

Compression d'images format JPEG

- 1) Introduction (système de compression, les normes)
- 2) Formats des images (les couleurs, formats 4:2:0 ..., QCIF ...)
- 3) JPEG (schéma de codage, modes de fonctionnement)

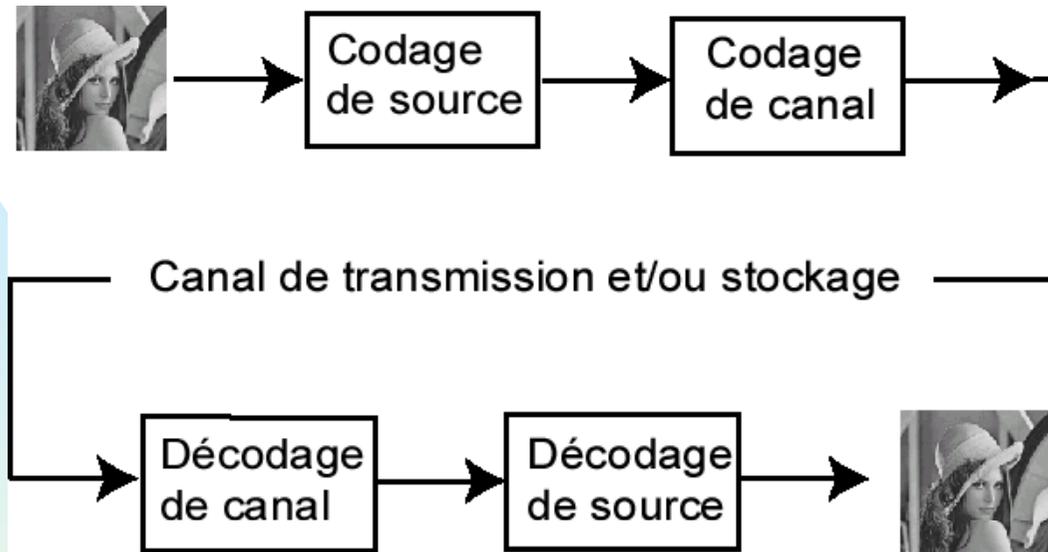
1) Introduction

- Contexte : codage ou compression des images numériques
- Pourquoi : réduction de la quantité d 'éléments binaires représentant l 'information « image »
=> codage de source

$$\text{Taux de comp.} = \frac{\text{Qté d'info. ima. originale [bit]}}{\text{Qté d'info. ima. compressée [bit]}}$$

- Finalité : archivage ou transmission

1.1) Systeme de compression



- Codage de canal :
adaptation signal / canal, taux d'erreur faible
- Codage de source, 2 types :
 - ◆ Codage sans perte (« entropique »)
Ex. : Huffman, Lempel-Ziv, Arithmétique, ...
=> taux de compression faible (1.5 à 2)

- ◆ Codage avec pertes (« irréversible ») :
Suppression des redondances
cad de l'information inutile car :
 - prévisible
 - invisible par système visuel humain (SVH)
- => Taux de compression élevés (>10)

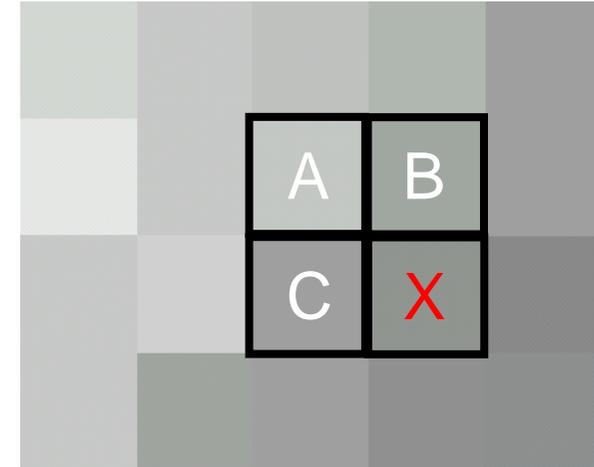
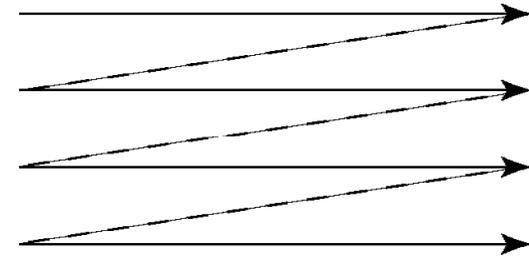
- Exemple :

- ◆ image – « matrice de pixels »
- ◆ pixel = « picture element »
- ◆ pixel codé sur un octet

- 8 bits => 256 niveaux de gris
- « 0 » => noir
- « 255 » => blanc

$$2^8 = 256$$

- ◆ image transmise en la balayant (« Zigzag scan »)



- ◆ constat : le niveau de gris d'un pixel dépend souvent de celui de ses voisins
- ◆ idée : prédire X en utilisant A, B ou C
la base du codage MICD
(Modulation d'Impulsions Codées Différentielles)
=> une phase d'analyse de l'image

1.2) Les normes

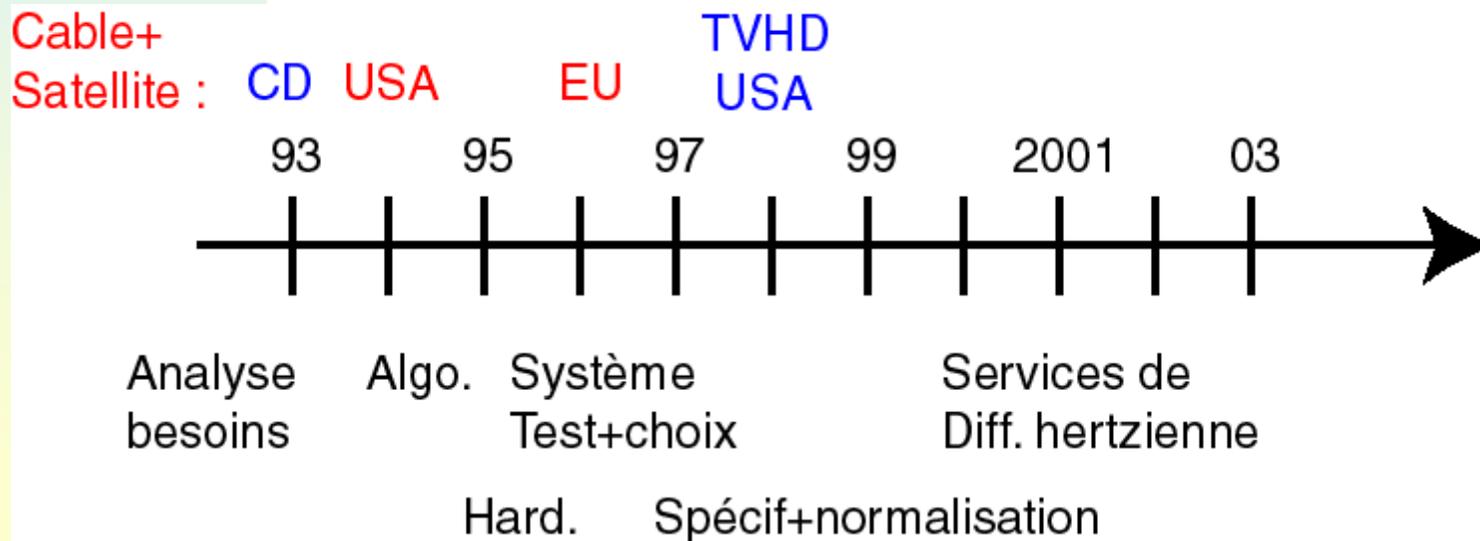
- Pourquoi : enjeux économiques pour les industriels
(Imposition d'un standart, manipulation de mêmes données, conception de « chips »)

- Par qui : organismes de normalisation
 - ◆ UIT (Union Internationales de Télécom.)
 - ISO (International Standart Organisation)
 - CCITT (Comité Consultatif International Télégraphique et Téléphonique)
 - CCIR (Comité Consultatif International pour la Radiodiffusion), ...

 - ◆ Groupes de travail réunissant des :
 - administrations (France serics, UK-DFI, Germany-DBPT, ITAIIA, ...)
 - opérateurs (France Telecom, NTL, TERACOM Scandinavia, ...)
 - industriels (Thomson, Philips, Grunding, Sony, Nokia, ...)
 - diffuseurs (RTL, TF1, RAI-ARD, BBC-RTVE, SDT Sweden, ...)

■ Comment :

- ◆ analyse du besoin
- ◆ appels d'offres
=> compétitions entre industriels et laboratoires
- ◆ analyse des résultats par les commissions
=> sélection + décision
- ◆ Ex. calendrier pour MPEG2 :



Historique

Il y a, au fur et à mesure, complémentarité entre les normes et complexification

■ Image fixe

◆ 1980 : Recommandation pour le fac simulé

◆ 1992 : **JPEG**

☞ « Joint Photographic Expert Group »

☞ images couleurs et N&B (Ex : satellite, médicales, ...)

☞ plusieurs modes (Ex : séquentiel, sans perte, progressif, hiérarchique)

☞ format image < (768x576)

☞ débits : de 8 M bit/s à 40 M bit/s

◆ 2000 : JPEG 2000

(débits inférieurs, haute robustesse aux erreurs de transmission, description basé contenu, large gamme d'images, interface avec MPEG4, ...)

■ Vidéo : visiophonie

◆ 1990 : H261

- ☞ téléphonie visuelle sur le RNIS [« ISDN »]
(Réseau Numérique à Intégration de Services, 64 k bit/s)
- ☞ formats image CIF (« Common Intermediate Format ») et QCIF (« Quarter CIF »)
- ☞ débits : p x 64 k bit/s (p : de 1 à 30)

◆ 1996 : H263

- ☞ débits inférieurs car pour réseau LAN (28.8 k bit/s)
- ☞ formats image SQCIF, 4CIF, 16CIF

Format	Résolution ima.		H261	H263	Débits Mbit/s	
	Lum.	Chrom.			N&B	Couleur
SQCIF	128x96	64x48		oui	3.0	4.4
QCIF	176x144	88x72	oui	oui	6.1	9.1
CIF	352x288	176x144	opt.	oui	24.3	36.5
4CIF	704x576	352x288		opt.	97.3	146.0
16CIF	1408x1152	704x576		opt.	389.3	583.9

30 images/s

■ Vidéo : archivage et diffusion

◆ 1988 : fondement de **MPEG**

- ☞ « Moving Picture Coding Experts Group »
- ☞ 3 parties : **Vidéo**, Audio et Système (le « stream »)

◆ 1993 : MPG1

- ☞ but : la qualité VHS sur un CD-ROM (=> stockage)
débit bas de 1.5 M bit/s (ima. 352x288 + audio)
- ☞ un accès aléatoire au sein de la séquence

◆ 1994 : MPEG2

- ☞ but : diffusion de la vidéo (=> application télévisuelle)
- ☞ adaptation aux formats (« scalabilité »)
 - taille image 4/3 16/9
 - entrelacé ou non, formats 4:2:0, 4:2:2, 4:4:4, compatibilité entre niveaux de qualité (normale<-> HDTV)
- ☞ MPEG3, originellement prévu pour la HDTV (TV Haute définition) est inclus dans MPEG2

MPEG2

"Level"	Taille	Débits M bit/s	Application
Low	352x288x30	4	enreg. vidéo
Main	720x576x30	15	studio TV
High 1440	1440x1152x60	60	HDTV
High	1920x1152x60	80	production film

◆ 1998 : MPEG4

☞ décomposition de la scène en éléments VOP

- VOP « Video Object Plane »
- chaque élément = une composante audio + une comp. Vidéo
- éléments réels ou de synthèse, forme arbitraire, ...
- scalabilité
- une « boîte à outils de compression » pour les VOP

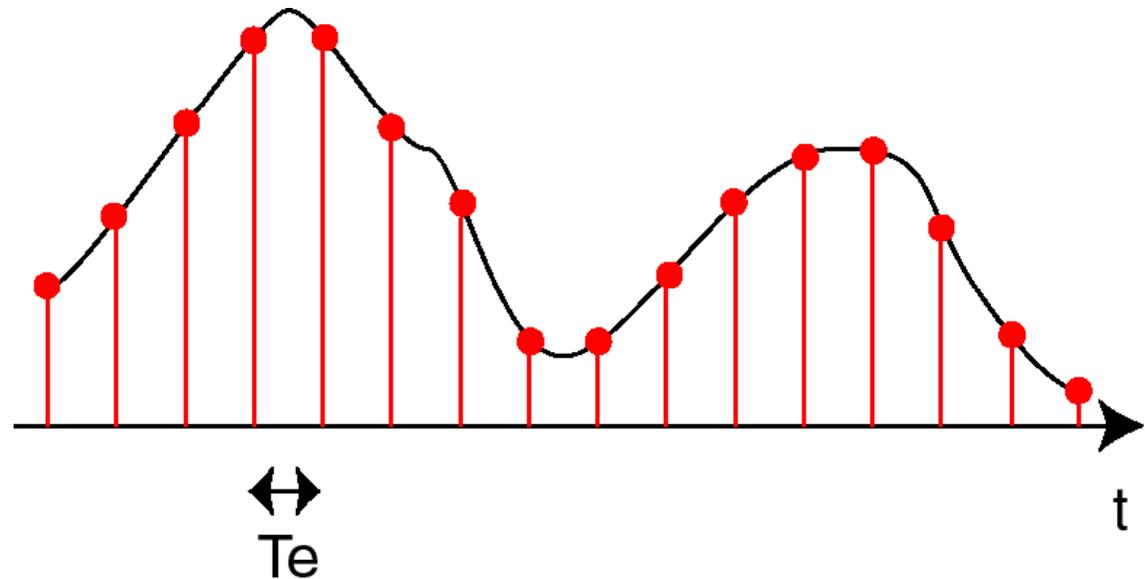
☞ débits : 5 k bit/s à 5 M bit/s pour la vidéo
2 k bit/s à 64 k bit/s pour l'audio

◆ 2000 : MPEG7 pour le multimédia (... jusqu'à la description sémantique de la scène)

2) Les formats image (représentations des images)

2.1) Echantillonnage et quantification

■ Echantillonnage

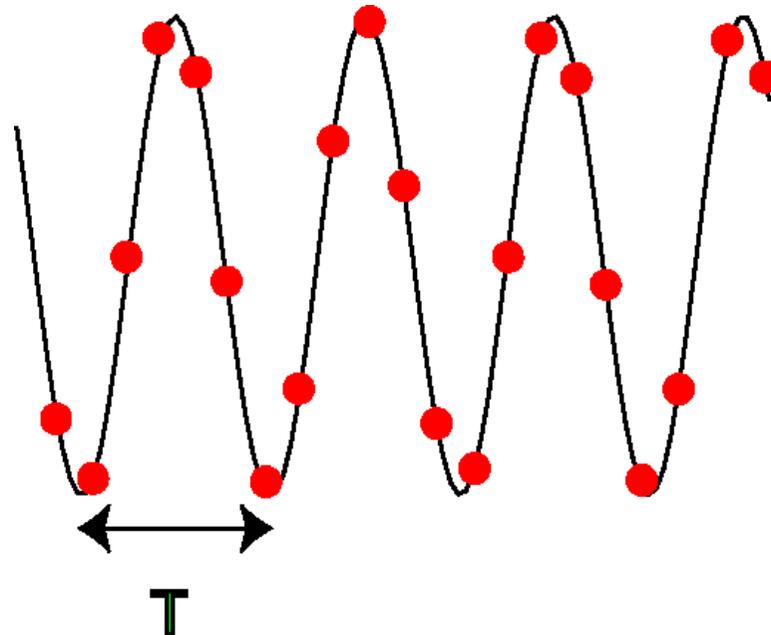
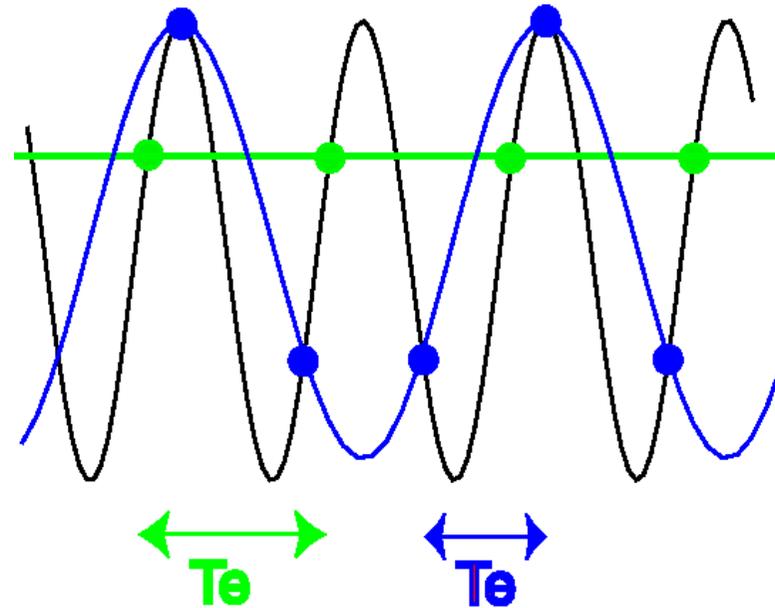


- ◆ un signal continu
- ◆ diviser l'axe des temps en intervalles réguliers (discrétisation)
- ◆ échantillonnage uniforme $\Rightarrow T_e$ constante

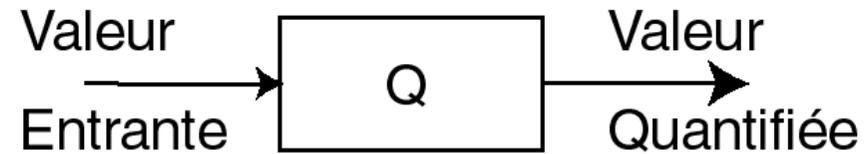
■ Théorème de Nyquist (échantillonnage) :

◆ Période
d'échantillonnage
trop grande
=> « aliasing »

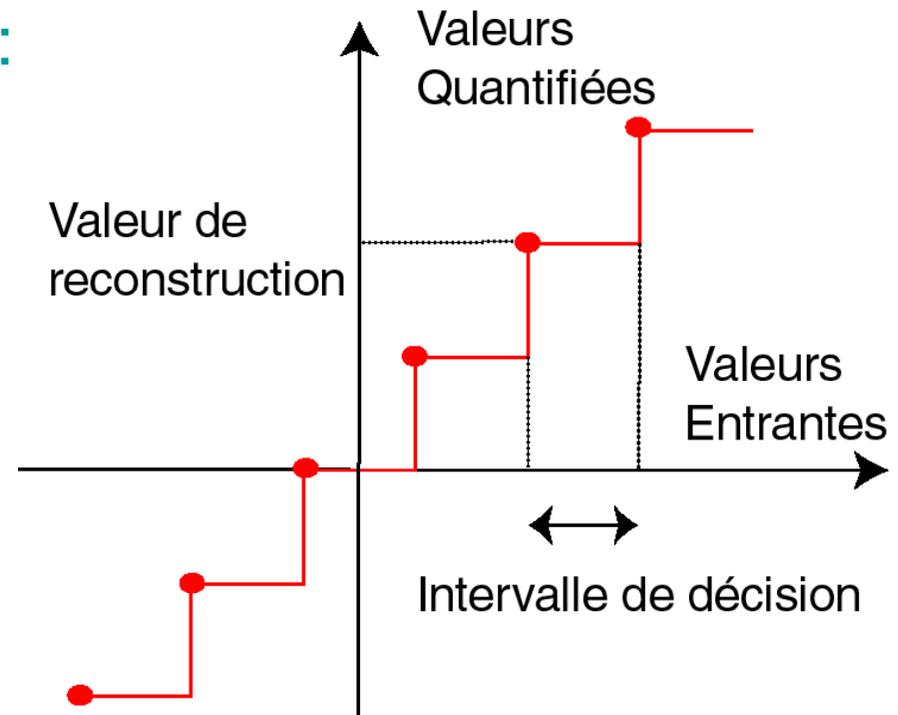
◆
$$T_e \leq \frac{T}{2}$$
$$F_e \geq 2 \times F$$



■ Quantification :



- ◆ discrétisation de l'axe vertical
- ◆ quantification uniforme : des intervalles égaux



- ◆ quantification non-uniforme ...

2.2) Les modèles des couleurs

- Bases :
 - ◆ spectre de la lumière visible : 400 -700 nm
 - ◆ oeil = une caméra
 - ◆ excitation des neurones par :
 - ☞ les batonnets (la luminosité)
 - ☞ les cones (la couleur)
 - ◆ 3 types de cones : 3 domaines de fréquences
 - => 3 couleurs suffises
 - pour reproduire tout le spectre visible
 - => 3 couleurs primaires

- Modèles pour les images :
 - ◆ RGB (« Red, Green, Blue »)
 - => les couleurs primaires additives
 - ◆ CMY (« Cyan, Magenta, Yellow »)
 - => les couleurs primaires soustractives

- Modèles pour la vidéo,
par transformation des RGB :
 - ◆ YUV :
 - ☞ Y : la luminance (image N&B)
 - ☞ U, V : les chrominances (la différence entre une couleur et un « blanc » de référence)
 - ◆ YCbCr
 - ◆ YIQ

2.3) Les types d'images (fixes)

- ◆ binaire ou monochrome :
pixel codé sur « 0 » ou « 1 »
- ◆ couleur 24 bits :
pixel sur 3 octets (Ex. RGB)
256x256x256 couleurs possibles
- ◆ couleur 8 bits
256 couleurs possibles
=> tables de conversion des couleurs
(LUT « Look-Up Table)
- ◆ formats : GIF (couleur 8 bits, compression sans perte, animée),
JPEG (tous types, compression avec pertes)
TIFF (tous types, compression sans pertes)
Postscript (sans compression)
BMP (par Microsoft, couleur 24 bits)

2.4) Les types de signaux vidéo

- ◆ 3 types de signaux vidéo, les 3 composantes sont :
 - ☞ soit séparées
=> large bande + synchronisation
 - ☞ soit mixées (vidéo composite) => interférences
 - ☞ S-VHS (mixage que des chrominances) => compromis

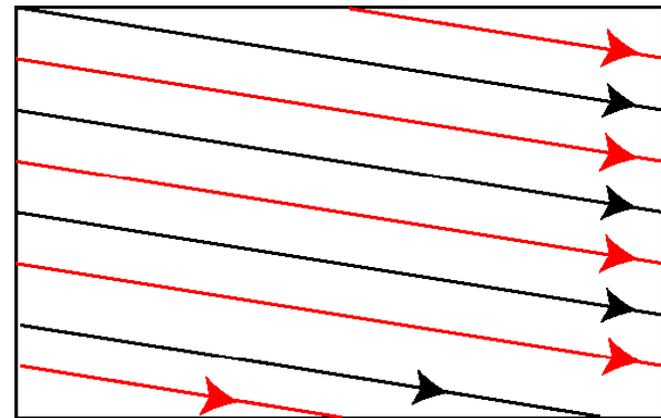
- ◆ Vidéo analogique :

- ☞ NTSC

- 525 lignes/image dont 20 lignes de contrôle (TV ~320 lignes)
 - image entrelacée
 - 30 images/s, modèle YIQ

- ☞ PAL

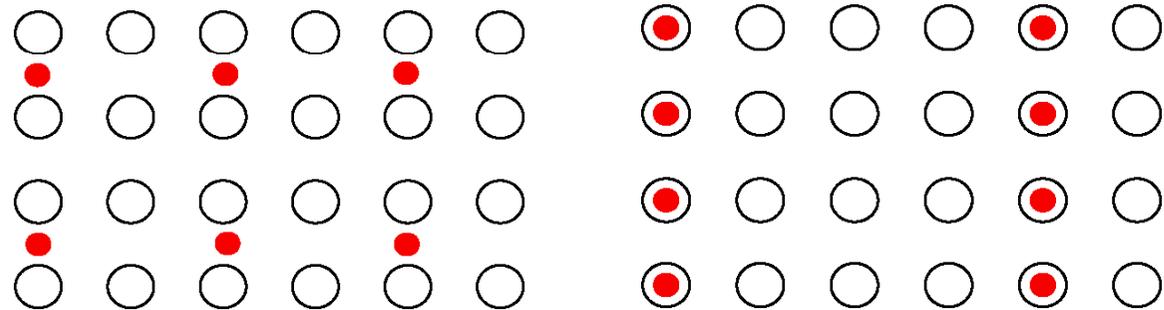
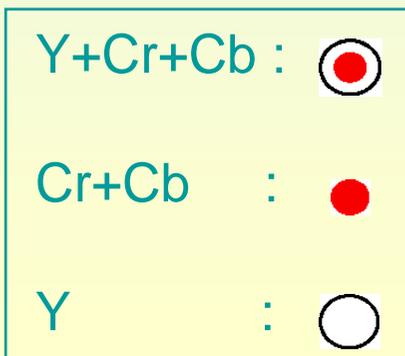
- 625 lignes/image
 - 25 images/s
 - image entrelacée, modèle YUV



◆ Vidéo numérique

👉 sous-échantillonnage de la chrominance :

- 4:4:4 : pas d'échant.
- 4:2:2 : échant. horizontal d'un facteur 2
- 4:1:1 : échant. horizontal d'un facteur 4
- 4:2:0 :
 - échant. horizontal et vertical d'un facteur 4
 - positionnement du pixel de chrominance



👉 Normes pour la vidéo numérique par le CCIR
(« Consultative Committee for International Radio »)

	CCIR 601 525/60 NTSC	CCIR 601 625/50 PAL/SECAM	CIF	QCIF
Résolution Luminance	720x485	720x576	352x288	176x144
Résolution Chrominance	360x485	360x576	176x144	88x72
Sous-échant.	4:2:2	4:2:2	4:2:0	4:2:0
Images/s	60	50	30	30
Entrelacé	oui	oui	non	non

→ progressif

- débit : ~165 M bits/s pour CCIR 601 (NTSC)
- CIF ~ qualité VHS
- format 4/3

👉 Normes pour la TV numérique par le ATSC
(« Advance Television Systems Comittee »)

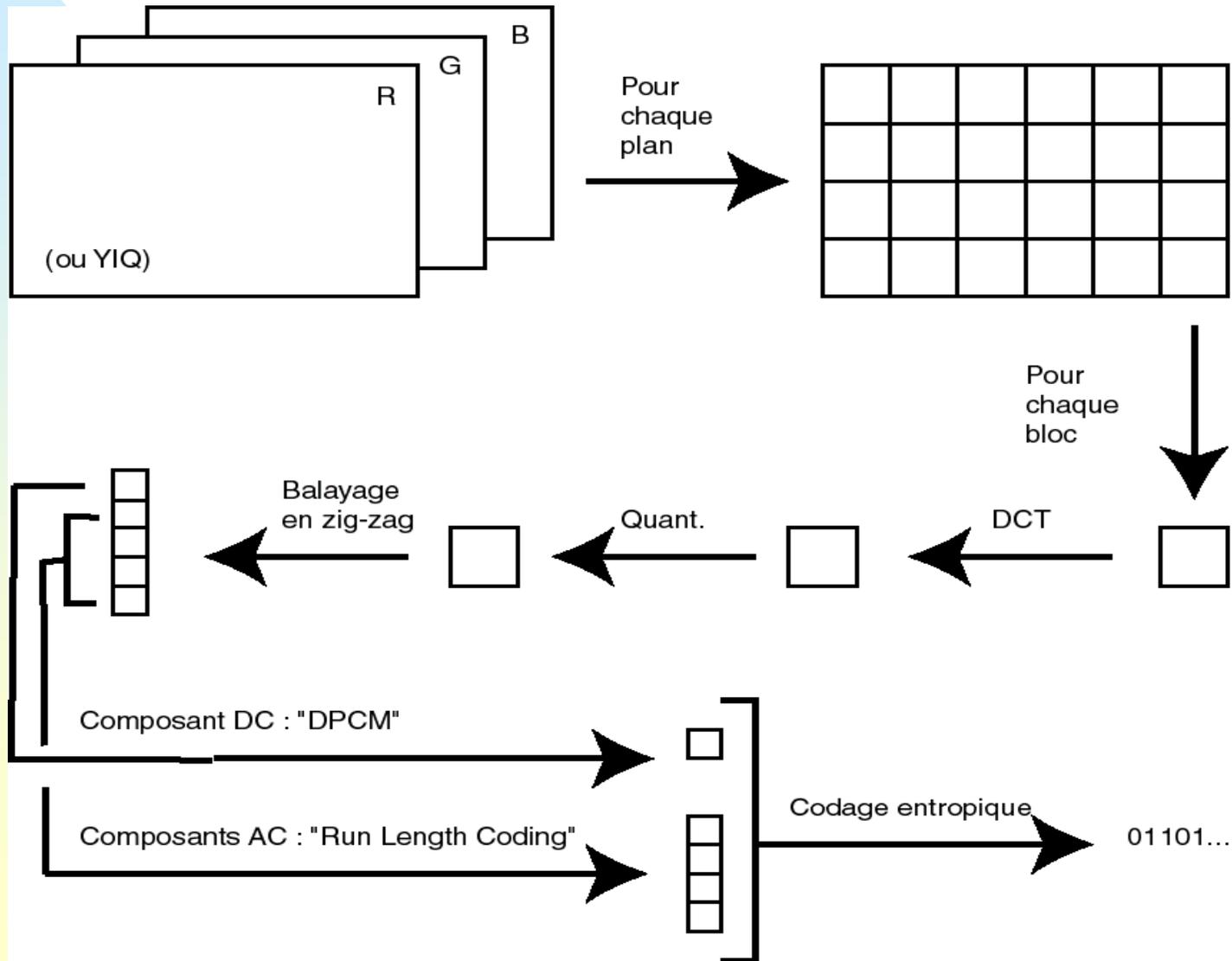
1995 : la norme pour la HDTV (« High Definition TV »)

Nb. lignes	Nb. colonnes	Format	Fréquence
1080	1920	16/9	60E 30P 24P
720	1280	16/9	60P 30P 24P
480	704	16/9 & 4/3	60E 60P 30P 24P
480	640	4/3	60E 60P 30P 24P

E : entrelacé P : progressif

3) JPEG

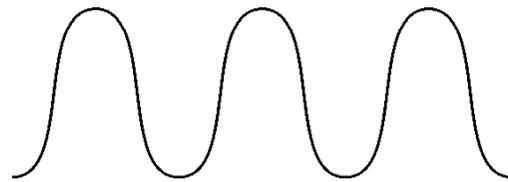
3.1) schéma général



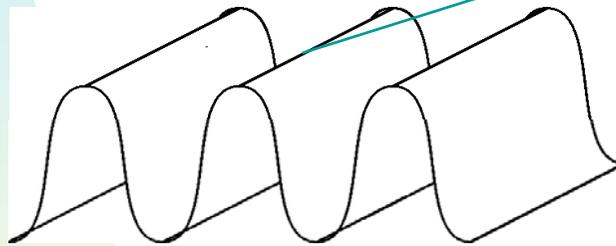
3.2) DCT

- Notion de fréquences :

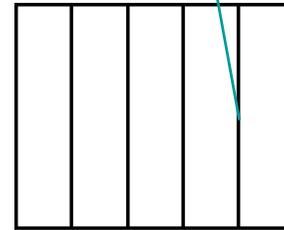
- ◆ 1D



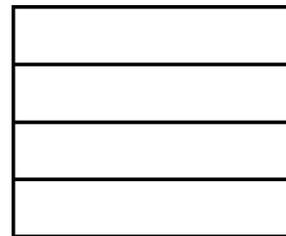
- ◆ 2D



crêtes

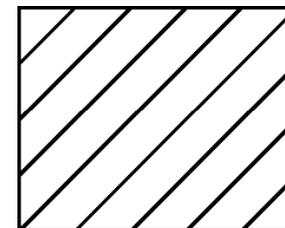


- ☞ horizontales

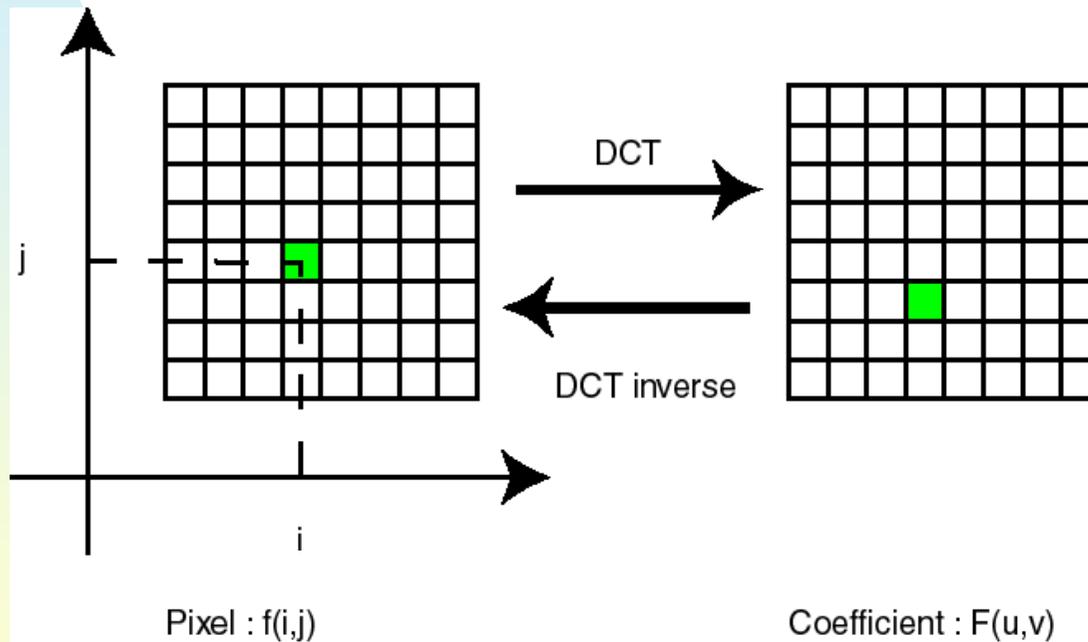


- ☞ verticales

- ☞ diagonales



- DCT (« Discrete Cosine Transformation »)
 - ◆ Transformation Discrète en Cosinus
 - ◆ changement de l'espace de représentation :
passage du domaine spatial au domaine fréquentiel



64 pixels

64 coefficients

■ DCT (suite), définitions

◆ DCT

$$F(u, v) = \frac{A(u) \times A(v)}{4} \times \sum_{i=0}^7 \sum_{j=0}^7 \cos\left(\frac{(2 \times i + 1) \times u \times \pi}{16}\right) \times \cos\left(\frac{(2 \times j + 1) \times v \times \pi}{16}\right) \times f(i, j)$$

$$A(\alpha) = \begin{cases} 1/\sqrt{2} & \text{si } \alpha = 0 \\ 1 & \text{sinon} \end{cases}$$

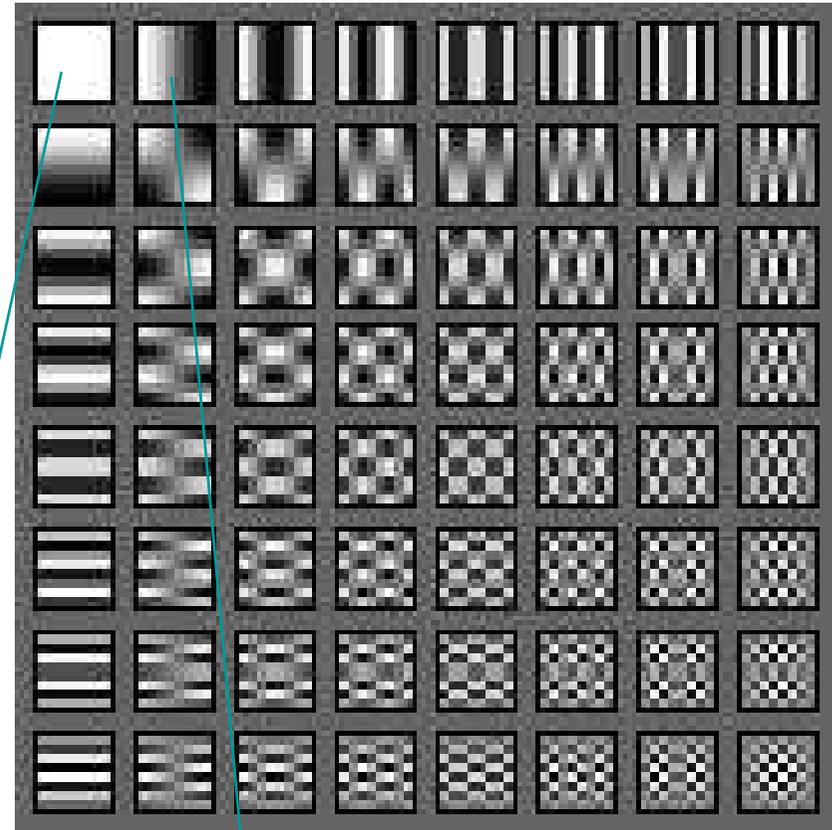
◆ DCT inverse

$$\hat{f}(i, j) = \frac{1}{4} \times \sum_{u=0}^7 \sum_{v=0}^7 A(u) \times A(v) \times \cos\left(\frac{(2 \times i + 1) \times u \times \pi}{16}\right) \times \cos\left(\frac{(2 \times j + 1) \times v \times \pi}{16}\right) \times F(u, v)$$

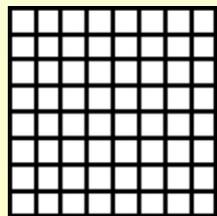
$$A(\alpha) = \begin{cases} 1/\sqrt{2} & \text{si } \alpha = 0 \\ 1 & \text{sinon} \end{cases}$$

■ DCT (suite)

- ◆ une décomposition sur 64 fonctions de base (ou sous-images de base)



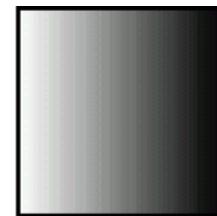
bloc



$$= F(0,0)x$$



$$+ F(0,1)x$$



+ ...

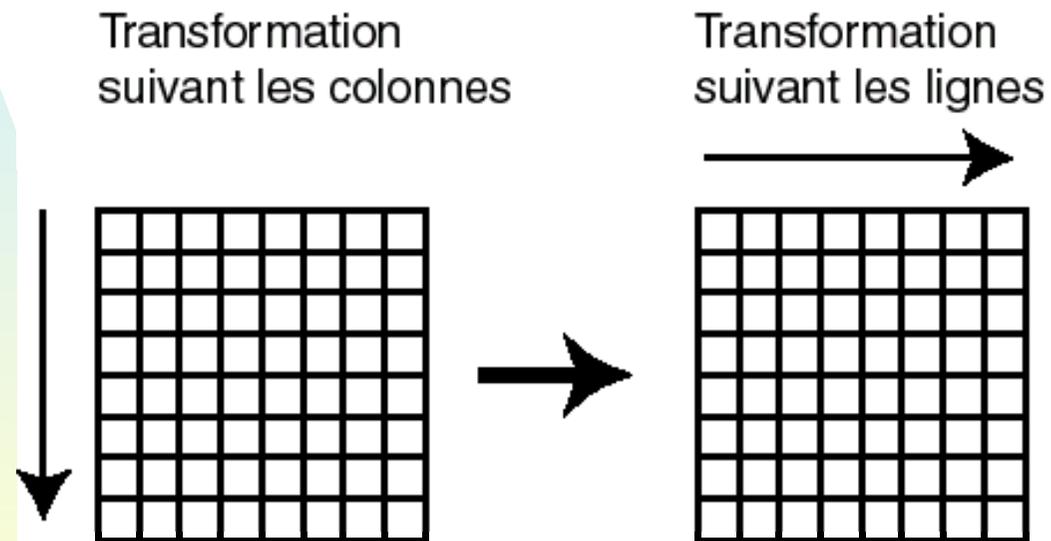
sous-images

- DCT (suite)

- ◆ implémentation

- 👉 des algorithmes rapides de calcul

- 👉 transformation 2D \rightsquigarrow 2 x transformations 1D



3.3) Quantification

- Pourquoi : SVH moins sensible aux hautes fréq.
- Idée : moins de bits pour les coeff. relatifs à ces fréq.
- Quantification : $F'(u, v) = \text{ent}[F(u, v) / q(u, v)]$
- Reconstruction : $F''(u, v) = F'(u, v) \times q(u, v)$
- Ex. $F(u, v) = 45$ [décimal] = 101101 [binaire] \Rightarrow 6 bits
 - ◆ $q(u, v) = 4$
 - ◆ $F'(u, v) = \text{ent}[45 / 4] = \text{ent}[11.25] = 11$ [décimal] = 1011 [binaire] \Rightarrow 4 bits
 - ◆ $F''(u, v) = 11 \times 4 = 44$ [décimal] $\Rightarrow F''(u, v) \approx F(u, v)$

Une erreur de quantification existe
(cf. codage avec pertes)

- Quantification uniforme : tous les $q(u,v)$ égaux
- Quantification non-uniforme : tables des $q(u,v)$

Vers les htes. Fréq.

16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	50	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

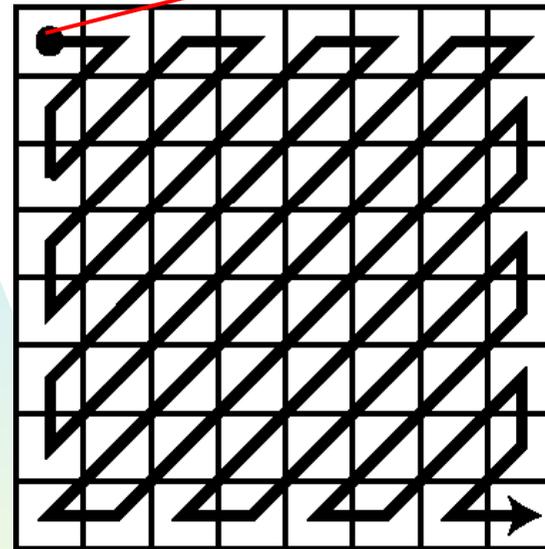
pour la Luminance

17	18	24	47	99	99	99	99
18	21	26	66	99	99	99	99
24	26	56	99	99	99	99	99
47	66	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99

pour les Chrominances

- ◆ $q(u,v)$ grand \Rightarrow quantification grossière
- ◆ tables peuvent-être transmises dans l'en-tête (« header ») de l'image

3.4) Balayage en zig-zag (du bloc)



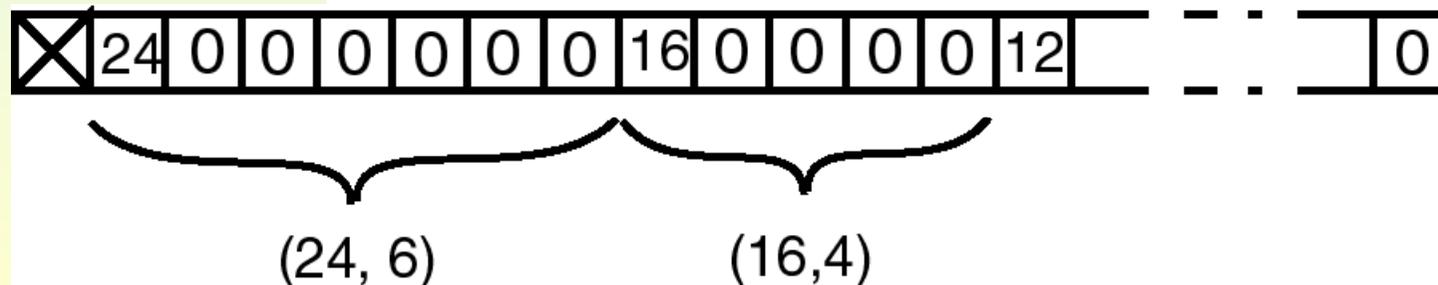
Coeff. DC

Coeff. AC
des
Basses fréq.

- Intérêt : former un vecteur où les coeff. relatifs aux basses fréq. sont regroupés

3.6) Codage des coefficients AC

- Constat : apparition de longues plages de 0 après quantification
- Méthode : codage de ces plages («Run Length Coding»)
 - ◆ un ensemble de paires (Coeff., Nb. de 0)
 - ◆ fin d'un bloc : paire (0, 0) :



3.7) Codage entropique

- Principe :



- ◆ codeur « classique » :

un alphabet de mots de code de même longueur

(Ex. pour 8 symboles : 000, 001, 010, 011, 100, 101, 110, 111)

- ◆ codeur entropique :

- ☞ un alphabet de mots de code de longueurs différentes
(Ex. 00, 01, 10, 110, 010, 111, ...)

- ☞ attribuer aux symboles les plus **probables**, les mots de code les plus courts

- ☞ effet : réduction du coût moyen (débit [en bit])
de la transmission

Codage entropique avec JPEG

- Coeff. DC :

- ◆ mise en correspondance :
valeurs coeff. /
taille du mot représentant [bit]
- ◆ Ex. : si DC=-9 => 4 bits
- ◆ transmission de la paire :
(Taille, Valeur)

Taille [bit]	Valeurs
1	-1, 1
2	-3, -2, 2, 3
3	-7, ..., -4, 4, ..., 7
4	-15, ..., -8, 8, ..., 15
⋮	⋮
10	-1023, ..., -512, 512, ..., 1023

- Coeff. AC :

codage de Huffman des paires (Coeff., Nb. de 0)

- Table de Huffman (cad l'alphabet des mots de code) :
 - ◆ pré-existante ou
 - ◆ construite (transmission dans l'entête)

3.9) Les 4 modes d'utilisation de JPEG

- **Mode séquentiel :**
 - ◆ le mode de base
 - ◆ chaque plan est codé directement par un balayage (gche->dt, haut->bas)

- **Mode sans perte :**
 - ◆ un codage sans perte
 - ◆ utilisation d'un prédicteur pour coder les coefficients
 - ◆ taux de compression de 1,5 à 2

Les 4 modes d'utilisation de JPEG

■ Mode **progressif** :

- ◆ idée : transmettre d'abord une image de basse qualité, puis l'améliorer par des ajouts successifs
- ◆ 2 façons :
 - ☞ sélection spectrale : transmettre d'abord les coeff. DC et quelques coeff. AC, puis d'autres coeff. AC
 - ☞ approximations successives : transmettre d'abord des coeff. grossièrement quantifiés, puis les quantifier plus finement et transmettre cette nouvelle information

■ Mode **hiérarchique** :

- ◆ répond aux besoins de « scalabilité » de certains décodeurs (dans un même flot binaire : plusieurs résolutions, plusieurs modes, ...)
- ◆ par l'imbrication d'opérations d'échantillonnages / codages / décodages / interpolations

3.10) Exemples : 1

Image originale (300 Ko)



Image compressée (10 ko)



Problème = effets de blocs



3.10) Exemples : 2

Image originale
(193 Ko)



Image compressée
QF 100 (56 ko)



Image compressée
QF 75 (12 ko)



Image
compressée
QF 50
(8 ko)



Image
compressée
QF 20
(5 ko)

