
Introduction au codage multimédia codage source

TS 200

2eme Année Electronique

ENSEIRB 2017

Electronique TS200

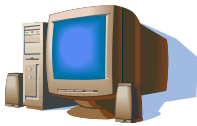
I Position du problème



ENSEIRB 2017

Electronique TS200

Avènement du numérique



Stockage

- Télévision
- Téléachat
- Téléformation
- Visiophone
- Vidéo conférence
- Vidéo surveillance
- Streaming ...



Transmission

I - Introduction

Télécommunications

► Supports physiques

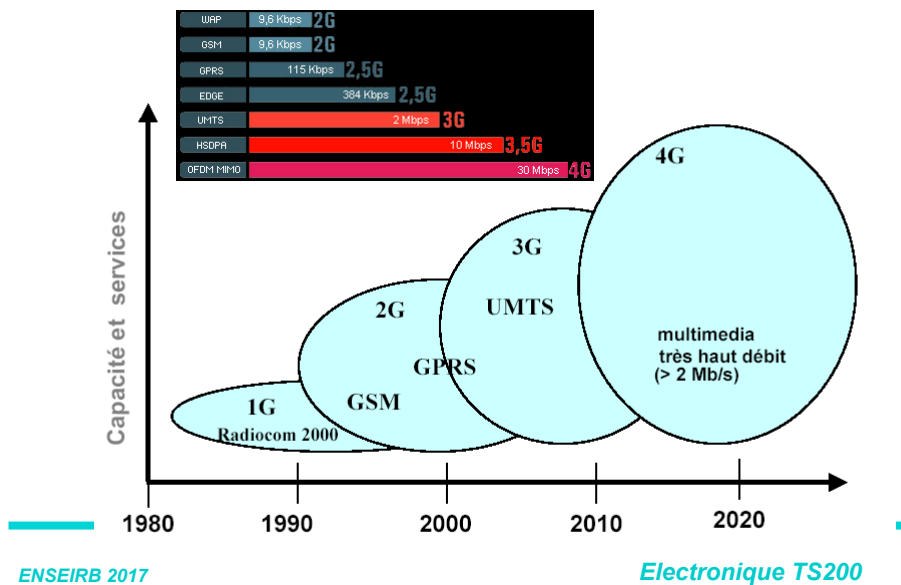
- Son
- Image
- Vidéo
- GPS ...

► Transmettre

- Réseau (streaming)
- Terrestre (TNT)
- Filaire (Adsl)
- Sans fils (WIFI, Wlmax)



Exemple : la téléphonie



Domaines scientifiques

- ▶ Mathématiques
 - Calcul intégral
 - Calcul différentiel
- ▶ Statistiques
 - Probabilité
 - Processus aléatoire
 - Théorie de l'information
- ▶ Théorie du signal
 - Signaux continus
 - Signaux discrets

Compression

JPEG
MP3
MPEG2

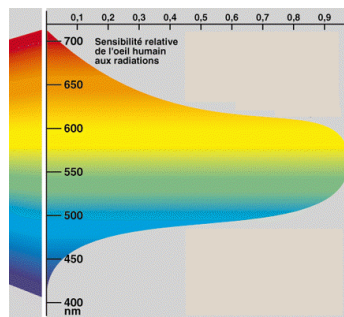
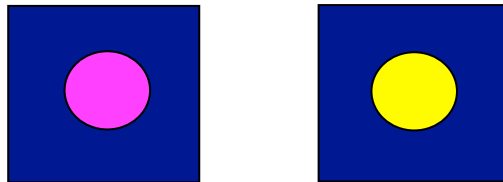
Canal

Code correcteur
Égalisation

Modulation

QAM
OFDM

Question de la couleur



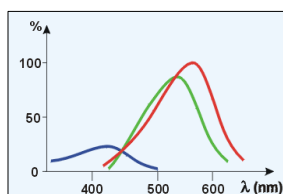
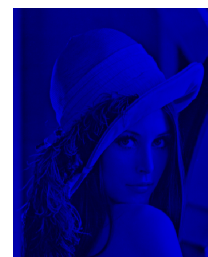
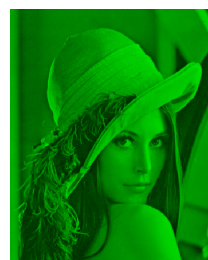
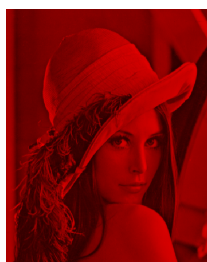
ENSEIRB 2017

Electronique TS200

Bilan codage



Niveaux de gris



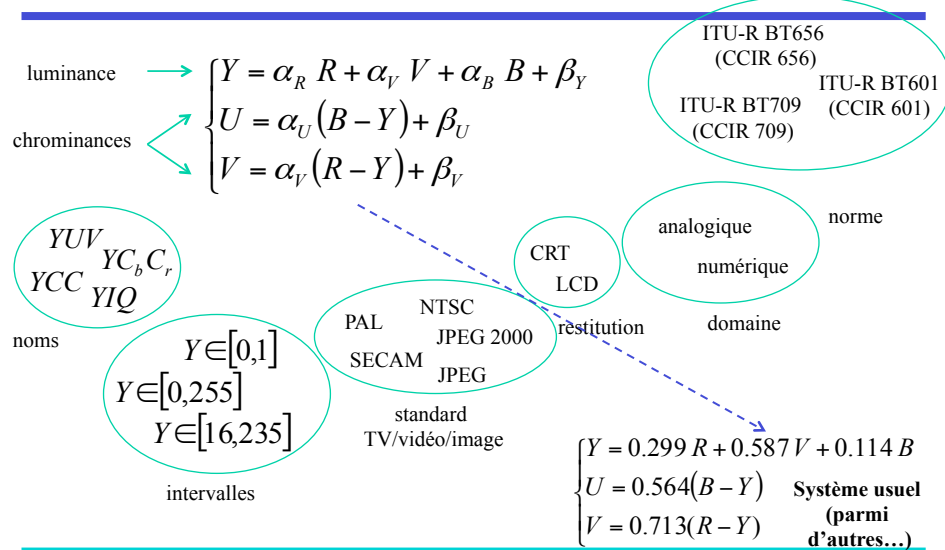
Palettes de primaires pures
(intensités identiques et autres composantes éteintes)

Sensibilité relative des photo-récepteurs de chrominance

ENSEIRB 2017

Electronique TS200

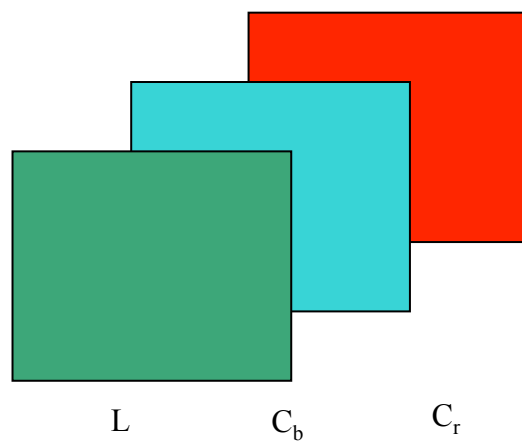
Bilan codage



ENSEIRB 2017

Electronique TS200

Format numérique

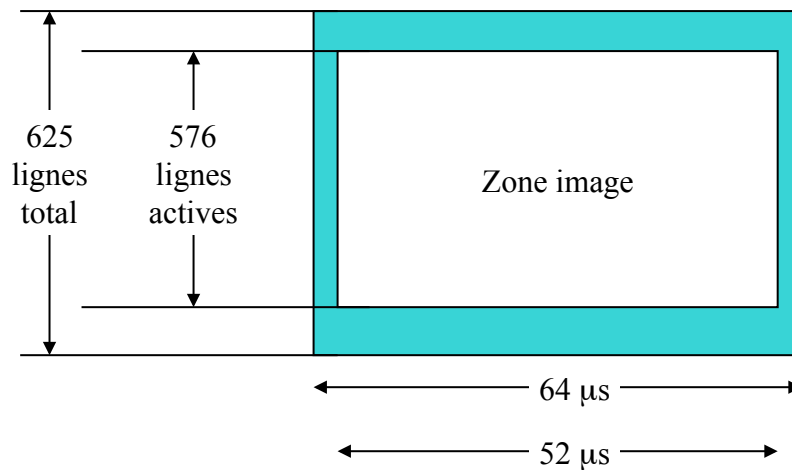


Conversion
A/N
 f_e

ENSEIRB 2017

Electronique TS200

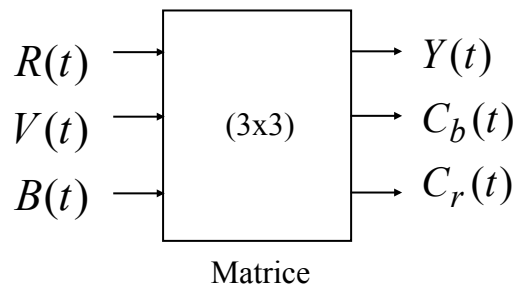
Format numérique



ENSEIRB 2017

Electronique TS200

Format numérique



$$Y = 0.299 R + 0.587 V + 0.114 B \quad \text{Signal de luminance}$$

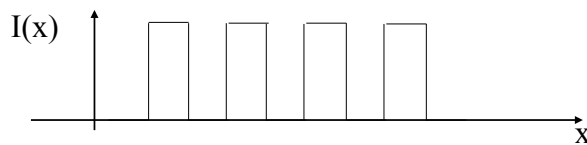
$$C_b = 0.564 (B - Y) \quad \text{Signal de Chrominance bleue}$$

$$C_r = 0.713 (R - Y) \quad \text{Signal de Chrominance rouge}$$

ENSEIRB 2017

Electronique TS200

Définition d'une BP



Période min = 2 points

$$F_{\max} = \frac{4 \cdot 576}{2 \cdot 3 \cdot 52 \mu s} \quad \rightarrow \quad \sim 6 \text{ MHz}$$

ENSEIRB 2017

Electronique TS200

Numérisation des signaux vidéo

Recommandation BT 601

Signal vidéo : bande passante 6MHz

Luminance

$F_{\text{ech}} = 13.5 \text{ MHz}$

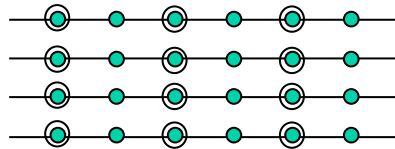
Chrominances

$F_{\text{ech}} = 6.75 \text{ MHz}$

ENSEIRB 2017

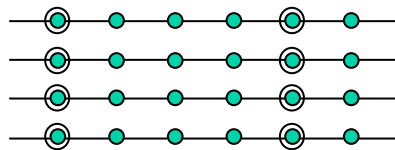
Electronique TS200

Formats et applications



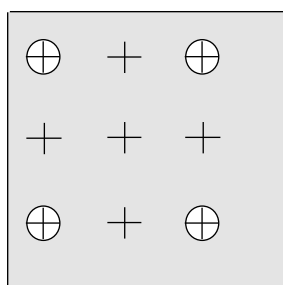
4:2:2

N1:N2:N3 N2/N1 Taux de sous échantillonnage ligne paire
N3/N1 Taux de sous échantillonnage ligne impaire



4:1:1

Le format 4:2:0



○ Chrominance

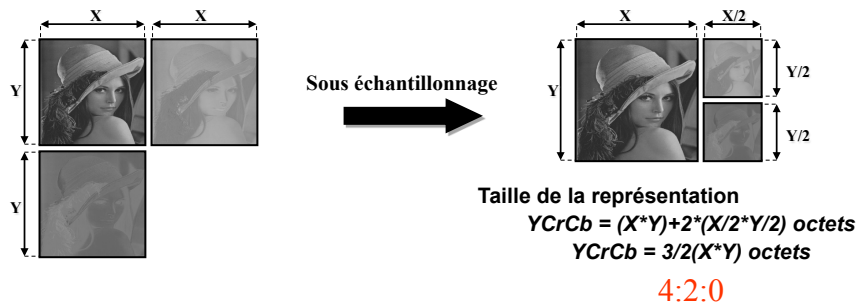
⊕ Luminance

Bilan : Luminance (720,576) standard 625 lignes

Chrominances (360,288)

Calcul du débit

$$(XY) \left[1 + 2 \left(\frac{N_2}{N_1} \frac{1}{2} + \frac{N_3}{N_1} \frac{1}{2} \right) \right]$$

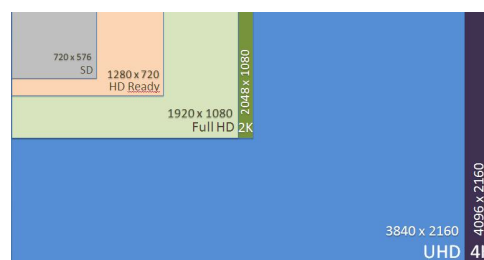


ENSEIRB 2017

Electronique TS200

Les formats

- VHS 200 lignes
- S-VHS (Hi-8) 400 lignes
- DV 500 lignes
- DVD 576 lignes (MPEG2)
- HD 720 lignes (MPEG4)
- HD (2K) 1080 lignes (p ou i)
- QFHD (4K) 2160 lignes
- UHD (8K) 4320 lignes (IMAX)



ENSEIRB 2017

Electronique TS200

Différentes fréquences d'échantillonnages

32 kHz (Télécommunication)

44.1 kHz (Cdrom grand public)

48 kHz (Studio, DAT, ...)

96 kHz (DVD)

ENSEIRB 2017

Electronique TS200

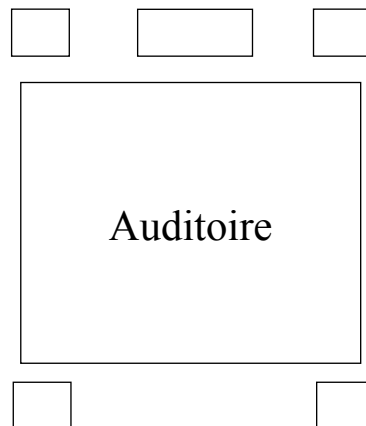
Dolby prologic AC3

- Haut parleurs (5)

- Gauche et droite (2)
- Devant et derrière (2)
- Renforcement des basses (1)

- Résolution

16 bits - 20 bits



ENSEIRB 2017

Electronique TS200

Bilan numérique

Signal vidéo : format 4:2:2 (8 bits)

Débit brut = $13.5 * 8 + 2 * 6.75 * 8 = \mathbf{216 \text{ Mbits/s}}$

Signal audio : 5 voies à 16 bits

Débit brut = $44.1 * 5 * 16 = \mathbf{3.6 \text{ Mbits/s}}$

Position du problème

- ▶ Débit brut d'une vidéo couleur (qualité studio)
 - Signal vidéo : format 4:2:2 (8 bits) **216 Mbits/s**
 - Signal audio : 5 voies à 44,1Hz 16 bits **3.6 Mbits/s**
- ▶ Lecture/écriture sur stockage optique
 - Cdrom 74 mm, 1,5 Mbits/s
 - Télé MPEG2 4~5 Mbits/s
 - DVD 133 mm, 15 Mbits/s
 - DV 270 mm, 5:1, 25 Mbits/s
- ▶ Transport
 - Réseau GSM 9,6 kbits/s
 - Modem V.90 56 kbits/s
 - Réseau Ethernet standard 10 Mbits/s
 - USB (v1) 1.5 Mbytes/s (v2) 480 Mbits/s
 - IEEE-1394 Firewire 800, 1600 et 3200 Mbits/s

Plan du cours

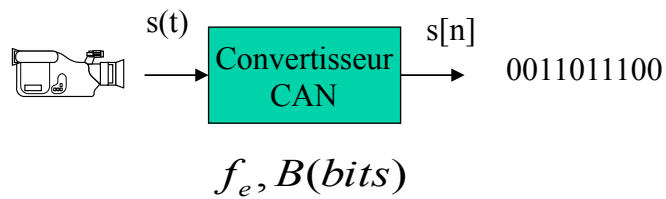
- ▶ Introduction
 - Formation des images
 - Acquisition/Représentation des images
 - Signaux vidéo
- ▶ Compression d'images fixes
 - Méthodes de codage source
 - Norme Fax
 - Norme JPEG
- ▶ Compression de vidéos
 - Estimation du mouvement
 - Normes MPEG 1 et 2, H261, H263
- ▶ Pour l'an prochain
 - Contenus multimédias interactifs
 - Norme MPEG 4
 - Transports
 - Streaming
 - Protocole IP

Objectifs

- ▶ Minimiser la bande passante ou la taille de stockage
 - Augmenter le nombre de programmes.
 - Augmenter la qualité de service.
 - Augmenter le nombre d'abonnés
- ▶ Cette minimisation peut être faite sans ou avec perte d'information
 - Secteur d'activité (biomédical, grand public, militaire ...)
 - Qualité de service (MPEG 4 avec le streaming)

Compression?

► Système numérique : Quantification/Echantillonnage



$$D_{brut} = f_e B$$

Où chercher ?

► Théorie de l'information

Source notée	S
Dictionnaire	D
Code noté	C_i

$$S = \{s_i\}$$
$$D = \{C_i\}_{i=1:N}$$

► Théorie du signal

- Fourier
- Corrélation
- Filtrage ...

Systeme numerique

► Formats binaires

► Code ASCII (1963) 5bits+2bits de parité

- 1 : 011 0001
- A : 100 0001
- r : 111 0010
- + : 010 1011

► Code EBCDIC (1975) 8 bits

- 1 : 1111 0001
- A : 1100 0001
- r : 1001 1001
- + : 0100 1110

SOS
10000011 01111001 10000011

Source aleatoire

► Problème probabiliste

► Source discrète

$$\mathbf{A}_n = \{x_i\}_{i=1:n}$$

$$\mathbf{X}_m = \{x_2, x_1, x_4, \dots, x_1\}$$

► Variable aléatoire discrète

$$p_{\mathbf{X}}(x) = \sum_{i=1}^n P(x_i) \delta(x - x_i)$$

$$P(x_i) = P_i$$

Théorie de l'information

► Estimation des probabilités

$$P(x_i) \approx \frac{\text{Card}(x_i)}{\text{Card}(X_m)} = \frac{m_i}{m}$$

► Messages

$$\mathbf{X}_4 = \{x_2, x_1, x_3, x_1\}$$

$$\mathbf{X}_4 = \{x_1, x_2, x_1, x_3\}$$

Quantité d'information

► Point de vue empirique

- Neige à Tahiti
- Neige à Chamonix

« *Moins* une observation est probable, *plus* son observation est porteuse d'information »

Quantité d'information

- I est une quantité positive
- I est inversement proportionnel à la probabilité de l'évènement

$$h(x) = -\log(P(x))$$

- Shannon

- Logarithme décimal : unité Hartley
- Logarithme base 2 : unité Shannon (bit/symbole)

Bit/symbole

- Exemple

$$\mathbf{X}_4 = \{x_2, x_1, x_3, x_1\}$$

$$P_2 = \frac{1}{4} \Rightarrow h_2 = 2 \text{ bits/symbole}$$

Le logarithme à base 2 fait le lien entre la quantité d'information et le codage binaire du symbole en fonction de sa probabilité d'occurrence.

Entropie (indépendance de l'alphabet)

- Caractérisation d'une source

$$H(\mathbf{X}) = E(h(x)) = - \int_{-\infty}^{+\infty} p_{\mathbf{X}}(x) \log(p_{\mathbf{X}}(x)) dx$$

- Forme discrète

$$H(\mathbf{X}) = - \sum_{i=1}^n P_i \log(P_i)$$

Entropie d'un système

$$H(S) = E\{Q(C_i)\} = - \sum_{i=1}^N \frac{n_i}{M} \log_2 \left(\frac{n_i}{M} \right) \quad \text{Bits/symbole}$$



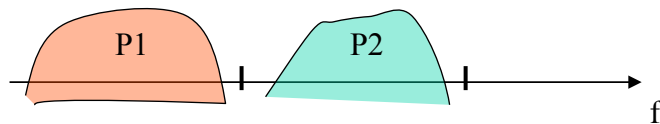
$H_1 = 1$ bit/symbole



$H_2 = 7.56$ bits/symbole

Compression?

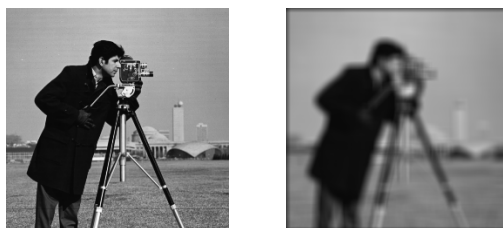
- Système analogique : Bande passante



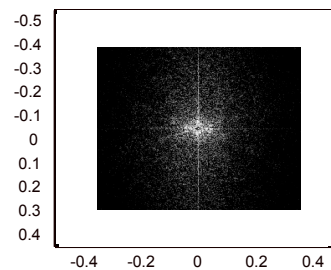
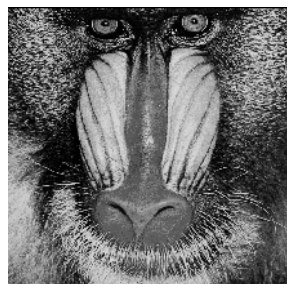
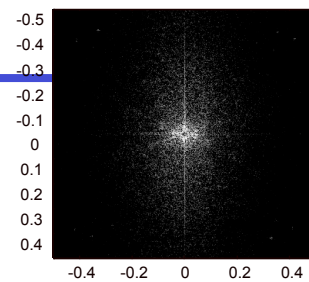
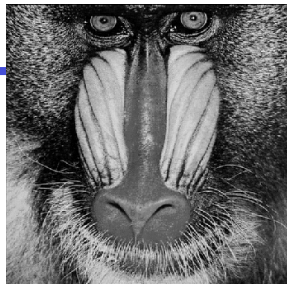
- Mesure de l'information : entropie bits/symbole

$$H(S) = E\{Q(C_i)\} = - \sum_{i=1}^N \frac{n_i}{M} \log_2 \left(\frac{n_i}{M} \right)$$

Passe-bas



Trop simpliste : mauvaise idée

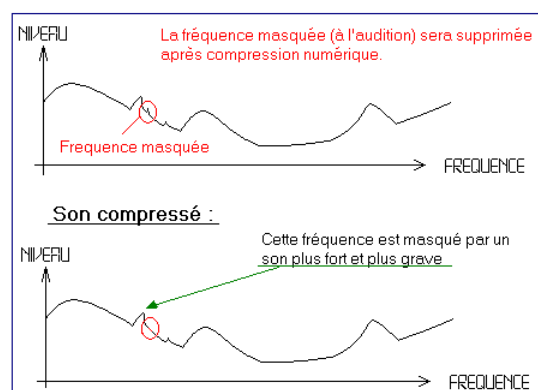


ENSEIRB 2017

Electronique TS200

Adaptation

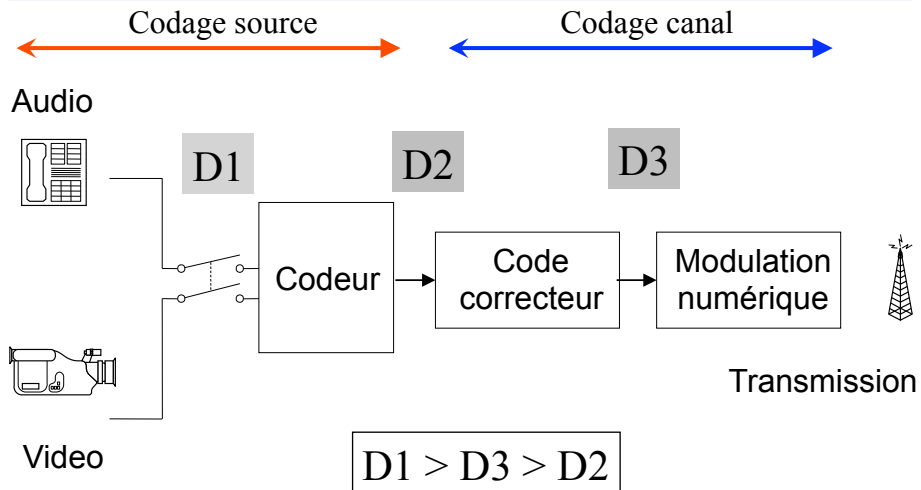
► MP3



ENSEIRB 2017

Electronique TS200

Synthèse transmission



ENSEIRB 2017

Electronique TS200

Contexte

résolution	format	Débit en Ko/s	Minutes sur CD
352x288	YUV	5 Mos/s	2
	MJPEG	768 Ko/s	14
	MPEG 1	175 Ko/s	60
	DIVX	114 Ko/s	95
640x480	YUV	15 Mos/s	40 secondes
	MJPEG	2.3 Mo/s	4,5
	MPEG 2	490 Ko/s	22
	DIVX	352 Ko/s	44
	DV	3.1 Mo/s	3.3

ENSEIRB 2017

Electronique TS200

II Rappels et compléments sur le codage source

II – 1 Rappels sur le codage source

ENSEIRB 2017

Electronique TS200

Formulation du problème

Image à coder

R	R	O	Y
R	O	O	Y
O	O	Y	G
Y	Y	Y	G

► Chaîne de symboles à compresser:

RR0YR00Y00YGYYYG

ENSEIRB 2017

Electronique TS200

Information codée

Séquence d'événements parmi un choix fini de possibilités (un alphabet).

Source notée	S
Dictionnaire	D
Code noté	C_i

$$S = \{s_i\}$$

$$D = \{C_i\}_{i=1:N}$$

Entropie d'un système

$$H(S) = E\{Q(C_i)\} = -\sum_{i=1}^N \frac{n_i}{M} \log_2 \left(\frac{n_i}{M} \right) \quad \text{Bits/symbole}$$



$H_1 = 1$ bit/symbole



$H_2 = 7.56$ bits/symbole

Longueur et efficacité d'un code

$$L_{moy}(S) = \sum_{i=1}^N p(C_i) l_i$$

Bits/symbole

Théorème du codage

$$L_{moy}(S) \geq H(S)$$

$$\eta = \frac{H(S)}{L_{moy}(S)}$$

II Rappels et compléments sur le codage source

II – 2 Un codeur intuitif

Codage des répétitions (RLC ou RLE)

- Choix d' un caractère de contrôle
- Coder les plages de k octets identiques

Exemple :

I= 01abbbbbbcZeed

I= 01a#6bcZeed

Problème : trouver un caractère de contrôle non élément du dictionnaire.

II Rappels et compléments sur le codage source

II – 3 Rappels sur le codage entropique

Code à longueur variable

Image à coder

R	R	O	Y
O	O	O	Y
O	Y	Y	G
Y	Y	Y	G

RRORR00Y00YGYYYG



0010010110101011000111000

Couleur	Code
R	001
O	01
Y	1
G	000

ENSEIRB 2017

Electronique TS200

Condition du préfixe

Aucun mot de code ne doit être un préfixe d'un autre mot de code.

=> Une seule interprétation possible

Exemple

Symboles :	x1	x2	x3	x4	x5	x6
Codes :	0	100	101	110	1110	1111

11001110101110 → 110 – 0 – 0 – 1110 – 101 – 110

ENSEIRB 2017

Electronique TS200

Codage de Huffman

Codage à arbre binaire

Idée : générer des codes de longueur variable

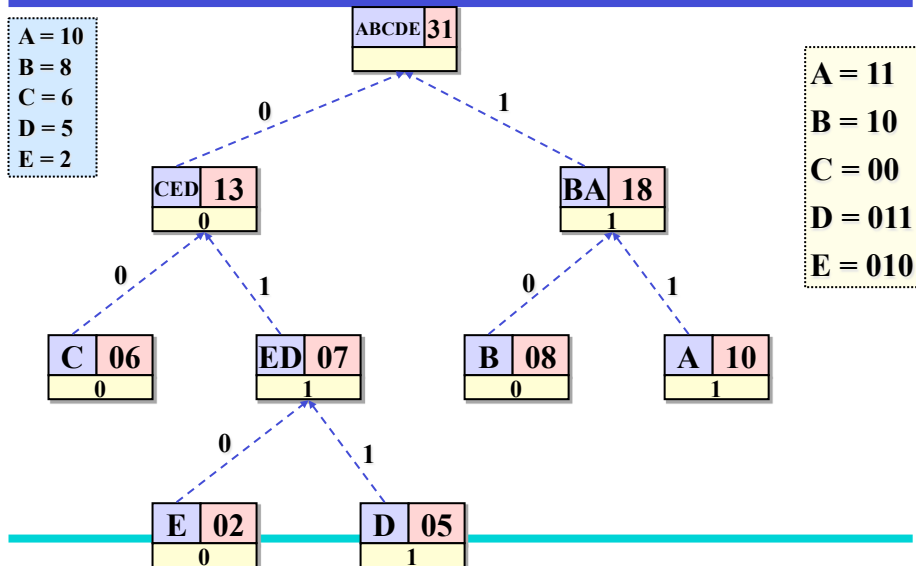
$p_i \Rightarrow 1$ $L(C_i)$ faible

$p_i \Rightarrow 0$ $L(C_i)$ grande

Méthodologie

- 1 - Classer les fréquences par ordre décroissant
- 2 - Deux noeuds libres de plus faible poids sont sélectionnés. Un noeud parent est alors créé. Son poids est la somme des poids des noeuds fils.
- 3 - Attribuer le bit 1 au fils gauche puis le bit 0 au fils droit.
- 4 - Itérer la construction de l'arbre jusqu'à ce qu'il ne reste plus qu'un seul noeud (de poids 1) $\Rightarrow \{$.

Exemple de Huffman



ENSEIRB 2017

Electronique TS200

Conclusion codage de Huffman

- Compression générique
- Efficace si probabilités d'apparition variées
 - Peu efficace sur les images ayant des histogrammes à distribution uniforme.
- Peu efficace sur les séquences répétitives
- Existe en version adaptative
 - Emission du code correspondant au code entrant
 - Mise à jour de l'arbre après chaque lecture

ENSEIRB 2017

Electronique TS200

Codage sans entête

- Table statistique prédéfinie: densité de probabilité moyenne représentative de l'application.

« *the text is written in english* »

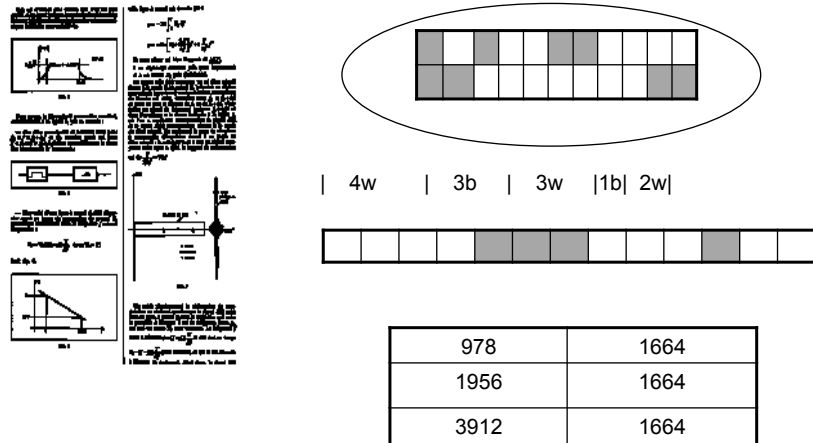
Par table :	111 bits		Entête < 27 bits
Par Huffman :	84 bits		

II Rappels et compléments sur le codage source

II – 4 Application au fax du codage entropique



Fax - Norme en télécommunications



ENSEIRB 2017

Electronique TS200

ITU-T

❖ T. 4 (Group 3 Standard)

- Developed in 1980
- Amended in 1984 and 1988
- Fax Transmission in PSTN (Public Switched Telephone Network) and International Leased Circuits

❖ T. 6 (Group 4 Standard)

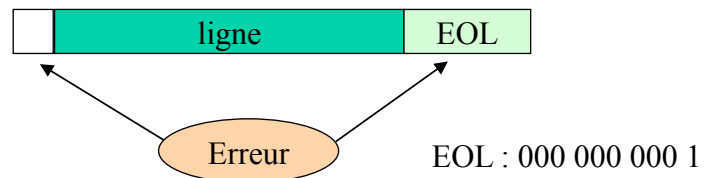
- Developed in 1988
- Error Free Digital fax transmission on Public Data networks.
- Also used in Document Imaging for Compression of Binary Images.

ENSEIRB 2017

Electronique TS200

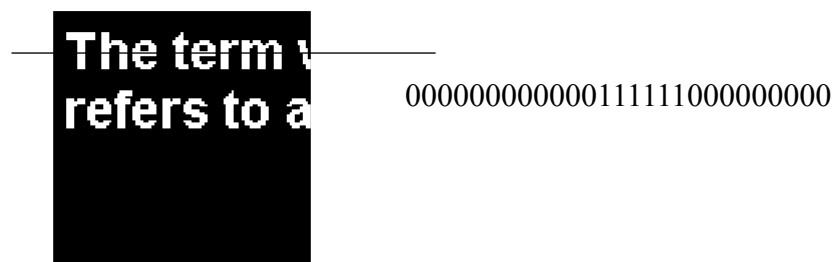
Process 1D RLE+Huffman

- ❖ Encodage pour chaque ligne + EOL (synchro.)



ligne	Pixel Blanc	EOL
page	1 EOL	6 EOL

Exemple



Mode MH

Run Length	Encoding
0-63	Use Codeword from 'T' table
64-1791	Use codewords from 'M' Table and 'T' Table M : Make up
1792-2560	Use codewords from 'AM' Table and 'T' Table

Projet de recommandation du CCITT T.4 RFC 804

ENSEIRB 2017

Electronique TS200

Exemples

- ▶ 12 pixels blancs (Terminating)

001000

- ▶ 140 pixels blancs : 128+12 (Make-up et Terminating)

10010|001000

ENSEIRB 2017

Electronique TS200

II Rappels et compléments sur le codage source

II – 5 Codage à dictionnaire

ENSEIRB 2017

Electronique TS200

Dictionnaire

- ▶ Transmission des adresses dans un dictionnaire à la place des symboles eux-mêmes
- ▶ Création du dictionnaire à la volée
 - Le décodeur peut reconstruire le dictionnaire
- ▶ Compactage sans lecture préalable du fichier

Dictionnaire	Adresse
NULL	0
...	...
a	97
b	98
...	...
si	256
ir	257
sid	258
ma	259
...	...

ENSEIRB 2017

Electronique TS200

LZW

```

s = premier octet du fichier;
tant que le fichier n' est pas fini :
{
    t = octet suivant;
    u = s  $\oplus$  t;
    si u  $\in$  D, alors faire s = u;
    sinon émettre adresse (s);
    ajouter u dans D;
    faire s = t;
}
émettre adresse (s);
  
```

ENSEIRB 2017

Electronique TS200

Fichier	émission	binaire	Dictionnaire	
10	rien			
130	10	00001010	256	10 130
130	130	10000010	257	130 130
121	130	10000010	258	130 121
10	121	01111001	259	121 10
10	10	00001010	260	10 10
10	SP	11111111	261	10 10 10
10	260	100000100	262	10 130 121
130			263	121 130
121	256	100000000		
130	121	001111001		
121				
	258	100000010		

ENSEIRB 2017

Electronique TS200

Décompression

a = premier caractère du fichier (8 bits)
émettre (a);
tant que le fichier n' est pas fini :
{
 b = lire adresse suivante;
 si $b \notin \text{adresse}(D)$, $s = a \oplus t$;
 si $b \in \text{adresse}(D)$, faire $s = \text{code}(b)$;
 émettre (s);
 t = premier caractère de s;
 ajouter $a \oplus t$ dans D;
 faire $a = s$;
}

Exemples

► Compress (UNIX) base LZW

- Dictionnaire init ($2^9=512$ entrées jusqu' à 2^{16})
- Saturation = dictionnaire statique

► ZIP et GZIP : LZ77+Huffman

- fenêtre 32Kbytes-long

► PKZIP : LZW et Shannon-Fano

- Dictionnaire dynamique

II Rappels et compléments sur le codage source

II – 6 Bilan et perspectives

ENSEIRB 2017

Electronique TS200

Exemples

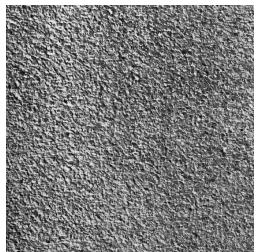


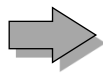
Image	BMP	TIF + LZW
Usine (Ko)	386	174
Route (Ko)	402	411

ENSEIRB 2017

Electronique TS200

Conclusion générale

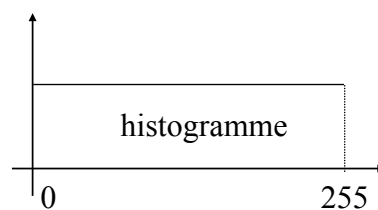
- RLC : Compression des plages de même niveau.
- Statistique (Huffman, arithmétique) : joue sur les disparités des fréquences d'apparition.
- Dictionnaire : recherche les répétitions de chaînes.



Aucune des méthodes ne traite des corrélations existantes entre les blocs de pixels.

Performances de méthodes statistiques

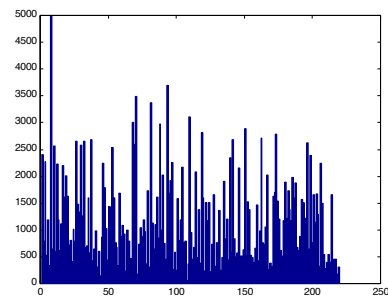
T_{Huff}



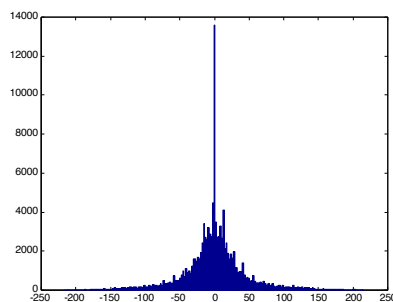
Observation sur une corrélation entre pixels voisins

$$\text{Err}(i) = x(i+1, j) - x(i, j)$$

Entropie(x) = 7.56

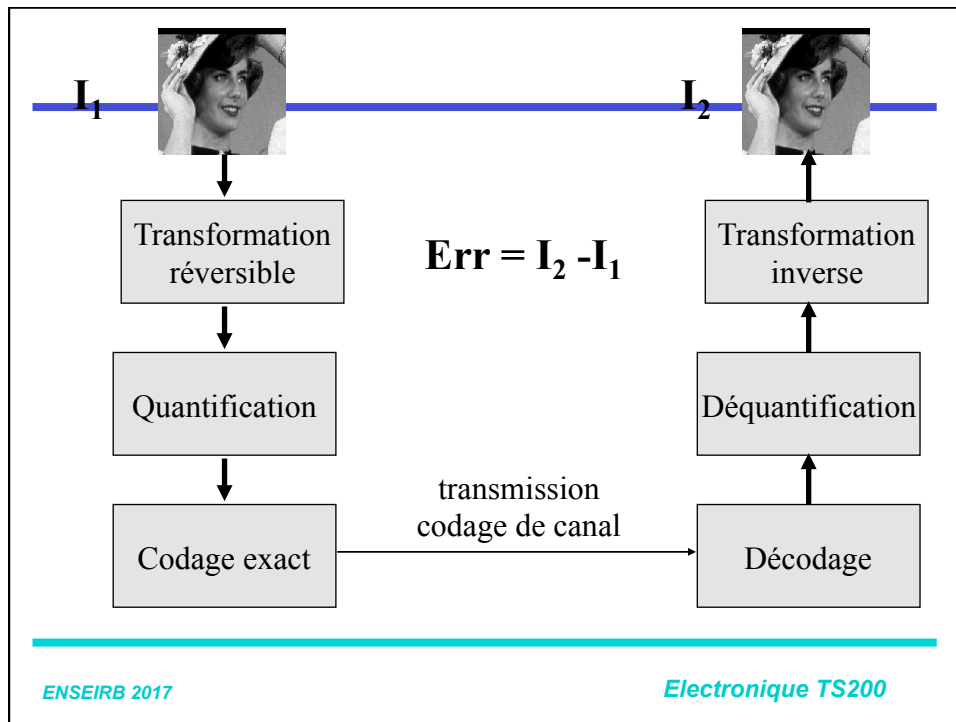


Entropie(err) = 5.02



Objectifs

- Trouver une transformation
 - Décorrélation.
 - Bonne répartition de l'énergie des nouveaux coefficients.
 - Peu coûteuse en temps de calcul
- Perte d'information
 - Troncature des coefficients de la transformée
 - Quantification scalaire ou multidimensionnelle



Extraction de la corrélation par transformée orthogonale

$$x(t) = \sum_{k=-\infty}^{+\infty} a_k \Phi_k(t) \quad a_k = \int_a^b x(t) \Phi_k^*(t) dt$$

$$\int_a^b \Phi_k(t) \Phi_l^*(t) dt = \begin{cases} 1 & \text{si } k = l \\ 0 & \text{si } k \neq l \end{cases}$$

Formes discrètes

$$x(n) = \sum_{k=0}^{N-1} a_k \Phi_k(n) \quad \Rightarrow \quad \underline{x} = \Phi \underline{a}$$

$$a_k = \sum_{n=0}^{N-1} x(n) \Phi_k^*(n) \quad \Rightarrow \quad \underline{a} = \Phi^H \underline{x}$$

\mathbf{H} : symbolise la matrice transposée et conjuguée

$$\text{orthogonalité} \quad \Phi^H \Phi = I_N$$

Objectifs de la compression

$$a_k \quad \text{pour } k \in [0, N-1]$$

$$x(n) \quad \text{pour } n \in [0, N-1]$$

- a_k décorrélés
- a_k fonction fortement décroissante

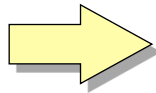
Troncature sur $\{a_k\}$ en contrôlant l'erreur de reconstruction

$$e(n) = x(n) - \tilde{x}(n) = x(n) - \sum_{k=0}^{M-1} a_k \Phi_k(n)$$

$$M < N$$

Principe psychovisuel

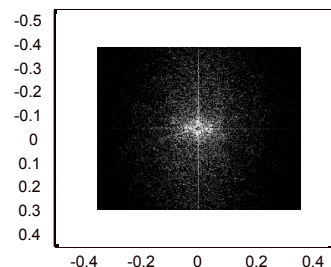
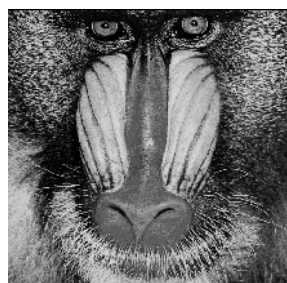
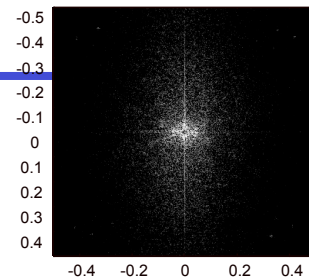
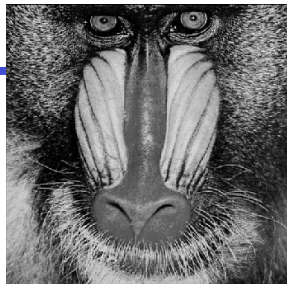
La perception visuelle
peu sensible à la suppression
des composantes hautes fréquences



Destruction de l'information haute fréquence

ENSEIRB 2017

Electronique TS200



ENSEIRB 2017

Electronique TS200