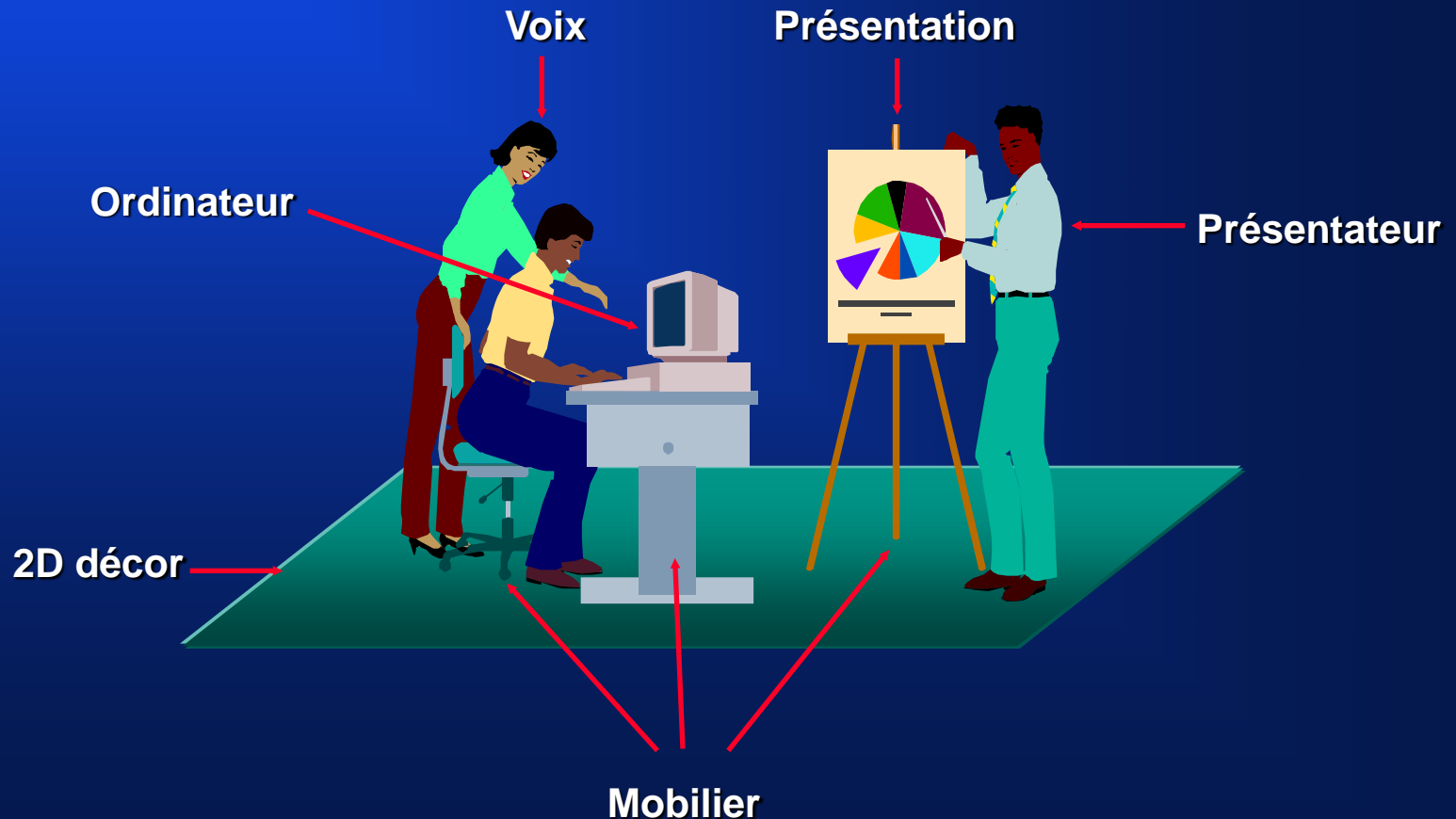
MPEG 4 : qualitativement très performante

Les flux MPEG 4

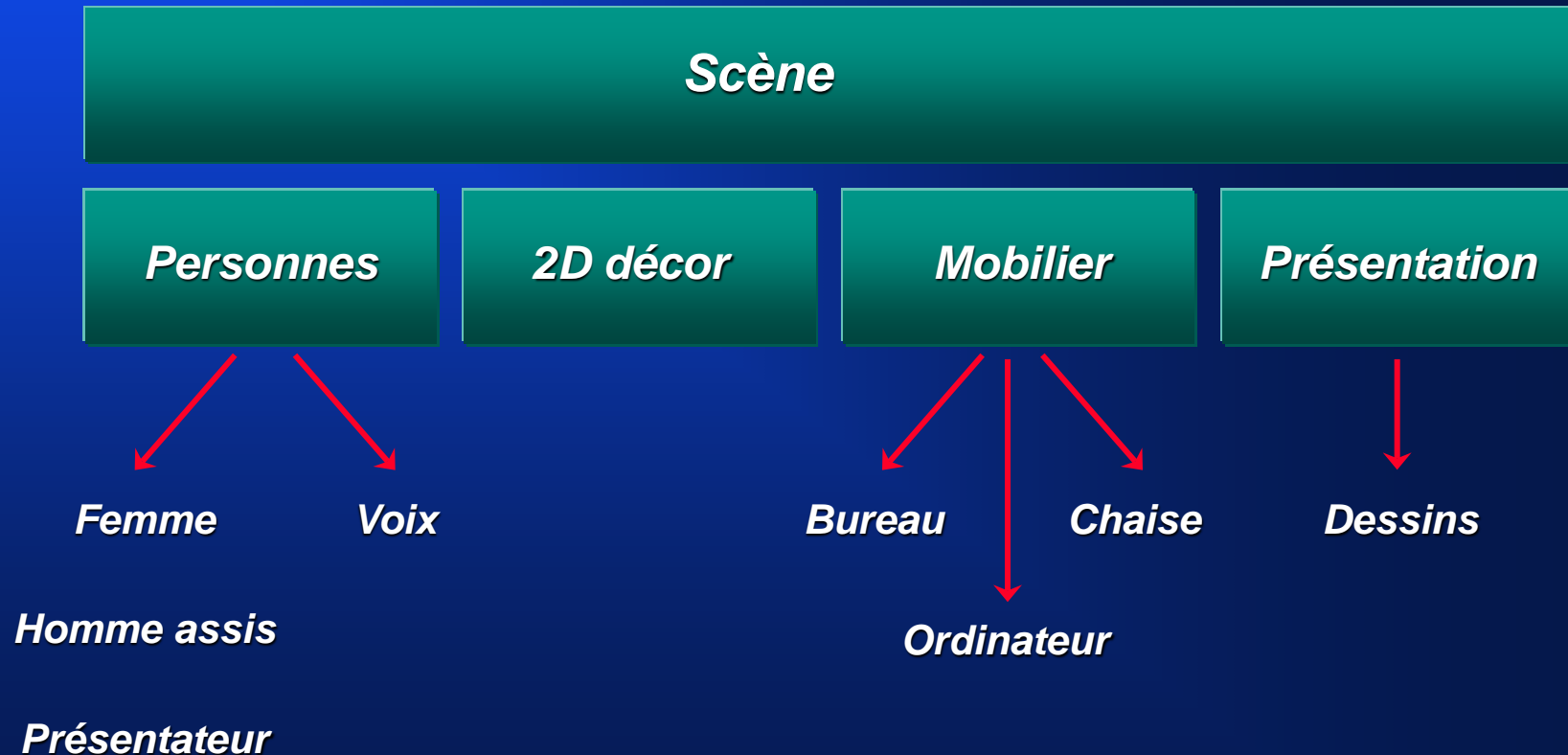


- La compression MPEG 4 est qualitativement très performante !
- Si on compare la qualité des images d'un match de tennis encodé en MPEG 4 AVC et MPEG 2 sur des machines comparables à 25 i/s, l'indice PSN est déjà au maximum pour un débit de 1500Kbps alors que pour la même qualité MPEG 2 nécessite un débit de 3000 Kbps *!

Les représentations objet de la norme MPEG-4



Systeme MPEG-4: composition de la scène



MPEG 4: Boite à outils générique

MPEG 4

*Encodage numérique descriptif d'objets audiovisuels,
des rapports entre ces objets et un contexte :*

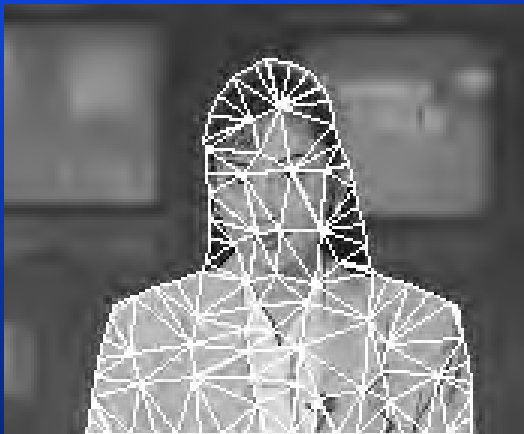
L'encodage MPEG4 exploite une "boite à outils" générique pour **segmenter** automatiquement et figurer une vidéo dynamique tout en **localisant et en caractérisant les objets** de manière compacte et individualisée. Facile à manipuler ! Toute manipulation engendre un flux d'informations variables dont les paquets sont hiérarchiquement emboîtés. On a une localisation spatiale et temporelle.



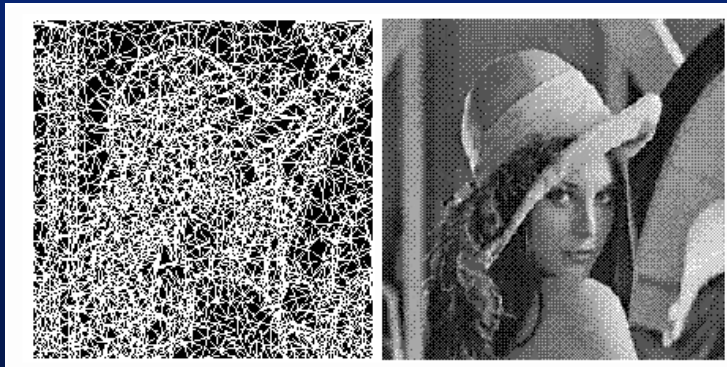
Exemple de segmentation d'une scène mobile par un encodeur MPEG4: reconnaissance des personnages en mouvement relatif.

MPEG 4 : maillage de chaque objet

Acquisition MPEG 4



L'acquisition exploite un système de **maillage automatique** produit par l'encodeur lequel découpe et structure les images 2D et 3D. La modélisation se fait par la projection d'un **maillage 3D composé de polygones** dont la finesse est déterminée par la structure du maillage et par une **liste plus ou moins importante de nœuds** (déterminé à l'acquisition en fonction du flux MPEG4 souhaité en final)



A gauche maillage automatique et à droite l'image "reconstruite" avec 3200 nœuds (c'est trop peu!). Dans ce cas 95% des échantillons ont été éliminés..

MPEG 4 : cartes de segmentation

Objets répertoriés, localisés, temporisés, à échelle variable

- Pour chaque objet vidéo et pour chaque séquence l'encodeur crée des **répertoires hiérarchiques** qui comprennent des couches d'informations successives pour décrire les objets (topologie, mouvement, formes, couleurs, textures, sons associés)
- La **localisation** est **spatiale** et **temporelle** selon une grille auto produite. Le maillage s'anime dans l'espace et le temps. La topologie est évolutive selon des cartes de segmentation qui prennent en compte diverses caractéristiques spatio-temporelles des mouvements
- Une autre caractéristique est la **scalabilité** = structuration multi échelle du travail d'analyse et de découpage :
 - Scalabilité des objets eux-mêmes
 - Scalabilité spatiale
 - Scalabilité temporelle
 - Scalabilité de la représentation
 - Scalabilité de la distribution (en termes de flux vers l'utilisateur)

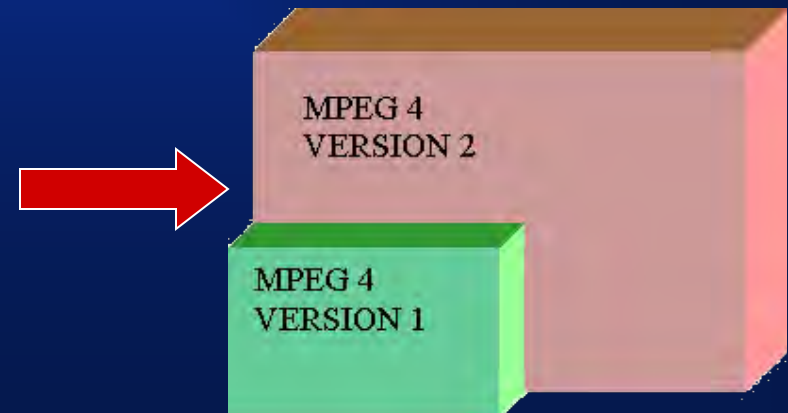
MPEG 4 : qualité à la demande

Les flux MPEG 4 :

- Les flux sont **variables**, **progressifs** et **hiérarchiquement emboîtés**.
- Les flux entrants dans les décodeurs usagers sont lus en fonction des caractéristiques des décodeurs - le consommateur pouvant intervenir sur des ingrédients de la composition séquentielle - c'est **la qualité à la demande** qui permet d'utiliser MPEG 4 sur des réseaux diversifiés à capacité \neq entre 1Mbps (xDSL) et 1000 Mbps (HD)

Les versions successives MPEG 4 sont "survit aminées" l'une par rapport à l'autre.

L' "intelligence" des dernières versions s'est déployée dans plusieurs dimensions...



MPEG 4 : exemple d'un traitement

**Exemple
d'un
traitement
MPEG 4 :
segmentation
et fusion
d'ingrédients**

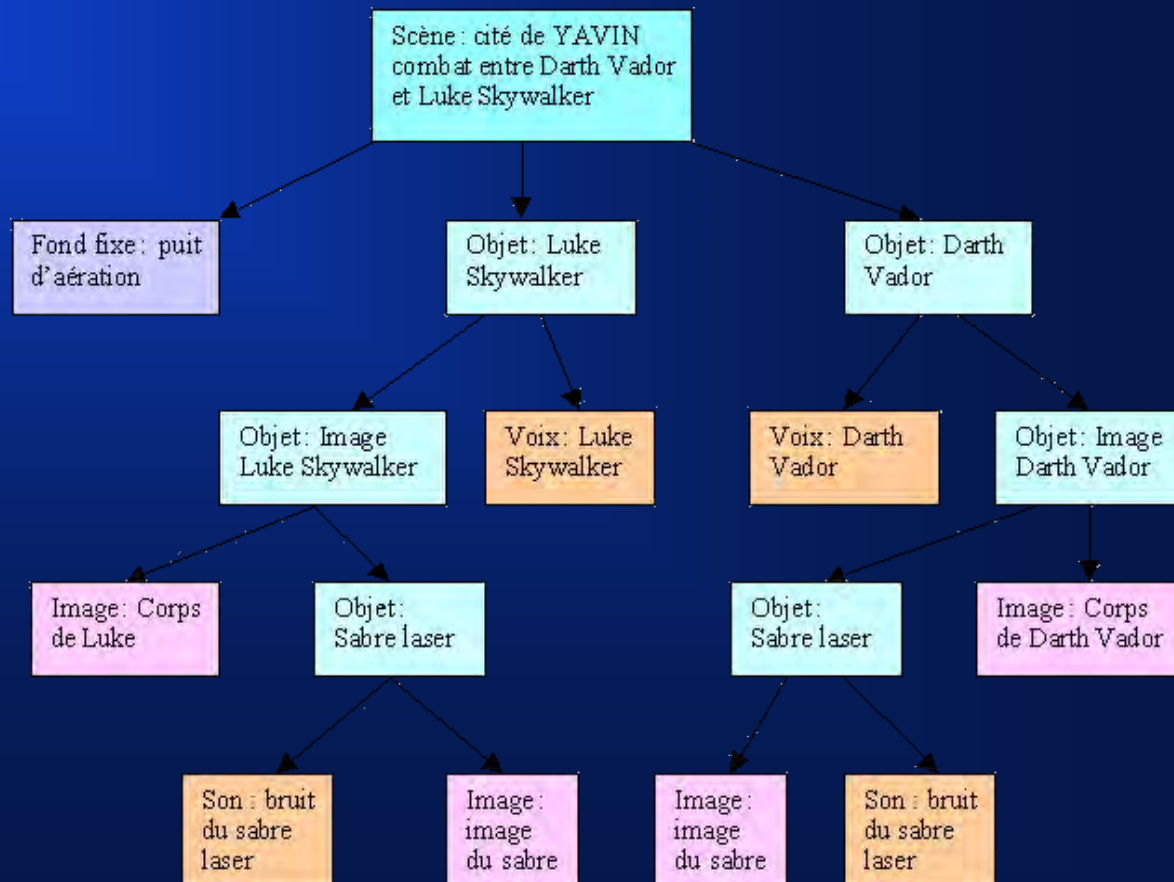


- L'encodeur isole le fond et recrée un pano du fond de scène complet (estimation et compensation de mouvement par blocs de 8 ou 16 pixels)
- L'encodeur extrait le personnage en mouvement
- Le fond est encodé une fois, seules les variations formelles y sont ré-encodées en fonction des besoins.
- L'encodeur incruste le joueur en mouvement tenant compte des zones masquées. Les zones non reconnues sont représentés par de la DCT
- Le décodeur recrée la scène grâce aux paramètres de la caméra pour le fond et au joueur envoyé dans sa position à chaque image

MPEG 4 : répertoires hiérarchiques

Structure hiérarchique de graphe acylique

- Le traitement MPEG 4 "détourne", "reconnaît" et répertorie automatiquement les Objets Vidéo-Son (OV) ainsi que leurs relations synchrones (Image/son).
- Chaque "feuille" représente un élément simple.
- Les relations (mouvement relatif...) entre les OV sont décrites



Le MPEG 4 : en développement constant

- Amélioration du codage des visuels "naturels"
- Animation des corps
- Codage en maillage des objets 3D
- Améliorations audio
- Améliorations systèmes
- BIFS avancé* (Binary Format for Scene Description)
- Définition d'un format de fichier M4F
- Améliorations protocolaires

MPEG 4 est une norme "en mouvement" régulièrement améliorée, certains développements pouvant être adressés aux décodeurs des usagers par les diffuseurs.

*Binary Format for Scene Description (descripteur de la synchronisation dynamique des objets dans une scène encodée/décodée)

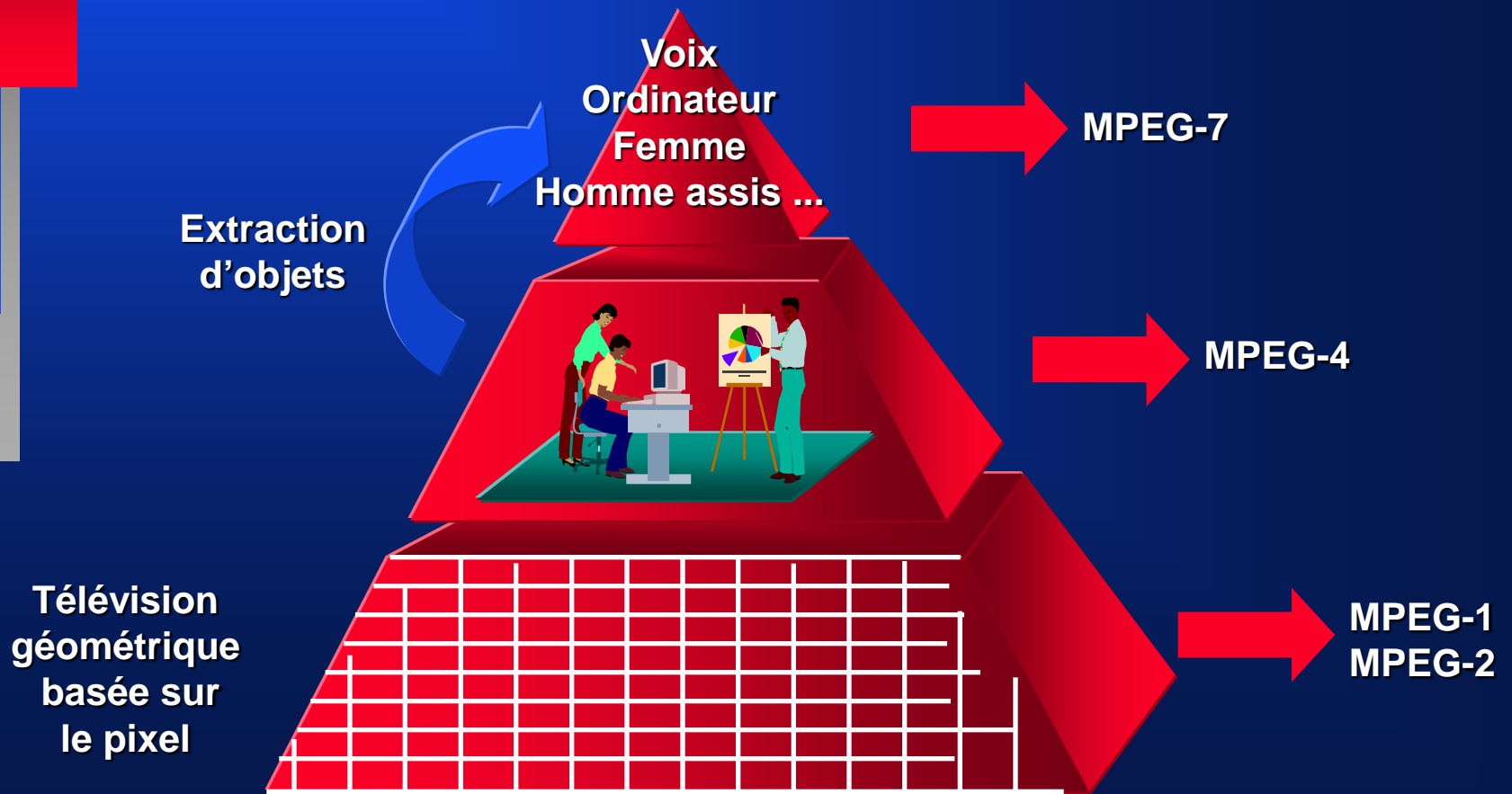
Nouveaux concepts: les métadonnées

- *Donnée: Informations autres que l'image et le son: time code, numéro de scène, etc.*
 - *Les essences: images et sons numérisés.*
 - *Les métadonnées: interactions entre ces éléments.*
 - *Le média intégré qui en résulte.*
 - *Les liens interactifs qui pointent vers des données ou des essences sont établis par les métadonnées.*
1. Le MPEG 4 et le MPEG 7
 2. Le XML
 3. L'AAF et le MXF

MPEG 4: Base pour l'indexation

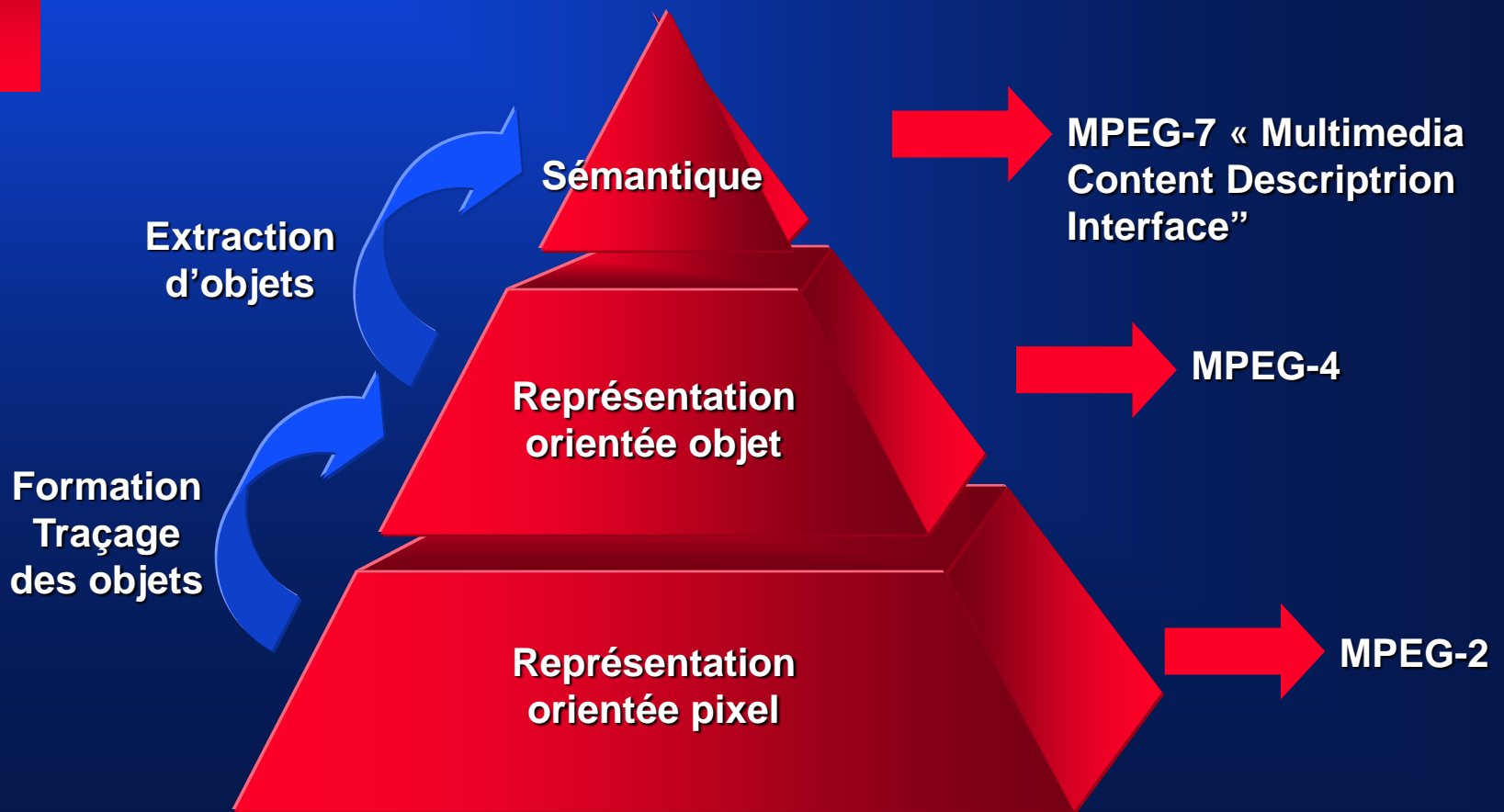
- ❑ ***Le MPEG 4 permet l'encodage descriptif d'objet visuel***
- ❑ ***Contrairement au MPEG 2, le MPEG 4 permet lors de la numérisation une indexation semi-automatique des flux de données***
- ❑ ***L'objectif est de localiser, de caractériser et d'indexer ces objets de manière compacte et individualisée, les fiches devant être faciles à manipuler, toutes ces opérations engendrant un flux d'informations variable dont les « paquets » sont hiérarchiquement emboîtés, dans un fichier synchronisé au fichier de base par un outil de type SMIL 2.0 (Synchronised Multimedia Interactive Language).***

Relations entre les différentes normes MPEG 1, 2, 4, et 7



Relations entre les différentes normes MPEG 2, 4 et 7

- *Du pixel à l'objet avec un codage sémantique :*



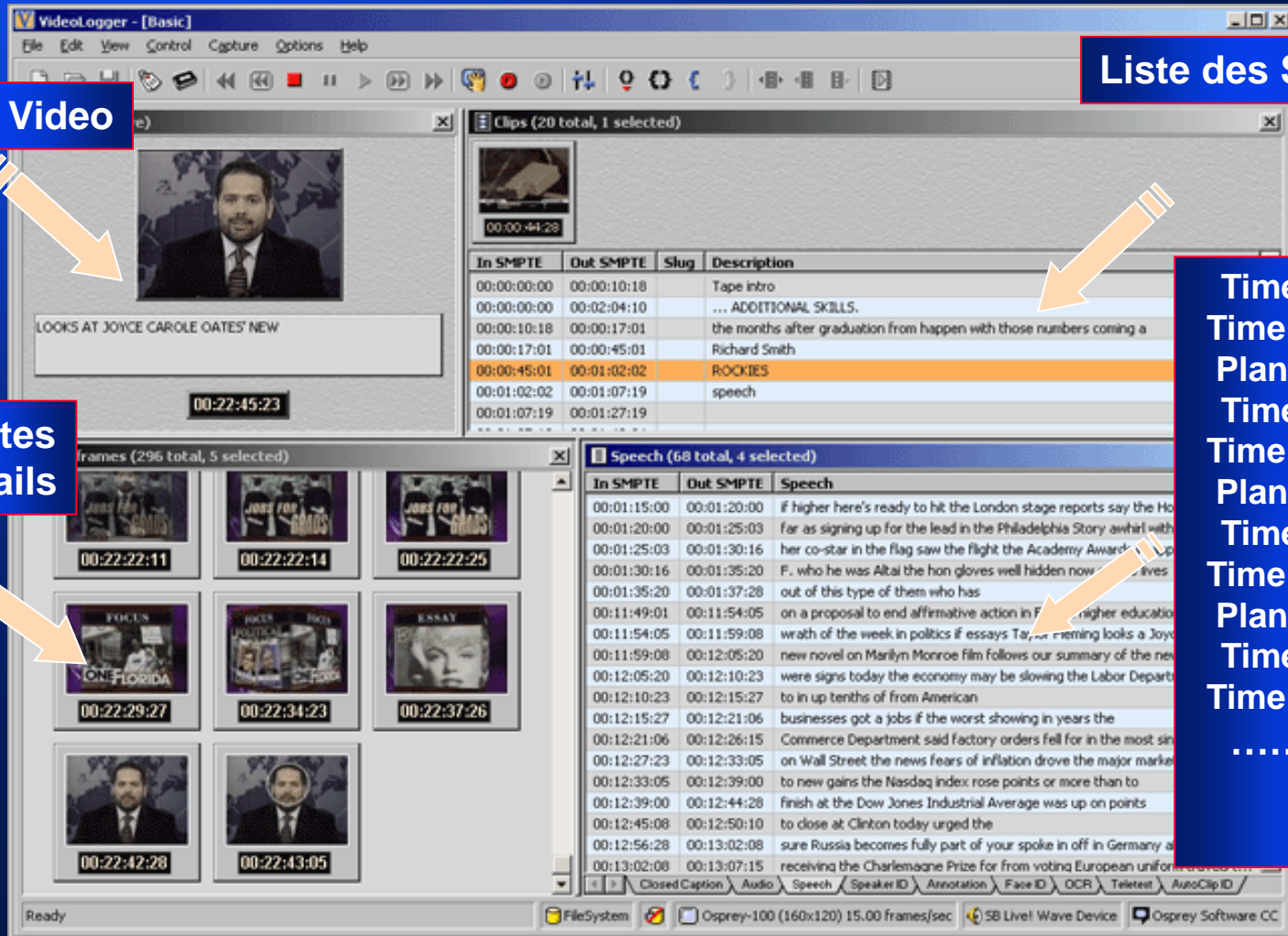
Le XML eXchange Markup Language

- ***Du SGML (Standard General Markup Language) créé en 1980 est issu un peu plus tard le HTML (Hyper Text Markup Language)***
 - *Exemple de balise HTML:*
 - `<Body><h1>Lemy</h1>
<h2>Patrick</h2></Body>`

- ***Le XML modifie la balise en ajoutant la sémantique.***
 - *Exemple de balise XML:*
 - `<Body><Nom>Lemy</Nom><Prénom>Patrick</Prénom></Body>`

- ***Grâce au XSD il offre une grande souplesse et une adaptabilité sur les serveurs ou Workflow.***

Ecran de travail en quatre zones : Rush image et son, imagettes plans, liste scène, données et méta données par plans :



Audio - Video

Liste des Scènes

Imagettes Thubnails

Time Code In
 Time Code Out
 Plan 1/ Shot 1
 Time Code In
 Time Code Out
 Plan 2/ Shot 2
 Time Code In
 Time Code Out
 Plan 3/ Shot 3
 Time Code In
 Time Code Out
